ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΑΓΡΟΝΟΜΩΝ ΚΑΙ ΤΟΠΟΓΡΑΦΩΝ ΜΗΧΑΝΙΚΩΝ
ΤΟΜΕΑΣ ΤΟΠΟΓΡΑΦΙΑΣ
ΕΡΓΑΣΤΗΡΙΟ ΦΩΤΟΓΡΑΜΜΕΤΡΙΑΣ

ΑΝΑΠΤΥΞΗ ΕΦΑΡΜΟΓΩΝ ΕΠΑΥΞΗΜΕΝΗΣ ΠΡΑΓΜΑΤΙΚΟΤΗΤΑΣ ΒΑΣΕΙ ΕΠΙΠΕΔΟΥ ΠΡΟΤΥΠΟΥ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΣΤΥΛΙΑΝΗ ΒΕΡΥΚΟΚΟΥ

ΑΘΗΝΑ, ΟΚΤΩΒΡΙΟΣ 2013

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΑΓΡΟΝΟΜΩΝ ΚΑΙ ΤΟΠΟΓΡΑΦΩΝ ΜΗΧΑΝΙΚΩΝ
ΤΟΜΕΑΣ ΤΟΠΟΓΡΑΦΙΑΣ
ΕΡΓΑΣΤΗΡΙΟ ΦΩΤΟΓΡΑΜΜΕΤΡΙΑΣ

Ανάπτυξη Εφαρμογών Επαυξημένης Πραγματικότητας Βάσει Επίπεδου Προτύπου

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

της

Στυλιανής Α. Βερυκόκου

Επιβλέπων Καθηγητής:

Χαράλαμπος Ιωαννίδης, Αναπληρωτής Καθηγητής Ε.Μ.Π.

Τριμελής Εξεταστική Επιτροπή:

Ανδρέας Γεωργόπουλος, Καθηγητής Ε.Μ.Π.

Χαράλαμπος Ιωαννίδης, Αναπληρωτής Καθηγητής Ε.Μ.Π.

Βασίλειος Βεσκούκης, Επίκουρος Καθηγητής Ε.Μ.Π.

Αθήνα, Οκτώβριος 2013

Στυλιανή Βερυκόκου
Διπλωματούχος Αγρονόμος και Τοπογράφος Μηχανικός Ε.Μ.Π.

Copyright @ Στυλιανή Βερυκόκου, 2013
Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τη συγγραφέα.

Στους γονείς μου, Ανδρέα και Τασούλα

ΠΡΟΛΟΓΟΣ

Η παρούσα διπλωματική εργασία ανατέθηκε από το Εργαστήριο Φωτογραμμετρίας της Σχολής Αγρονόμων και Τοπογράφων Μηχανικών του Εθνικού Μετσόβιου Πολυτεχνείου, με σκοπό την ανάπτυξη εφαρμογών επαυξημένης πραγματικότητας βάσει επίπεδου προτύπου, με χρήση αλγορίθμων και μεθόδων της φωτογραμμετρίας και της όρασης υπολογιστών.

Η επιστήμη του Αγρονόμου και Τοπογράφου Μηχανικού συντίθεται από ένα ευρύ φάσμα επιστημονικών πεδίων, όπως είναι η γεωδαισία, η φωτογραμμετρία, η φωτοερμηνεία - τηλεπισκόπηση, η χαρτογραφία, η επιστήμη της γεωγραφικής πληροφορίας και τα συστήματα γεωγραφικών πληροφοριών, το κτηματολόγιο, η ανάλυση, ο σχεδιασμός και η οργάνωση του πολυδιάστατου χώρου, καθώς και η μελέτη συγκοινωνιακών, υδραυλικών και τεχνικών έργων. Τα τελευταία χρόνια, λόγω της ραγδαίας εξέλιξης της τεχνολογίας, όλα τα παραδοσιακά αντικείμενα της Σχολής έχουν λάβει νέα διάσταση και έχουν αποκτήσει νέες δυνατότητες. Έτσι, θεωρώ ότι ο σύγχρονος Αγρονόμος και Τοπογράφος Μηχανικός δεν πρέπει να παραμένει στάσιμος στις ευρέως εφαρμοσμένες τεχνικές που ήδη γνωρίζει, αλλά οφείλει να είναι ανοικτός στις νέες τεχνολογίες.

Στο πλαίσιο αυτό, ανέλαβα την εκπόνηση της συγκεκριμένης διπλωματικής εργασίας, ώστε να ασχοληθώ με ένα νέο επιστημονικό πεδίο, αυτό της επαυξημένης πραγματικότητας, και να διερευνήσω τα αποτελέσματα που μπορούν να επιτευχθούν μέσω αλγορίθμων, τεχνικών και σχέσεων της φωτογραμμετρίας και του συγγενούς επιστημονικού κλάδου της όρασης υπολογιστών. Με την ενασχόλησή μου με το παρόν θέμα, επεδίωξα να επεκτείνω τις γνώσεις μου στη φωτογραμμετρία, να εξετάσω τη σχέση της με την όραση υπολογιστών, να χρησιμοποιήσω σε πρακτικό επίπεδο αλλά και να εμπλουτίσω τις γνώσεις μου στον προγραμματισμό σε C++ και τέλος να διερευνήσω τις δυνατότητες της βιβλιοθήκης OpenCV και της διεπαφής προγραμματισμού εφαρμογών OpenGL και να εξοικειωθώ με αυτές, μέσω της χρήσης τους στην ανάπτυξη των εφαρμογών. Παράλληλα, μου δόθηκε η ευκαιρία να ασχοληθώ με τις μεγάλες δυνατότητες της επαυξημένης πραγματικότητας, όχι μόνο σε φωτογραμμετρικές εφαρμογές, αλλά και σε εφαρμογές γενικού ενδιαφέροντος.

ΕΥΧΑΡΙΣΤΙΕΣ

Έχοντας ολοκληρώσει την παρούσα διπλωματική εργασία, αισθάνομαι την ανάγκη να ευχαριστήσω όλους όσους συνέβαλαν με οποιοδήποτε τρόπο στην πραγματοποίησή της. Τις θερμότατες ευχαριστίες μου οφείλω στον επιβλέποντα Καθηγητή μου, κ. Χαράλαμπο Ιωαννίδη, για την εμπιστοσύνη που μου έδειξε με την ανάθεση του συγκεκριμένου θέματος, για το χρόνο που διέθεσε απλόχερα, την προθυμία του να με βοηθήσει σε οποιοδήποτε πρόβλημα αντιμετώπιζα κατά τη διάρκεια εκπόνησης της εργασίας, τις πολύτιμες συμβουλές του, τις εύστοχες παρατηρήσεις του και για το άριστο κλίμα που χαρακτήριζε τη συνεργασία μας.

Επίσης, αισθάνομαι την ανάγκη να ευχαριστήσω τους Καθηγητές μου στα μαθήματα της Φωτογραμμετρίας, κ. Γιώργο Καρρά και κ. Χαράλαμπο Ιωαννίδη, καθώς μέσω των παραδόσεων των μαθημάτων μου κέντρισαν το ενδιαφέρον για το αντικείμενο της Φωτογραμμετρίας και μου μεταλαμπάδευσαν πολύτιμες γνώσεις που στάθηκαν καθοριστικές για την ολοκλήρωση της παρούσας εργασίας.

Ακόμη, θα ήθελα να ευχαριστήσω τον κ. Νίκο Δουλάμη, Επίκουρο Καθηγητή Ε.Μ.Π., για την προτροπή του να ασχοληθώ με τον προγραμματισμό σε C++ και με τη βιβλιοθήκη OpenCV και για την προθυμία του να με βοηθήσει σε τεχνικά θέματα που αντιμετώπισα κατά τη διάρκεια της διπλωματικής εργασίας. Επίσης, ευχαριστώ την κα. Σοφία Σοϊλέ, μέλος του Εργαστηρίου Φωτογραμμετρίας, για το χρόνο που διέθεσε για την παροχή στοιχείων αναγκαίων για την ολοκλήρωση της εργασίας.

Το μεγαλύτερο ευχαριστώ το οφείλω στους γονείς μου, οι οποίοι υπήρξαν ανεκτίμητο στήριγμα για μένα, σε κάθε στάδιο της ζωής μου, και στους οποίους οφείλω όλη τη διαδρομή των σπουδών μου μέχρι σήμερα. Η πίστη τους στις δυνατότητές μου, η στήριξη όλων των επιλογών μου, αλλά και η αμέριστη συμπαράστασή τους κατά τη διάρκεια εκπόνησης της παρούσας εργασίας συνέβαλαν σε μεγάλο βαθμό στην επιτυχή διεκπεραίωσή της.

Τέλος, θα ήθελα να ευχαριστήσω και την υπόλοιπη οικογένειά μου, που είναι δίπλα μου όλα αυτά τα χρόνια, καθώς επίσης και τους φίλους και τις φίλες μου για τη στήριξη και την παρουσία τους, η οποία ομορφαίνει τη ζωή μου.

Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Αγρονόμων και Τοπογράφων Μηχανικών
Τομέας Τοπογραφίας
Εργαστήριο Φωτογραμμετρίας

ΑΝΑΠΤΥΞΗ ΕΦΑΡΜΟΓΩΝ ΕΠΑΥΞΗΜΕΝΗΣ ΠΡΑΓΜΑΤΙΚΟΤΗΤΑΣ ΒΑΣΕΙ ΕΠΙΠΕΔΟΥ ΠΡΟΤΥΠΟΥ

Στυλιανή Βερυκόκου

Διπλωματική Εργασία
Οκτώβριος 2013

Περίληψη

Η επαυξημένη πραγματικότητα είναι μία ραγδαίως εξελισσόμενη τεχνολογία που εμπλουτίζει την πραγματικότητα με πληροφορίες που παράγονται από έναν υπολογιστή. Μολονότι η έννοια αυτή συναντάται για πρώτη φορά τη δεκαετία του 1960, τα τελευταία χρόνια έχει αρχίσει να παρουσιάζει πρακτικές εφαρμογές σε διάφορους τομείς και σήμερα κεντρίζει το ενδιαφέρον πολλών ερευνητών, επιστημόνων αλλά και εταιριών.

Σκοπός της συγκεκριμένης διπλωματικής εργασίας είναι η ανάπτυξη εφαρμογών επαυξημένης πραγματικότητας βάσει επίπεδου προτύπου, εσωτερικού και εξωτερικού χώρου, με τη χρήση αλγορίθμων και μεθόδων της φωτογραμμετρίας και της όρασης υπολογιστών. Οι εφαρμογές αυτές αναπτύχθηκαν στη γλώσσα προγραμματισμού C++, με χρήση της βιβλιοθήκης OpenCV, της διεπαφής προγραμματισμού εφαρμογών OpenGL και της βιβλιοθήκης GLM: An Alias Wavefront OBJ file Library και είναι εφαρμογές κονσόλας Win32, που εκτελούνται σε ηλεκτρονικούς υπολογιστές με λειτουργικό σύστημα Microsoft Windows. Στην παρούσα εργασία, γίνεται μία εισαγωγή στην έννοια της επαυξημένης πραγματικότητας, παρουσιάζεται το θεωρητικό υπόβαθρο που είναι απαραίτητο για την κατανόηση της διαδικασίας που ακολουθήθηκε στις εφαρμογές και αναφέρονται ορισμένα βασικά στοιχεία για τη γλώσσα προγραμματισμού, τις βιβλιοθήκες που χρησιμοποιήθηκαν και το περιβάλλον εργασίας. Ακολουθεί η περιγραφή των εφαρμογών και του τρόπου υλοποίησής τους, η παρουσίαση των αποτελεσμάτων που επιτεύχθηκαν και η παράθεση κάποιων προτάσεων που χρήζουν περαιτέρω διερεύνησης για τη μελλοντική επέκταση και βελτίωση των εφαρμογών.

Λέξεις Κλειδιά: Επαυξημένη Πραγματικότητα, Φωτογραμμετρία, Όραση Υπολογιστών, SURF, RANSAC, Ομογραφία, Ομογενείς Συντεταγμένες, Βαθμονόμηση, OpenCV, OpenGL, OBJ

National Technical University of Athens, Greece
School of Rural and Surveying Engineering
Department of Surveying
Laboratory of Photogrammetry

DEVELOPMENT OF PLANAR PATTERN BASED AUGMENTED REALITY APPLICATIONS

Styliani Verykokou

Diploma Thesis
October 2013

Abstract

Augmented reality is a rapidly evolving technology that enriches reality with computer generated information. Although it was first presented in the 1960s, actually in recent years it has begun to have practical applications in various fields and today arouses the interest of many researchers, scientists, as well as companies.

The purpose of this diploma thesis is the development of indoor and outdoor planar pattern based augmented reality applications, through the use of methods and algorithms of photogrammetry and computer vision. These applications were developed in C++ programming language, with the use of the OpenCV library, the OpenGL application programming interface and the library GLM: An Alias Wavefront OBJ file Library and they are Win32 console applications, intended for computers running Microsoft Windows. In this thesis, an introduction is made concerning the concept of augmented reality, followed by a presentation of the theoretical background, necessary to understand the procedure that was held during the applications. Furthermore, certain basic elements of information are provided, referring to the programming language, the libraries that have been used and the development environment. Attending is the description of the applications, of the way to materialize them, as well as the presentation of the results obtained and the formation of some proposals that require further research for future development and improvement of the applications.

Key Words: Augmented Reality, Photogrammetry, Computer Vision, SURF, RANSAC, Homography, Homogeneous Coordinates, Calibration, OpenCV, OpenGL, OBJ

ΠΕΡΙΕΧΟΜΕΝΑ

ΠΡΟΛΟΓΟΣ I

ΕΥΧΑΡΙΣΤΙΕΣ III

ΠΕΡΙΛΗΨΗ V

ABSTRACT VII

ΚΑΤΑΛΟΓΟΣ ΕΙΚΟΝΩΝ XV

ΚΑΤΑΛΟΓΟΣ ΔΙΑΓΡΑΜΜΑΤΙΚΩΝ ΠΑΡΟΥΣΙΑΣΕΩΝ XXVII

ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ XXIX

ΕΙΣΑΓΩΓΗ 1

ΚΕΦΑΛΑΙΟ 1: ΕΠΑΥΞΗΜΕΝΗ ΠΡΑΓΜΑΤΙΚΟΤΗΤΑ 5

1.1 Γενική Περιγραφή 5

1.2 Ιστορική Αναδρομή 7

1.3 Εφαρμογές 15

1.3.1 Ιατρική 15

1.3.2 Εκπαίδευση 16

1.3.3 Στρατός 17

1.3.4 Ψυχαγωγία - ενημέρωση - αθλητισμός 18

1.3.5 Τέχνη - πολιτισμός - τουρισμός 19

1.3.6 Πλοήγηση 21

1.3.7 Εμπόριο και διαφήμιση 23

1.3.8 Αρχιτεκτονική - Σχεδιασμός εσωτερικού χώρου 23

1.3.9 Υποστήριξη εργασιών 24

1.3.10 Άλλες εφαρμογές 24

1.4 Τεχνολογίες Θέασης της Επαυξημένης Σκηνής 25
1.4.1 Φορετές συσκευές 25
1.4.1.1 Προσαρτημένα επί κεφαλής (ή επί κράνους) συστήματα παρουσίασης 25
1.4.1.2 Συσκευές απεικόνισης στον αμφιβληστροειδή του ματιού (VRDs) 27
1.4.1.3 Ειδικά γυαλιά 27
1.4.1.4 Ειδικοί φακοί επαφής 28
1.4.2 Φορητές συσκευές χειρός 29
1.4.3 Χωρικές συσκευές απεικόνισης 30
1.4.3.1 Συσκευές απεικόνισης σε οθόνη με χρήση βίντεο τεχνολογίας 30
1.4.3.2 Χωρικές συσκευές απεικόνισης με χρήση οπτικής τεχνολογίας 30
1.4.3.3 Χωρικές συσκευές απεικόνισης μέσω προβολής 31
1.5 Μέθοδοι Επαύξησης της Πραγματικότητας 32
1.5.1 Επαυξημένη πραγματικότητα βάσει προτύπου 32
1.5.2 Επαυξημένη πραγματικότητα βάσει περιγράμματος 33
1.5.3 Επαυξημένη πραγματικότητα βάσει θέσης 33
1.5.4 Επαυξημένη πραγματικότητα σε επιφάνεια 33
1.6 Προκλήσεις και Προβλήματα Επαύξησης της Πραγματικότητας 34
ΚΕΦΑΛΑΙΟ 2: ΨΗΦΙΑΚΗ ΣΥΝΤΑΥΤΙΣΗ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ ΣΗΜΕΙΩΝ 37
2.1 Κατηγορίες Ψηφιακής Συνταύτισης 38
2.2 Ανίχνευση και Περιγραφή Σημείων Ενδιαφέροντος 40
2.2.1 Γενικές πληροφορίες και διάφοροι αλγόριθμοι 41
2.2.2 Αλγόριθμος SURF 42
2.2.2.1 Δομές blobs 43
2.2.2.2 Αθροιστικές εικόνες 44
2.2.2.3 Εσσιανός Πίνακας και παραλλαγή του που χρησιμοποιείται από τον αλγόριθμο 45
2.2.2.4 Αναπαράσταση του χώρου κλίμακας 47

2.2.2.5 Εντοπισμός των σημείων ενδιαφέροντος 50

2.2.2.6 Καθορισμός του κυρίαρχου προσανατολισμού κάθε σημείου ενδιαφέροντος 52

2.2.2.7 Εξαγωγή του διανύσματος περιγραφικών χαρακτηριστικών κάθε σημείου ενδιαφέροντος 54

2.2.2.8 Βήματα του αλγορίθμου SURF για τον προσδιορισμό των σημείων ενδιαφέροντος 56

2.3 Εύρεση των Ομόλογων Σημείων Ενδιαφέροντος 57

2.3.1 Κριτήρια ταύτισης των διανυσμάτων περιγραφικών χαρακτηριστικών 58

2.3.2 Εύρεση των ομόλογων σημείων ενδιαφέροντος από τον αλγόριθμο SURF ... 59

2.3.3 Απόρριψη μη έγκυρων ομολογιών 59

2.3.3.1 Λόγος των αποστάσεων των εγγύτερων γειτόνων (ratio test) 60

2.3.3.2 Διασταυρωμένος έλεγχος (cross-check test) 60

2.3.3.3 Αλγόριθμος RANSAC 61

ΚΕΦΑΛΑΙΟ 3: ΚΕΝΤΡΙΚΗ ΠΡΟΒΟΛΗ ΚΑΙ ΕΠΙΛΥΣΗ ΒΑΣΙΚΩΝ ΦΩΤΟΓΡΑΜΜΕΤΡΙΚΩΝ ΠΡΟΒΛΗΜΑΤΩΝ 65

3.1 Προβολική Γεωμετρία 65

3.2 Ομογενείς Συντεταγμένες 66

3.2.1 Ομογενείς συντεταγμένες σημείων στο προβολικό επίπεδο - θεωρητική προσέγγιση 66

3.2.2 Περιγραφή και πλεονεκτήματα των ομογενών συντεταγμένων 68

3.3 Κεντρική Προβολή 69

3.3.1 Μαθηματικό μοντέλο που χρησιμοποιείται στη φωτογραμμετρία 70

3.3.2 Μαθηματικό μοντέλο που χρησιμοποιείται στην όραση υπολογιστών 73

3.3.3 Σύγκριση συνθήκης συγγραμμικότητας και προβολικού μετασχηματισμού 79

3.4 Δισδιάστατοι Γραμμικοί Μετασχηματισμοί 80

3.4.1 Ιεραρχία των δισδιάστατων γραμμικών μετασχηματισμών 80

3.4.1.1 Μετασχηματισμός μετάθεσης 82

3.4.1.2 Ευκλείδειος μετασχηματισμός 82
3.4.1.3 Μετασχηματισμός ομοιότητας 83
3.4.1.4 Αφινικός μετασχηματισμός 84
3.4.1.5 Προβολικός μετασχηματισμός (ομογραφία) 85
3.4.2 Μετασχηματισμοί μεταξύ εικονοσημείων διαφόρων εικόνων ενός δισδιάστατου αντικειμένου 87
3.5 Βαθμονόμηση 90
3.5.1 Ανασκόπηση διαφόρων μεθόδων βαθμονόμησης 90
3.5.2 Βαθμονόμηση με λήψη εικόνων επίπεδης σκακιέρας 91
3.5.2.1 Αρχική επεξεργασία κάθε εικόνας 92
3.5.2.2 Έλεγχος εντοπισμού της σκακιέρας σε κάθε εικόνα 92
3.5.2.3 Ανίχνευση των εσωτερικών γωνιών της σκακιέρας σε κάθε εικόνα 98
3.5.2.4 Καθορισμός των συντεταγμένων χώρου των εσωτερικών γωνιών της σκακιέρας 100
3.5.2.5 Εκτίμηση προσεγγιστικών τιμών για τα στοιχεία του εσωτερικού προσανατολισμού 101
3.5.2.6 Υπολογισμός αρχικών τιμών για τα στοιχεία του εξωτερικού προσανατολισμού κάθε λήψης 107
3.5.2.7 Υπολογισμός αρχικών τιμών για τους συντελεστές των πολυωνύμων της διαστροφής των φακών 110
3.5.2.8 Ενιαία βελτιστοποίηση των στοιχείων του εσωτερικού και εξωτερικού προσανατολισμού 111
3.6 Προσδιορισμός των Στοιχείων του Εξωτερικού Προσανατολισμού 111
ΚΕΦΑΛΑΙΟ 4: ΠΡΟΓΡΑΜΜΑΤΙΣΤΙΚΟ ΠΕΡΙΒΑΛΛΟΝ ΕΡΓΑΣΙΑΣ - ΜΕΤΑΣΧΗΜΑΤΙΣΜΟΙ ΣΥΣΤΗΜΑΤΩΝ 115
4.1 Γλώσσα Προγραμματισμού C++ 115
4.1.1 Σύντομα ιστορικά στοιχεία 116
4.1.2 Βασικά χαρακτηριστικά της C++ 116

4.1.3 Επιλογή της C++ ως γλώσσας προγραμματισμού για την ανάπτυξη των εφαρμογών ... 117
4.2 Περιβάλλον Ανάπτυξης Προγραμμάτων C++ ... 118
4.2.1 Βασικά στοιχεία ενός τυπικού περιβάλλοντος ανάπτυξης C++ ... 118
4.2.2 Microsoft Visual C++ 2010 Express ... 119
4.3 OpenCV και OpenGL ... 121
4.3.1 Βιβλιοθήκη OpenCV ... 122
4.3.2 OpenGL και σχετικές βιβλιοθήκες ... 122
4.3.3 Συστήματα συντεταγμένων κάμερας των OpenCV και OpenGL ... 124
4.4 Μετασχηματισμοί στην OpenGL ... 125
4.4.1 Μετασχηματισμός μοντέλου - κάμερας ... 125
4.4.2 Μετασχηματισμός προβολής και κανονικοποίηση συντεταγμένων ... 129
4.4.2.1 Προοπτική προβολή ... 131
4.4.2.2 Ορθογραφική προβολή ... 139
4.4.2.3 Κανονικοποίηση συντεταγμένων ... 141
4.4.3 Μετασχηματισμός παραθύρου παρατήρησης και κωδικοποίηση της πληροφορίας βάθους ... 142
ΚΕΦΑΛΑΙΟ 5: ΕΦΑΡΜΟΓΕΣ ΕΠΑΥΞΗΜΕΝΗΣ ΠΡΑΓΜΑΤΙΚΟΤΗΤΑΣ ... 145
5.1 Γενική Μεθοδολογία ... 146
5.2 Βαθμονόμηση ... 152
5.3 Μορφή Αρχείων OBJ ... 156
5.4 Εφαρμογή 1 - Επαυξημένη Πραγματικότητα σε Εσωτερικό Χώρο ... 159
5.4.1 Αρχικά βήματα για τον καθορισμό των επιλογών του χρήστη ... 161
5.4.2 Ενδιάμεσα βήματα πριν την επαναληπτική διαδικασία επαύξησης της σκηνής κάθε στιγμιότυπου ... 165
5.4.3 Επαναληπτική διαδικασία επαύξησης της σκηνής κάθε στιγμιότυπου ... 176
5.4.3.1 Υπολογισμός του εξωτερικού προσανατολισμού του στιγμιότυπου ... 177
5.4.3.2 Εμφάνιση των κατάλληλων παραθύρων μέσω της OpenCV ... 182

5.4.3.3 Σχεδίαση του στιγμιότυπου ως φόντου της επιφάνειας σχεδίασης του παραθύρου επαύξησης της πραγματικότητας 187

5.4.3.4 Σχεδίαση του τρισδιάστατου μοντέλου στο παράθυρο επαύξησης της πραγματικότητας 192

5.5 Εφαρμογή 2 - Επαυξημένη Πραγματικότητα σε Εξωτερικό Χώρο 200

5.5.1 Αρχικά βήματα για τον καθορισμό των επιλογών του χρήστη 201

5.5.2 Ενδιάμεσα βήματα πριν την επαναληπτική διαδικασία επαύξησης της σκηνής κάθε στιγμιότυπου 207

5.5.3 Επαναληπτική διαδικασία επαύξησης της σκηνής κάθε στιγμιότυπου 213

5.5.3.1 Υπολογισμός του εξωτερικού προσανατολισμού του στιγμιότυπου 214

5.5.3.2 Εμφάνιση των κατάλληλων παραθύρων μέσω της OpenCV 216

5.5.3.3 Σχεδίαση της επαυξημένης σκηνής 221

ΣΥΜΠΕΡΑΣΜΑΤΑ 233

ΑΝΑΦΟΡΕΣ 237

ΠΑΡΑΡΤΗΜΑ Α: ΧΡΗΣΙΜΟΠΟΙΟΥΜΕΝΑ ΣΥΜΒΟΛΑ ΤΩΝ ΔΙΑΓΡΑΜΜΑΤΩΝ ΡΟΗΣ 257

ΠΑΡΑΡΤΗΜΑ Β: ΑΠΟΔΕΙΞΗ ΤΗΣ ΣΧΕΣΗΣ (3.39) 259

ΠΑΡΑΡΤΗΜΑ Γ: ΕΓΧΕΙΡΙΔΙΟ ΧΡΗΣΗΣ ΤΗΣ ΕΦΑΡΜΟΓΗΣ 1 261

Γ.1 Επιλογή της Κανονικής Εφαρμογής Επαυξημένης Πραγματικότητας 262

Γ.2 Εξέταση των Εναλλακτικών Επιλογών που Παρέχονται από την Εφαρμογή 264

Γ.2.1 Χρήση της κάμερας του υπολογιστή 264

Γ.2.2 Επαύξηση ενός αρχείου βίντεο 265

Γ.2.2.1 Εισαγωγή ενός αρχείου βίντεο από το χρήστη 266

Γ.2.2.2 Χρήση του προεπιλεγμένου αρχείου βίντεο 266

Γ.3 Επιπρόσθετες Διευκρινίσεις 267

ΠΑΡΑΡΤΗΜΑ Δ: ΕΓΧΕΙΡΙΔΙΟ ΧΡΗΣΗΣ ΤΗΣ ΕΦΑΡΜΟΓΗΣ 2 271

Δ.1 Παρακολούθηση Προεπισκόπησης της Εφαρμογής 271

Δ.2 Εισαγωγή Αρχείων από το Χρήστη 272

Δ.3 Επιπρόσθετες Διευκρινίσεις 275

ΚΑΤΑΛΟΓΟΣ ΕΙΚΟΝΩΝ

Εικόνα 1-1. Το Συνεχές Πραγματικότητας - Εικονικότητας (Reality - Virtuality Continuum) 7

Εικόνα 1-2. Το πρώτο σύστημα επαυξημένης πραγματικότητας από τον Ivan Sutherland (The Sword of Damocles) - Η συσκευή απεικόνισης του Sutherland που τοποθετείται στο κεφάλι του χρήστη 8

Εικόνα 1-3. Σκηνές από το Videoplace του Krueger 9

Εικόνα 1-4. Σύστημα επαυξημένης πραγματικότητας KARMA και εικόνα που βλέπει ο χρήστης μέσω της συσκευής θέασης 10

Εικόνα 1-5. Παραδείγματα χρήσεων του συστήματος NaviCam 11

Εικόνα 1-6. Παραδείγματα επίπεδων στόχων του 2D barcode συστήματος του Rekimoto 11

Εικόνα 1-7. Σύστημα «Map-in-hat» 11

Εικόνα 1-8. Παράδειγμα εικόνας που προβάλλεται στο χρήστη του παιχνιδιού ARQuake από το HMD, όπου φαίνεται τόσο ο πραγματικός κόσμος όσο και ο κόσμος που παράγεται από τον υπολογιστή 12

Εικόνα 1-9. Εφαρμογή MARA από τη Nokia 13

Εικόνα 1-10. Εφαρμογή Wikitude 14

Εικόνα 1-11. Παράδειγμα επαυξημένης πραγματικότητας στην ιατρική 16

Εικόνα 1-12. Η επαυξημένη πραγματικότητα δίνει τη δυνατότητα στο μαθητή να αλληλεπιδράσει με τη δομή ενός μορίου, το οποίο εμφανίζεται σε αυτόν σε μία εικόνα κάμερας όταν ειδικό marker που κρατά στο χέρι του εντοπίζεται από την κάμερα 17

Εικόνα 1-13. Παράδειγμα εφαρμογής επαυξημένης πραγματικότητας στον τομέα της εκπαίδευσης στην ιατρική 17

Εικόνα 1-14. Παράδειγμα Head-Up Display (HUD) 18

Εικόνα 1-15. Η τεχνολογία της επαυξημένης πραγματικότητας δημιουργεί την απεικονιζόμενη κίτρινη γραμμή σε γήπεδα ποδοσφαίρων, που φαίνεται μόνο στους τηλεοπτικούς δέκτες 19

Εικόνα 1-16. Η επαυξημένη πραγματικότητα στο Μουσείο Μοντέρνας Τέχνης της Νέας Υόρκης 20

Εικόνα 1-17. Ο χρήστης της εφαρμογής Konstruct δημιουργεί και βλέπει ένα έργο τέχνης στην οθόνη της κινητής συσκευής του σε περιβάλλον επαυξημένης πραγματικότητας 20

Εικόνα 1-18. Ερείπια του Ναού της Ήρας στη σημερινή τους κατάσταση και επαυξημένος Ναός με το τρισδιάστατο μοντέλο του να υπερτίθεται σε ζωντανό βίντεο 21

Εικόνα 1-19. Εικονικοί αθλητές συναγωνίζονται στο στάδιο της Αρχαίας Ολυμπίας 21

Εικόνα 1-20. Εικόνες από την εφαρμογή Wikitude Drive 22

Εικόνα 1-21. Η επαυξημένη πραγματικότητα σε εφαρμογή πλοήγησης σε εσωτερικό χώρο 22

Εικόνα 1-22. Επαυξημένη εικόνα από το «Digital Box» της Lego 23

Εικόνα 1-23. Εφαρμογή μετάφρασης επαυξημένης πραγματικότητας Word Lens 25

Εικόνα 1-24. Εννοιολογικό διάγραμμα οπτικού συστήματος HMD 26

Εικόνα 1-25. Εννοιολογικό διάγραμμα βίντεο συστήματος HMD 26

Εικόνα 1-26. Head-Mounted Display, Helmet-Mounted Display και Head-Mounted Projective Display 27

Εικόνα 1-27. Παραδείγματα γυαλιών επαυξημένης πραγματικότητας: EyeTap (2η έκδοση, 1999), Vuzix SMART Glasses και Google Glass 28

Εικόνα 1-28. Παράδειγμα εφαρμογής επαυξημένης πραγματικότητας μέσω συσκευής απεικόνισης σε οθόνη με χρήση βίντεο τεχνολογίας 30

Εικόνα 1-29. Παράδειγμα εφαρμογής επαυξημένης πραγματικότητας με χρήση οπτικής τεχνολογίας - διάφανη οθόνη 31

Εικόνα 1-30. Παράδειγμα χωρικής επαυξημένης πραγματικότητας με χρήση συσκευής απεικόνισης μέσω προβολής 31

Εικόνα 1-31. Παράδειγμα επαυξημένης πραγματικότητας βασισμένης σε markers (marker-based AR) και σε φυσικά χαρακτηριστικά του αντικειμένου (markerless AR) 32

Εικόνα 1-32. Παραδείγματα επαυξημένης πραγματικότητας σε επιφάνεια: SurfaceFusion και πάτωμα επαυξημένης πραγματικότητας που μιμείται την εμφάνιση και την αίσθηση της άμμου..... 34

Εικόνα 2-1. Δομές blobs..... 43

Εικόνα 2-2. Το άθροισμα των τιμών των εικονοστοιχείων του ορθογωνίου εντός της εικόνας μπορεί να υπολογιστεί μέσω της απλής πράξης: Σ=A+C-B-D, όπου A, B, C και D είναι οι τιμές της αθροιστικής εικόνας στις αντίστοιχες θέσεις..... 44

Εικόνα 2-3. Φίλτρα, διαστάσεων 9x9, που προκύπτουν από τη διακριτοποίηση και κοπή των παραγώγων δεύτερης τάξης της συνάρτησης Gauss δύο διαστάσεων στις διευθύνσεις x, y και xy 46

Εικόνα 2-4. Φίλτρα, διαστάσεων 9x9, που προσεγγίζουν τις παραγώγους δεύτερης τάξης της συνάρτησης Gauss δύο διαστάσεων για σ=1.2, στις διευθύνσεις x, y και xy..... 46

Εικόνα 2-5. Γκαουσιανή πυραμίδα και πυραμίδα DoG (Difference of Gaussian) 48

Εικόνα 2-6. Αντί για την επαναληπτική μείωση του μεγέθους της εικόνας, ο αλγόριθμος SURF χρησιμοποιεί την αύξηση του μεγέθους του φίλτρου για την αναπαράσταση του χώρου κλίμακας 49

Εικόνα 2-7. Μεγέθη l_o και l_1 των 9x9 φίλτρων που προσεγγίζουν τις παραγώγους δεύτερης τάξης της συνάρτησης Gauss δύο διαστάσεων για σ=1.2, στις διευθύνσεις x, y και xy..... 50

Εικόνα 2-8. Φίλτρα 9x9 και 15x15 που προσεγγίζουν τις παραγώγους δεύτερης τάξης της συνάρτησης Gauss δύο διαστάσεων στις διευθύνσεις y και xy για δύο διαδοχικά επίπεδα κλίμακας της πρώτης οκτάβας 50

Εικόνα 2-9. Περιοχή 3x3x3 στην οποία εφαρμόζεται η μέθοδος απαλοιφής των μη μεγίστων 51

Εικόνα 2-10. Δομές blobs με διαφορετικό τύπο αντίθεσης..... 52

Εικόνα 2-11. Haar wavelet φίλτρα..... 53

Εικόνα 2-12. Διαδικασία ανάθεσης προσανατολισμού του σημείου ενδιαφέροντος..... 53

Εικόνα 2-13. Τετράγωνα παράθυρα που σχηματίζονται γύρω από κάθε σημείο ενδιαφέροντος, με μέγεθος εικοσαπλάσιο της κλίμακας στην οποία εντοπίστηκε και προσανατολισμένα κατά μήκος του κυρίαρχου προσανατολισμού κάθε σημείου 54

Εικόνα 2-14. Δημιουργία μίας τετράγωνης περιοχής γύρω από κάθε σημείο ενδιαφέροντος, προσανατολισμένης με βάση τον κυρίαρχο προσανατολισμό, και διαίρεσή της σε 16 τετράγωνες υποπεριοχές 55

Εικόνα 3-1. Προβολή στο επίπεδο w=1 των ομογενών διανυσμάτων $[x\ y\ w_i]^T$ για διάφορες τιμές του i 67

Εικόνα 3-2. Γεωμετρική ερμηνεία των ομογενών συντεταγμένων - προβολή στο επίπεδο w=1 68

Εικόνα 3-3. Το μοντέλο κεντρικής προβολής (pinhole camera model) 70

Εικόνα 3-4. Απλοποιημένη γεωμετρία κεντρικής προβολής 75

Εικόνα 3-5. Σύστημα εικονοσυντεταγμένων και θεωρούμενο σύστημα αναφοράς των μετρήσεων της εικόνας 76

Εικόνα 3-6. Μετάθεση και στροφή του επίγειου συστήματος συντεταγμένων για να ταυτιστεί με το φωτογραμμετρικό 77

Εικόνα 3-7. Αλληλουχία των γραμμικών δισδιάστατων μετασχηματισμών 81

Εικόνα 3-8. Διπλός λόγος 86

Εικόνα 3-9. Προβολικός μετασχηματισμός μεταξύ επίπεδου αντικειμένου του χώρου και της εικόνας του και μεταξύ εικόνων του ίδιου δισδιάστατου αντικειμένου 89

Εικόνα 3-10. Εφαρμογή αφινικού μετασχηματισμού και ομογραφίας 90

Εικόνα 3-11. Αρχική εικόνα, συστολή της και διαστολή της 93

Εικόνα 3-12. 4-γείτονες (4-neighbours) και 8-γείτονες (8-neighbours) ενός εικονοστοιχείου P 94

Εικόνα 3-13. Παράδειγμα 4-συνδεδεμένου συστατικού (4-connected component) και 8-συνδεδεμένου συστατικού (8-connected component) 94

Εικόνα 3-14. Παράδειγμα 4-διαδρομής (4-path) και 8-διαδρομής (8-path) 94

Εικόνα 3-15. Παράδειγμα συστολής, διαστολής και κατωφλίωσης σε ασπρόμαυρη εικόνα σκακιέρας 97

Εικόνα 3-16. Εντοπισμός των κορυφών των τετραγώνων της σκακιέρας με υποψηφιδική ακρίβεια, με βάση την κλίση της εικόνας 100

Εικόνα 3-17. Αρίθμηση των εσωτερικών γωνιών της σκακιέρας 100

Εικόνα 4-1. Περιβάλλον ανάπτυξης προγραμμάτων C++: Microsoft Visual C++ 2010 Express 121

Εικόνα 4-2. Συστήματα συντεταγμένων κάμερας της OpenGL και της OpenCV 124

Εικόνα 4-3. Όγκος θέασης στην περίπτωση επιλογής της προοπτικής προβολής 130

Εικόνα 4-4. Τομή στο επίπεδο XZ του όγκου θέασης 131

Εικόνα 4-5. Τομή στο επίπεδο YZ του όγκου θέασης 132

Εικόνα 4-6. Ένα σημείο εντός της κόλουρης ορθογωνικής πυραμίδας (σε συντεταγμένες κάμερας) μετασχηματίζεται σε σημείο εντός ενός κύβου (σε κανονικοποιημένες συντεταγμένες συσκευής) 134

Εικόνα 4-7. Γραμμική σχέση μεταξύ των X_{NDC} και X_{CLIP} και μεταξύ των Y_{NDC} και Y_{CLIP} 134

Εικόνα 4-8. Σχέση μεταξύ των συντεταγμένων του πρωτεύοντος σημείου και των διαστάσεων της εικόνας με τις αποστάσεις των τεσσάρων επιπέδων αποκοπής (top, bottom, left, right) από το σημείο λήψης 137

Εικόνα 4-9. Θεωρούμενος όγκος θέασης στην ορθογραφική προβολή 139

Εικόνα 4-10. Σχέση μεταξύ συστήματος κανονικοποιημένων συντεταγμένων συσκευής και συστήματος συντεταγμένων παραθύρου 142

Εικόνα 5-1. Εικόνες σκακιέρας από το κινητό τηλέφωνο Samsung Galaxy Y Pro young GT-B5510 που χρησιμοποιήθηκαν για τη βαθμονόμηση της κάμεράς του 152

Εικόνα 5-2. Περιεχόμενο αρχείου Imagelist_GTB5510.xml 153

Εικόνα 5-3. Εύρεση των εσωτερικών γωνιών της σκακιέρας κατά τη διαδικασία της βαθμονόμησης 154

Εικόνα 5-4. Στιγμιότυπα από τη διαδικασία της βαθμονόμησης της κάμερας του φορητού υπολογιστή 156

Εικόνα 5-5. Τρισδιάστατα μοντέλα που χρησιμοποιήθηκαν στις εφαρμογές 159

Εικόνα 5-6. Σύστημα συντεταγμένων εικόνας, επίγειο σύστημα συντεταγμένων και συντεταγμένες των τεσσάρων κορυφών του επίπεδου αντικειμένου στα δύο συστήματα 168

Εικόνα 5-7. Παράθυρο «Pattern Image» στο οποίο εμφανίζεται η εικόνα - πρότυπο της Εφαρμογής 1 169

Εικόνα 5-8. Απεικόνιση των σημείων ενδιαφέροντος στην εικόνα - πρότυπο της Εφαρμογής 1 με ταυτόχρονη εμφάνιση οπτικών πληροφοριών για την κλίμακα και τον προσανατολισμό κάθε σημείου και απεικόνιση των σημείων ενδιαφέροντος μέσω απλών κύκλων χωρίς διαφοροποίηση του μεγέθους αυτών, η οποία τελικά χρησιμοποιήθηκε 171

Εικόνα 5-9. Παράθυρο «Live Feed», στο οποίο απεικονίζονται τα ζωντανά στιγμιότυπα της κάμερας, και παράθυρο «Tracked Features», στο οποίο απεικονίζονται τα στιγμιότυπα με τα σημεία ενδιαφέροντος που εντοπίστηκαν σε αυτά, για μία τυχαία σκηνή της ζωντανής μετάδοσης της κάμερας για την Εφαρμογή 1 182

Εικόνα 5-10. Παράθυρο «Input Video», στο οποίο απεικονίζονται τα στιγμιότυπα του αρχείου βίντεο, και παράθυρο «Tracked Features», στο οποίο απεικονίζονται τα στιγμιότυπα του βίντεο με τα σημεία ενδιαφέροντος που εντοπίστηκαν σε αυτά, για μία τυχαία σκηνή ενός αρχείου βίντεο της Εφαρμογής 1 183

Εικόνα 5-11. Παράθυρο «Matches» με τις ομολογίες που προέκυψαν μετά το διασταυρωμένο έλεγχο και την κατωφλίωση και αντίστοιχο παράθυρο «Inliers» με τις έγκυρες ομολογίες που προέκυψαν από τον αλγόριθμο RANSAC και χρησιμοποιήθηκαν για τον υπολογισμό του πίνακα της ομογραφίας, για ένα τυχαίο στιγμιότυπο της μετάδοσης της κάμερας του υπολογιστή (Εφαρμογή 1) 184

Εικόνα 5-12. Παράθυρο «Matches» με τις ομολογίες που προέκυψαν μετά το διασταυρωμένο έλεγχο και την κατωφλίωση και αντίστοιχο παράθυρο «Inliers» με τις έγκυρες ομολογίες που προέκυψαν από τον αλγόριθμο RANSAC και χρησιμοποιήθηκαν για τον υπολογισμό του πίνακα της ομογραφίας, για μία τυχαία σκηνή ενός αρχείου βίντεο της Εφαρμογής 1 185

Εικόνα 5-13. Παράθυρο «Live Feed», που εικονίζει ένα τυχαίο στιγμιότυπο της ζωντανής μετάδοσης της κάμερας του υπολογιστή, και παράθυρο «Warped Frame», που εικονίζει το αντίστοιχο αναδομημένο στιγμιότυπο που δείχνει το πρότυπο επίπεδο αντικείμενο (Εφαρμογή 1) 186

Εικόνα 5-14. Παράθυρο «Input Video», που εικονίζει ένα τυχαίο στιγμιότυπο του βίντεο, και παράθυρο «Warped Frame», που εικονίζει το αντίστοιχο αναδομημένο στιγμιότυπο που δείχνει το πρότυπο επίπεδο αντικείμενο..... 186

Εικόνα 5-15. Στιγμιότυπα από το παράθυρο «Pattern in the Frame» της Εφαρμογής 1, με μία τυχαία σκηνή της ζωντανής μετάδοσης της κάμερας και ενός αρχείου βίντεο. Σε κάθε στιγμιότυπο έχει σχεδιαστεί μία κόκκινη κλειστή γραμμή που ενώνει τις τέσσερις κορυφές του πρότυπου αντικειμένου 187

Εικόνα 5-16. Θεωρούμενο σύστημα συντεταγμένων μοντέλου και σύστημα συντεταγμένων παραθύρου. Σχεδίαση του τετραγώνου ως ορθογωνίου που καλύπτει όλη τη διαθέσιμη σχεδιαστική επιφάνεια του παραθύρου 188

Εικόνα 5-17. Αρχή του φωτογραμμετρικού συστήματος συντεταγμένων που θεωρήθηκε στην περίπτωση της απεικόνισης του τετραγώνου στο παράθυρο μέσω ορθογραφικής προβολής, όγκος θέασης και επίπεδα αποκοπής. 189

Εικόνα 5-18. Συντεταγμένες υφής των κορυφών του τετραγώνου..... 192

Εικόνα 5-19. Νέο σύστημα συντεταγμένων του μοντέλου, μετά το μετασχηματισμό που πραγματοποιείται μέσω της βιβλιοθήκης GLM..... 193

Εικόνα 5-20. Στιγμιότυπα από το παράθυρο «AR Window» της Εφαρμογής 1, στο οποίο επαυξάνονται οι σκηνές των στιγμιότυπων της ζωντανής μετάδοσης της κάμερας. Το μέγεθος της σχεδιαστικής επιφάνειας του παραθύρου είναι ίδιο με το μέγεθος των στιγμιότυπων της κάμερας του υπολογιστή 194

Εικόνα 5-21. Στιγμιότυπα από το παράθυρο «AR Window» της Εφαρμογής 1, το μέγεθος του οποίου έχει μεταβληθεί από το χρήστη (περίπτωση επαύξησης της πραγματικότητας με χρήση της κάμερας του υπολογιστή σε πραγματικό χρόνο) 195

Εικόνα 5-22. Στιγμιότυπα από το παράθυρο «AR Window» της Εφαρμογής 1, στο οποίο επαυξάνονται οι σκηνές των αρχείων βίντεο. Το μέγεθος της σχεδιαστικής επιφάνειας του παραθύρου είναι ίδιο με το μέγεθος των στιγμιότυπων του αρχείου βίντεο 195

Εικόνα 5-23. Στιγμιότυπα από το παράθυρο «AR Window» της Εφαρμογής 1, το μέγεθος του οποίου έχει μεταβληθεί από το χρήστη (περίπτωση επαύξησης των σκηνών ενός αρχείου βίντεο) 196

Εικόνα 5-24. Στιγμιότυπο από την Εφαρμογή 1, στην περίπτωση επαύξησης των σκηνών ενός αρχείου βίντεο 197

Εικόνα 5-25. Σύστημα συντεταγμένων εικόνας και επίγειο και συντεταγμένες των τεσσάρων κορυφών του πρότυπου αντικειμένου (τμήμα της νότιας πλευράς του κτηρίου Καθαρτήρια) στα δύο συστήματα, για την περίπτωση που το μοντέλο επαύξησης είναι το αγγείο (Εφαρμογή 2) 209

Εικόνα 5-26. Σύστημα συντεταγμένων εικόνας και επίγειο και συντεταγμένες των τεσσάρων γωνιών του πρότυπου αντικειμένου (τμήμα της βόρειας πλευράς του κτηρίου Μακρυγιάννη) στα δύο συστήματα, για την περίπτωση που το μοντέλο επαύξησης είναι το άγαλμα (Εφαρμογή 2) 210

Εικόνα 5-27. Εικόνα - πρότυπο του τμήματος της νότιας πλευράς του κτηρίου Καθαρτήρια και εντοπισμένα σημεία ενδιαφέροντος με χρήση του αλγορίθμου SURF για την Εφαρμογή 2 211

Εικόνα 5-28. Εικόνα - πρότυπο του τμήματος του κτηρίου Μακρυγιάννη και εντοπισμένα σημεία ενδιαφέροντος με χρήση του αλγορίθμου SURF για την Εφαρμογή 2 211

Εικόνα 5-29. Παράθυρο «Input Video», με ένα τυχαίο στιγμιότυπο του βίντεο, και παράθυρο «Tracked Features», με τα σημεία ενδιαφέροντος που εντοπίστηκαν στο ίδιο στιγμιότυπο, για την περίπτωση του βίντεο που θα επαυξηθεί με το αγγείο (Εφαρμογή 2) 217

Εικόνα 5-30. Παράθυρο «Input Video», με ένα τυχαίο στιγμιότυπο του βίντεο, και παράθυρο «Tracked Features», με τα σημεία ενδιαφέροντος που εντοπίστηκαν στο ίδιο στιγμιότυπο, για την περίπτωση του βίντεο που θα επαυξηθεί με το άγαλμα (Εφαρμογή 2) 217

Εικόνα 5-31. Παράθυρο «Matches» με τις ομολογίες που προκύπτουν μετά τον έλεγχο του λόγου των αποστάσεων των εγγύτερων γειτόνων και την κατωφλίωση και αντίστοιχο παράθυρο «Inliers» με τις έγκυρες ομολογίες που προέκυψαν από τον αλγόριθμο RANSAC και χρησιμοποιήθηκαν για τον υπολογισμό του πίνακα της ομογραφίας, για μία τυχαία σκηνή ενός αρχείου βίντεο της Εφαρμογής 2 που θα επαυξηθεί με το τρισδιάστατο μοντέλου του αγγείου 218

Εικόνα 5-32. Παράθυρο «Matches» με τις ομολογίες που προκύπτουν μετά τον έλεγχο του λόγου των αποστάσεων των εγγύτερων γειτόνων και την κατωφλίωση και αντίστοιχο παράθυρο «Inliers» με τις έγκυρες ομολογίες που προέκυψαν από τον αλγόριθμο RANSAC και χρησιμοποιήθηκαν για τον υπολογισμό του πίνακα της ομογραφίας, για μία τυχαία σκηνή ενός αρχείου βίντεο της Εφαρμογής 2 που θα επαυξηθεί με το τρισδιάστατο μοντέλο του αγάλματος 219

Εικόνα 5-33. Παράθυρο «Input Video», που εικονίζει ένα τυχαίο στιγμιότυπο του βίντεο, και παράθυρο «Warped Frame», που εικονίζει το αντίστοιχο αναδομημένο στιγμιότυπο που δείχνει το πρότυπο αντικείμενο, στην περίπτωση που το βίντεο θα επαυξηθεί με το αγγείο (Εφαρμογή 2)..... 219

Εικόνα 5-34. Παράθυρο «Input Video», που εικονίζει ένα τυχαίο στιγμιότυπο του βίντεο, και παράθυρο «Warped Frame», που εικονίζει το αντίστοιχο αναδομημένο στιγμιότυπο που δείχνει το πρότυπο αντικείμενο, στην περίπτωση που το βίντεο θα επαυξηθεί με το άγαλμα (Εφαρμογή 2)..... 220

Εικόνα 5-35. Παράθυρο «Pattern in the Frame» της Εφαρμογής 2, με ένα τυχαίο στιγμιότυπο του βίντεο της περιοχής που θα επαυξηθεί με το αγγείο. Στο στιγμιότυπο έχει σχεδιαστεί μία κόκκινη κλειστή γραμμή που ενώνει τις τέσσερις κορυφές του πρότυπου αντικειμένου που εντοπίστηκαν..... 220

Εικόνα 5-36. Παράθυρο «Pattern in the Frame» της Εφαρμογής 2, με ένα τυχαίο στιγμιότυπο του βίντεο της περιοχής που θα επαυξηθεί με το άγαλμα. Στο στιγμιότυπο έχει σχεδιαστεί μία κόκκινη γραμμή που ενώνει τις τέσσερις κορυφές του πρότυπου αντικειμένου που εντοπίστηκαν..... 220

Εικόνα 5-37. Στιγμιότυπα από το παράθυρο «AR Window» της Εφαρμογής 2, στο οποίο επαυξάνεται το πραγματικό περιεχόμενο του βίντεο με το τρισδιάστατο μοντέλο του αγγείου. Το μέγεθος της σχεδιαστικής επιφάνειας των παραθύρων είναι ίδιο με το μέγεθος των στιγμιότυπων του βίντεο 223

Εικόνα 5-38. Στιγμιότυπα από το παράθυρο «AR Window» της Εφαρμογής 2, το μέγεθος του οποίου έχει μεταβληθεί από το χρήστη (περίπτωση επαύξησης του περιεχομένου του βίντεο με το αγγείο) 224

Εικόνα 5-39. Στιγμιότυπα από το παράθυρο «AR Window» της Εφαρμογής 2, στο οποίο επαυξάνεται το πραγματικό περιεχόμενο του βίντεο με το τρισδιάστατο μοντέλο του αγάλματος. Το μέγεθος της σχεδιαστικής επιφάνειας των παραθύρων είναι ίδιο με το μέγεθος των στιγμιότυπων του βίντεο 225

Εικόνα 5-40. Στιγμιότυπα από το παράθυρο «AR Window» της Εφαρμογής 2, το μέγεθος του οποίου έχει μεταβληθεί από το χρήστη (περίπτωση επαύξησης του περιεχομένου του βίντεο με το άγαλμα) 226

Εικόνα 5-41. Στιγμιότυπο από την Εφαρμογή 2, στην περίπτωση επαύξησης της περιοχής μπροστά από τη νότια πλευρά του κτηρίου Καθαρτήρια 227

Εικόνα 5-42. Στιγμιότυπο από την Εφαρμογή 2, στην περίπτωση επαύξησης της βόρειας πλευράς του κτηρίου Μακρυγιάννη 227

Εικόνα Α-1. Υπόμνημα των χρησιμοποιούμενων συμβόλων των διαγραμμάτων ροής 257

Εικόνα Γ-1. Εισαγωγικό μήνυμα των Εφαρμογών 1 και 2 261

Εικόνα Γ-2. Μήνυμα που εμφανίζεται στο χρήστη της Εφαρμογής 1 για την πληκτρολόγηση της επιλογής του σχετικά με το αν επιθυμεί μία πραγματική εφαρμογή επαυξημένης πραγματικότητας ή αν επιδιώκει να εξετάσει τις δυνατές επιλογές του 261

Εικόνα Γ-3. Επισήμανση της λανθασμένης απάντησης του χρήστη της Εφαρμογής 1 και προτροπή να πληκτρολογήσει τη σωστή απάντηση, ανάμεσα στους χαρακτήρες «a» ή «c» 262

Εικόνα Γ-4. Μήνυμα που εμφανίζεται στο χρήστη των εφαρμογών με το ερώτημα αν θέλει να πληκτρολογήσει τα στοιχεία του εσωτερικού προσανατολισμού της κάμερας 262

Εικόνα Γ-5. Επισήμανση της λανθασμένης απάντησης του χρήστη των εφαρμογών και προτροπή να πληκτρολογήσει τη σωστή απάντηση, ανάμεσα στους χαρακτήρες «y» ή «n» 262

Εικόνα Γ-6. Αλληλεπίδραση του χρήστη με την εφαρμογή, κατά την οποία πληκτρολογεί τα στοιχεία του εσωτερικού προσανατολισμού 263

Εικόνα Γ-7. Μήνυμα που εμφανίζεται στο χρήστη της Εφαρμογής 1 με το ερώτημα αν θέλει να επαυξήσει την πραγματικότητα ζωντανά μέσω της κάμερας του υπολογιστή του ή μέσω ενός αρχείου βίντεο 264

Εικόνα Γ-8. Επισήμανση της λανθασμένης απάντησης του χρήστη της Εφαρμογής 1 και προτροπή να πληκτρολογήσει τη σωστή απάντηση, ανάμεσα στους χαρακτήρες «l» ή «v» 264

Εικόνα Γ-9. Μήνυμα με το οποίο ερωτάται ο χρήστης των εφαρμογών αν θέλει να δει τη βασική διαδικασία που πραγματοποιείται για την επαύξηση της πραγματικότητας 265

Εικόνα Γ-10. Επισήμανση της λανθασμένης απάντησης του χρήστη και προτροπή να πληκτρολογήσει τη σωστή απάντηση, ανάμεσα στους χαρακτήρες «p» και «a» 265

Εικόνα Γ-11. Μήνυμα που εμφανίζεται στο χρήστη της Εφαρμογής 1 προκειμένου να επιλέξει να δει μία προεπισκόπηση της εφαρμογής ή να εισαγάγει το δικό του αρχείο βίντεο 266

Εικόνα Γ-12. Επισήμανση της λανθασμένης απάντησης του χρήστη της Εφαρμογής 1 και προτροπή να πληκτρολογήσει τη σωστή απάντηση, ανάμεσα στους χαρακτήρες «p» και «f» 266

Εικόνα Γ-13. Μήνυμα που καλεί το χρήστη της Εφαρμογής 1 να πληκτρολογήσει το όνομα του αρχείου βίντεο, το περιεχόμενο του οποίου επιθυμεί να επαυξήσει 266

Εικόνα Γ-14. Προειδοποιητικό μήνυμα προς το χρήστη της Εφαρμογής 1 σχετικά με την αποτυχία κατά το άνοιγμα της κάμερας του υπολογιστή 267

Εικόνα Γ-15. Προειδοποιητικό μήνυμα προς το χρήστη σχετικά με την αποτυχία κατά το άνοιγμα του αρχείου βίντεο 267

Εικόνα Γ-16. Μήνυμα σφάλματος κατά την αρχικοποίηση της βιβλιοθήκης GLEW 268

Εικόνα Γ-17. Παράδειγμα αλληλεπίδρασης του χρήστη με την Εφαρμογή 1 269

Εικόνα Δ-1. Μήνυμα που εμφανίζεται στο χρήστη της Εφαρμογής 2 για το αν επιθυμεί να δει μία προεπισκόπηση της εφαρμογής ή να εισαγάγει τα δικά του αρχεία 271

Εικόνα Δ-2. Επισήμανση της λανθασμένης απάντησης του χρήστη της Εφαρμογής 2 και προτροπή να πληκτρολογήσει τη σωστή απάντηση, ανάμεσα στους χαρακτήρες «p» και «f» 271

Εικόνα Δ-3. Ερώτηση που απευθύνεται στο χρήστη της Εφαρμογής 2 σχετικά με την επιλογή του τρισδιάστατου μοντέλου που θα αποτελέσει το μοντέλο επαύξησης της πραγματικότητας 272

Εικόνα Δ-4. Επισήμανση της λανθασμένης απάντησης του χρήστη της Εφαρμογής 2 και προτροπή να πληκτρολογήσει τη σωστή απάντηση, ανάμεσα στους χαρακτήρες «s» και «v» 272

Εικόνα Δ-5. Μήνυμα που καλεί το χρήστη της Εφαρμογής 2 να πληκτρολογήσει το όνομα του αρχείου βίντεο, του οποίου το περιεχόμενο επιθυμεί να επαυξήσει 273

Εικόνα Δ-6. Μήνυμα που καλεί το χρήστη της Εφαρμογής 2 να πληκτρολογήσει το όνομα του αρχείου κειμένου που περιλαμβάνει τις τιμές του GPS 273

Εικόνα Δ-7. Μήνυμα σφάλματος της Εφαρμογής 2 στο άνοιγμα του αρχείου με τις τιμές του GPS 274

Εικόνα Δ-8. Μήνυμα προειδοποίησης προς το χρήστη της Εφαρμογής 2 ότι η περιοχή που επιθυμεί να επαυξήσει δεν υποστηρίζεται από την εφαρμογή 274

Εικόνα Δ-9. Μήνυμα σφάλματος της Εφαρμογής 2 κατά το άνοιγμα του αρχείου βίντεο 275

Εικόνα Δ-10. Παράθυρο «Input Video» της Εφαρμογής 2 στο οποίο αναπαράγεται ένα αρχείο βίντεο που δεν προορίζεται για επαύξηση από την Εφαρμογή 275

ΚΑΤΑΛΟΓΟΣ ΔΙΑΓΡΑΜΜΑΤΙΚΩΝ ΠΑΡΟΥΣΙΑΣΕΩΝ

Διαγραμματική παρουσίαση 2-1. Απλοποιημένο διάγραμμα ροής του αλγορίθμου RANSAC, προσαρμοσμένο στην εύρεση των χονδροειδών σφαλμάτων στις ομολογίες σημείων 63

Διαγραμματική παρουσίαση 3-1. Διαδικασία βαθμονόμησης με χρήση επίπεδης σκακιέρας 112

Διαγραμματική παρουσίαση 3-2. Υπολογισμός των στοιχείων του εξωτερικού προσανατολισμού μίας εικόνας σε περίπτωση που διατίθενται ομολογίες σημείων ενός επίπεδου αντικειμένου 114

Διαγραμματική παρουσίαση 4-1. Απλοποιημένη διαδικασία που εκτελείται σε ένα τυπικό περιβάλλον ανάπτυξης προγραμμάτων C++ 120

Διαγραμματική παρουσίαση 4-2. Μετασχηματισμοί συντεταγμένων που πραγματοποιούνται από την OpenGL 126

Διαγραμματική παρουσίαση 5-1. Γενικό διάγραμμα ροής των εφαρμογών επαυξημένης πραγματικότητας βάσει επίπεδου προτύπου 151

Διαγραμματική παρουσίαση 5-2. Εφαρμογή 1, μέρος I 163

Διαγραμματική παρουσίαση 5-3. Εφαρμογή 1, μέρος II 164

Διαγραμματική παρουσίαση 5-4. Εφαρμογή 1, μέρος III 173

Διαγραμματική παρουσίαση 5-5. Εφαρμογή 1, μέρος IV 174

Διαγραμματική παρουσίαση 5-6. Εφαρμογή 1, πλαίσιο συμπληρώσεων 1 174

Διαγραμματική παρουσίαση 5-7. Εφαρμογή 1 - Εφαρμογή 2, πλαίσιο συμπληρώσεων 2 175

Διαγραμματική παρουσίαση 5-8. Εφαρμογή 1 - Εφαρμογή 2, πλαίσιο συμπληρώσεων 3 175

Διαγραμματική παρουσίαση 5-9. Εφαρμογή 1, πλαίσιο συμπληρώσεων 4 177

Διαγραμματική παρουσίαση 5-10. Εφαρμογή 1, πλαίσιο συμπληρώσεων 5 177

Διαγραμματική παρουσίαση 5-11. Εφαρμογή 1, πλαίσιο συμπληρώσεων 6, μέρος I 198

Διαγραμματική παρουσίαση 5-12. Εφαρμογή 1, πλαίσιο συμπληρώσεων 6, μέρος II 199

Διαγραμματική παρουσίαση 5-13. Εφαρμογή 1 - Εφαρμογή 2, πλαίσιο συμπληρώσεων 7 199

Διαγραμματική παρουσίαση 5-14. Εφαρμογή 1, πλαίσιο συμπληρώσεων 8 200

Διαγραμματική παρουσίαση 5-15. Εφαρμογή 1 - Εφαρμογή 2, πλαίσιο συμπληρώσεων 9 200

Διαγραμματική παρουσίαση 5-16. Εφαρμογή 2, μέρος I 205

Διαγραμματική παρουσίαση 5-17. Εφαρμογή 2, μέρος II 206

Διαγραμματική παρουσίαση 5-18. Εφαρμογή 2, πλαίσιο συμπληρώσεων 10 207

Διαγραμματική παρουσίαση 5-19. Εφαρμογή 2, μέρος III 212

Διαγραμματική παρουσίαση 5-20. Εφαρμογή 2, μέρος IV 213

Διαγραμματική παρουσίαση 5-21. Εφαρμογή 2, πλαίσιο συμπληρώσεων 11 213

Διαγραμματική παρουσίαση 5-22. Εφαρμογή 2, πλαίσιο συμπληρώσεων 12 228

Διαγραμματική παρουσίαση 5-23. Εφαρμογή 2, πλαίσιο συμπληρώσεων 13 229

Διαγραμματική παρουσίαση 5-24. Εφαρμογή 2, πλαίσιο συμπληρώσεων 14 230

Διαγραμματική παρουσίαση 5-25. Εφαρμογή 2, πλαίσιο συμπληρώσεων 15 230

Διαγραμματική παρουσίαση 5-26. Εφαρμογή 2, πλαίσιο συμπληρώσεων 16 231

ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ

Πίνακας 3-1. Ιεραρχία των γραμμικών δισδιάστατων μετασχηματισμών..... 81

Πίνακας 5-1. Αποτελέσματα βαθμονόμησης της κάμερας του κινητού τηλεφώνου. 155

Πίνακας 5-2. Αποτελέσματα βαθμονόμησης της κάμερας του φορητού ηλεκτρονικού υπολογιστή..... 156

Πίνακας 5-3. Αναπαράσταση της τρισδιάστατης γεωμετρίας μέσω αρχείου OBJ. 158

ΕΙΣΑΓΩΓΗ

Η εξαγωγή πληροφοριών από ψηφιακές εικόνες απασχολεί τις τελευταίες δεκαετίες εκτός από τους φωτογραμμέτρες και άλλους ερευνητές και επιστήμονες, οι οποίοι αποτελούν την κοινότητα της όρασης υπολογιστών (computer vision) και προσπαθούν να δώσουν απαντήσεις και λύσεις σε προβλήματα τα οποία διερευνά και η κοινότητα της φωτογραμμετρίας. Η έρευνα στο πεδίο της όρασης υπολογιστών έχει καταλήξει σε διάφορες μεθόδους που είναι μεγάλης σημασίας για την ψηφιακή φωτογραμμετρία και χρησιμοποιούνται όλο και περισσότερο, λόγω της αυτοματοποίησης και της υπολογιστικής αποτελεσματικότητας που συνεπάγονται. Ο συνδυασμός αλγορίθμων και μεθόδων των δύο αυτών συγγενών επιστημών είναι το κλειδί για την επίτευξη σημαντικών αποτελεσμάτων σε εφαρμογές που επεξεργάζονται εικόνες και εξάγουν πληροφορίες από αυτές.

Σκοπός της παρούσας διπλωματικής εργασίας είναι ο συνδυασμός των φωτογραμμετρικών γνώσεων και μεθόδων με αλγορίθμους της όρασης υπολογιστών, για την ανάπτυξη εφαρμογών επαυξημένης πραγματικότητας βάσει επίπεδου προτύπου και τη διερεύνηση των αποτελεσμάτων που μπορούν να επιτευχθούν. Στο πλαίσιο αυτό, ορίστηκαν τρεις στόχοι που πρέπει να έχουν εκπληρωθεί με το πέρας της εργασίας αυτής. Ο πρώτος στόχος αφορά στην κατανόηση της έννοιας της επαυξημένης πραγματικότητας και στον ορισμό μίας διαδικασίας που μπορεί να ακολουθηθεί προκειμένου να επιτευχθεί το τελικό αποτέλεσμα, βάσει αναγνώρισης ενός επίπεδου πρότυπου αντικειμένου με χρήση φωτογραμμετρικών μεθόδων και τεχνικών της όρασης υπολογιστών. Ο δεύτερος στόχος είναι η κατανόηση του θεωρητικού υποβάθρου των αλγορίθμων και των διαδικασιών που ορίστηκαν να εφαρμοστούν. Τέλος, ο τρίτος στόχος είναι ο προγραμματισμός των εφαρμογών με βάση τη διαδικασία που ορίστηκε για την εξέταση των τελικών αποτελεσμάτων.

Η επαυξημένη πραγματικότητα, διεθνώς γνωστή ως augmented reality (AR), είναι το επιστημονικό πεδίο που ασχολείται με το συνδυασμό του πραγματικού κόσμου με δεδομένα παραγόμενα από έναν υπολογιστή. Αυτά προσθέτουν επιπλέον πληροφορίες στον πραγματικό κόσμο και βελτιώνουν την αντίληψη που έχει ο άνθρωπος για την πραγματικότητα. Σήμερα γίνονται προσπάθειες για την αξιοποίηση των δυνατοτήτων της πολλά υποσχόμενης εν λόγω τεχνολογίας σε διάφορους τομείς, μέσω της ανάπτυξης εφαρμογών που απευθύνονται σε πληθώρα ανθρώπων, με διαφορετικά ενδιαφέροντα και ασχολίες. Οι εφαρμογές επαυξημένης πραγματικότητας που αναπτύχθηκαν είναι εσωτερικού

χώρου, η πρώτη, και εξωτερικού χώρου, η δεύτερη, και στηρίζονται στην αναγνώριση ενός επίπεδου αντικειμένου για την επαύξηση της σκηνής. Πρόκειται για εφαρμογές κονσόλας Win32 που εκτελούνται σε ηλεκτρονικούς υπολογιστές με λειτουργικό σύστημα Windows και υποστηρίζουν την επαύξηση της πραγματικότητας με τρισδιάστατα μοντέλα σε πραγματικό χρόνο μέσω της κάμερας του υπολογιστή και μέσω αρχείων βίντεο. Η γενική μεθοδολογία που ακολουθήθηκε είναι η είσοδος των στοιχείων του εσωτερικού προσανατολισμού της κάμερας, της εικόνας του πρότυπου επίπεδου αντικειμένου και του τρισδιάστατου μοντέλου, η εξαγωγή χαρακτηριστικών στην εικόνα του επίπεδου αντικειμένου και στο εκάστοτε στιγμιότυπο προς επαύξηση, η ψηφιακή συνταύτιση των εικόνων αυτών, η εξάλειψη των άκυρων ομολογιών, η εύρεση της γεωμετρικής σχέσης που συνδέει τις δύο εικόνες, ο εντοπισμός του επίπεδου αντικειμένου στο στιγμιότυπο και τέλος ο υπολογισμός των στοιχείων του εξωτερικού προσανατολισμού του στιγμιότυπου. Με βάση τα τελευταία, με γνώση της εσωτερικής γεωμετρίας της κάμερας και με θεώρηση της κλίμακας, της θέσης και του προσανατολισμού του τρισδιάστατου μοντέλου επαύξησης εντός του πραγματικού κόσμου, είναι δυνατή η τοποθέτησή του με το σωστό τρόπο και στην κατάλληλη θέση στο παράθυρο θέασης της επαυξημένης σκηνής.

Η εργασία διαιρείται σε πέντε κεφάλαια, εκ των οποίων τα τέσσερα πρώτα είναι κυρίως θεωρητικά, αλλά αναγκαία για την κατανόηση της έννοιας της επαυξημένης πραγματικότητας, της διαδικασίας που ακολουθήθηκε και του προγραμματιστικού υποβάθρου, ενώ το τελευταίο επικεντρώνεται στην ανάπτυξη των εφαρμογών, στην περιγραφή τους, στην παράθεση της μεθοδολογίας που ακολουθήθηκε και στην παρουσίαση των αποτελεσμάτων.

Συγκεκριμένα, το πρώτο κεφάλαιο αποτελεί μία εισαγωγή στην επαυξημένη πραγματικότητα. Δίνεται ο ορισμός της και τα βασικά χαρακτηριστικά των εφαρμογών της, περιγράφεται η εξέλιξή της στο χρόνο και παρουσιάζονται οι κύριοι τομείς στους οποίους μπορεί να συνεισφέρει. Παράλληλα, αναφέρονται οι τεχνολογίες θέασης των επαυξημένων σκηνών αλλά και οι μέθοδοι τοποθέτησης των εικονικών πληροφοριών στον πραγματικό κόσμο. Τέλος, καταγράφονται οι κύριες προκλήσεις τις οποίες αντιμετωπίζουν οι δημιουργοί εφαρμογών επαυξημένης πραγματικότητας για την επίτευξη της ρεαλιστικής επαύξησης των σκηνών του πραγματικού κόσμου.

Το δεύτερο κεφάλαιο αναφέρεται στην ψηφιακή συνταύτιση χαρακτηριστικών σημείων, η οποία εφαρμόστηκε κατά την ανάπτυξη των εφαρμογών για την εύρεση των ομολογιών μεταξύ της εικόνας του επίπεδου πρότυπου αντικειμένου και της εικόνας του στο εκάστοτε στιγμιότυπο. Για λόγους πληρότητας, παρουσιάζονται στην αρχή του κεφαλαίου εν συντομία οι βασικές κατηγορίες της ψηφιακής συνταύτισης, ενώ το υπόλοιπο κεφάλαιο επικεντρώνεται στην ψηφιακή συνταύτιση χαρακτηριστικών σημείων. Η τελευταία διεξάγεται σε δύο γενικά στάδια, το πρώτο εκ των οποίων είναι η ανίχνευση και η περιγραφή των σημείων

ενδιαφέροντος. Σχετικά με τη διαδικασία αυτή, δίνονται κάποιες γενικές πληροφορίες και αναφέρονται συνοπτικά διάφοροι αλγόριθμοι, ώστε στη συνέχεια να παρουσιαστεί αναλυτικά ο αλγόριθμος SURF που χρησιμοποιήθηκε στις εφαρμογές. Την ανίχνευση και περιγραφή των σημείων ενδιαφέροντος ακολουθεί το δεύτερο στάδιο, που είναι η κυρίως διαδικασία της συνταύτισης, δηλαδή η εύρεση των ομολογιών. Σχετικά με την τελευταία, παρουσιάζονται διάφορα κριτήρια ταύτισης των σημείων ενδιαφέροντος, γίνεται αναφορά στην εύρεση των ομολογιών μέσω του αλγορίθμου SURF και παρουσιάζονται διάφορες τεχνικές απόρριψης των μη έγκυρων ομολογιών, οι οποίες χρησιμοποιήθηκαν στο πλαίσιο των εφαρμογών. Αυτές είναι ο έλεγχος του λόγου των αποστάσεων των εγγύτερων γειτόνων (ratio test), ο διασταυρωμένος έλεγχος (cross-check test) και ο αλγόριθμος RANSAC.

Στο τρίτο κεφάλαιο παρουσιάζονται τα μαθηματικά μοντέλα που χρησιμοποιούνται στη φωτογραμμετρία και στην όραση υπολογιστών και εκφράζουν την κεντρική προβολή. Για την πληρέστερη κατανόηση του μοντέλου του προβολικού μετασχηματισμού, που υιοθετείται από την κοινότητα της όρασης υπολογιστών, γίνεται αναφορά στην έννοια της προβολικής γεωμετρίας και περιγράφονται αναλυτικά οι ομογενείς συντεταγμένες, οι οποίες χρησιμοποιούνται αντί των καρτεσιανών. Επίσης, παρουσιάζονται τα βασικά φωτογραμμετρικά προβλήματα της εύρεσης των στοιχείων του εσωτερικού προσανατολισμού της μηχανής και του εξωτερικού προσανατολισμού της κάθε λήψης που επιλύθηκαν στο πλαίσιο των εφαρμογών. Συγκεκριμένα, αναφέρονται βασικά στοιχεία για την κατανόηση της διαδικασίας της βαθμονόμησης, γίνεται ανασκόπηση διαφόρων μεθόδων βαθμονόμησης και δίνεται έμφαση στην αυτόματη διαδικασία βαθμονόμησης με χρήση εικόνων επίπεδης σκακιέρας που χρησιμοποιήθηκε στις εφαρμογές, η οποία στηρίζεται στον προβολικό μετασχηματισμό. Επίσης, παρουσιάζεται η διαδικασία υπολογισμού των στοιχείων του εξωτερικού προσανατολισμού, με χρήση του προβολικού μετασχηματισμού, που ακολουθήθηκε στις εφαρμογές. Επιπλέον, αντικείμενο του κεφαλαίου αυτού αποτελούν οι δισδιάστατοι γραμμικοί μετασχηματισμοί, και κυρίως ο προβολικός μετασχηματισμός ή ομογραφία. Συγκεκριμένα, οι κυριότεροι γραμμικοί δισδιάστατοι μετασχηματισμοί καταγράφονται ως μία «ιεραρχία», από τον ειδικότερο στο γενικότερο, και παρουσιάζονται οι μαθηματικές σχέσεις αυτών, με χρήση τόσο των καρτεσιανών, όσο και των ομογενών συντεταγμένων. Τέλος, αιτιολογείται η επιλογή της ομογραφίας ως της σχέσης που συνδέει τα εικονοσημεία διαφόρων εικόνων του ίδιου δισδιάστατου αντικειμένου ώστε να γίνει κατανοητός ο λόγος χρήσης της στις εφαρμογές επαυξημένης πραγματικότητας.

Το τέταρτο κεφάλαιο αναφέρεται στο προγραμματιστικό περιβάλλον εργασίας. Ειδικότερα, γίνεται μία εισαγωγή στη γλώσσα προγραμματισμού των εφαρμογών C++, περιγράφονται τα βασικά στοιχεία ενός περιβάλλοντος ανάπτυξης προγραμμάτων C++, γίνεται αναφορά στο συγκεκριμένο περιβάλλον που χρησιμοποιήθηκε για την υλοποίηση των εφαρμογών και παρουσιάζονται οι βιβλιοθήκες που χρησιμοποιήθηκαν. Αυτές είναι η βιβλιοθήκη OpenCV,

οι σχετικές βιβλιοθήκες με τη διεπαφή προγραμματισμού εφαρμογών OpenGL, καθώς επίσης και η βιβλιοθήκη GLM: An Alias Wavefront OBJ file Library. Η βιβλιοθήκη OpenCV χρησιμοποιείται προκειμένου να υπολογιστούν τα στοιχεία του εξωτερικού προσανατολισμού του εκάστοτε στιγμιότυπου, αλλά και για τη βαθμονόμηση της μηχανής. Η OpenGL χρησιμοποιείται για τη σωστή τοποθέτηση του τρισδιάστατου μοντέλου στο στιγμιότυπο και τη θέαση της επαυξημένης σκηνής, ενώ η βιβλιοθήκη GLM χρησιμοποιείται για την εισαγωγή και εμφάνιση των τρισδιάστατων μοντέλων που διατίθενται σε μορφότυπο OBJ. Επιπλέον, στο κεφάλαιο αυτό επισημαίνονται κάποιες διαφορές των OpenCV και OpenGL, σχετικά με τα συστήματα συντεταγμένων που χρησιμοποιούν, και περιγράφονται οι μετασχηματισμοί που πραγματοποιούνται στην OpenGL προκειμένου ένα μοντέλο να παρουσιαστεί με το σωστό τρόπο σε ένα παράθυρο του υπολογιστή.

Στο πέμπτο και τελευταίο κεφάλαιο της παρούσας εργασίας παρουσιάζονται οι εφαρμογές επαυξημένης πραγματικότητας που αναπτύχθηκαν. Περιγράφεται λεπτομερώς η μεθοδολογία που ακολουθήθηκε, αναφέρονται οι δυνατότητες των εφαρμογών και παρουσιάζονται τα αποτελέσματά τους. Επίσης, γίνεται αναφορά στις βαθμονομήσεις που διεξήχθησαν για την κάλυψη των απαιτήσεων των εφαρμογών και στο μορφότυπο OBJ, στο οποίο διατέθηκαν τα τρισδιάστατα μοντέλα επαύξησης της πραγματικότητας. Τέλος, μετά το κεφάλαιο αυτό ακολουθούν γενικά συμπεράσματα και σκέψεις για το μέλλον, που αποτελούν προτάσεις για μελλοντική επέκταση και βελτίωση των εφαρμογών.

ΚΕΦΑΛΑΙΟ 1

ΕΠΑΥΞΗΜΕΝΗ ΠΡΑΓΜΑΤΙΚΟΤΗΤΑ

Η επαυξημένη πραγματικότητα αποτελεί μία πολλά υποσχόμενη τεχνολογία που εξελίσσεται συνεχώς και μολονότι προς το παρόν δεν είναι ευρέως γνωστή, ενδέχεται στο μέλλον να προκαλέσει σημαντικές αλλαγές στον τρόπο με τον οποίο ο χρήστης αποκτά πρόσβαση στην πληροφορία. Η τεχνολογία αυτή εμπλουτίζει την πραγματικότητα του χρήστη με πληροφορίες τις οποίες αυτός δεν μπορεί να αντιληφθεί μόνο με τα αισθητήρια όργανά του και έχει κεντρίσει τα τελευταία χρόνια το ενδιαφέρον τόσο επιστημόνων και ερευνητών, όσο και διαφόρων εταιριών που προσπαθούν να αξιοποιήσουν τις δυνατότητές της.

Το κεφάλαιο αυτό αποτελεί μία εισαγωγή στην επαυξημένη πραγματικότητα. Δίνονται πληροφορίες για την εξέλιξή της στο χρόνο, με έμφαση σε κάποια σημαντικά επιτεύγματα που σηματοδοτούν την ιστορία της, ενώ περιγράφονται και τα κύρια πεδία εφαρμογών στα οποία μπορεί να συνεισφέρει. Παράλληλα, παρουσιάζονται οι κύριες τεχνολογίες θέασης μίας επαυξημένης σκηνής αλλά και οι διάφορες μέθοδοι που χρησιμοποιεί η επαυξημένη πραγματικότητα για να τοποθετήσει τις πληροφορίες στον πραγματικό κόσμο. Τέλος, καταγράφονται οι κύριες προκλήσεις με τις οποίες έρχονται αντιμέτωποι οι δημιουργοί εφαρμογών επαυξημένης πραγματικότητας, στις οποίες πρέπει να δίνεται εξέχουσα σημασία προκειμένου αυτές να εκπληρώσουν το σκοπό τους.

1.1 Γενική Περιγραφή

Ο όρος επαυξημένη πραγματικότητα αναφέρεται στην τεχνολογία που επιτρέπει την προσθήκη πληροφοριών, που παράγονται από έναν υπολογιστή, στον πραγματικό κόσμο, μέσω κατάλληλων συσκευών. Πρόκειται για ένα ταχέως εξελισσόμενο ερευνητικό πεδίο, το οποίο στοχεύει στον εμπλουτισμό της αντίληψης της πραγματικότητας από το χρήστη και συνδέεται με ένα γενικότερο όρο, τη διαμεσολαβημένη πραγματικότητα (mediated reality), σύμφωνα με την οποία η αντίληψη της πραγματικότητας από το χρήστη τροποποιείται από έναν υπολογιστή [1], [2].

Ιδανικά, τα αντικείμενα που συμπληρώνουν τον πραγματικό κόσμο παρουσιάζονται να συνυπάρχουν στον ίδιο χώρο με αυτόν. Ωστόσο, δεν αποτελούν μόνο οπτικές πληροφορίες. Η επαυξημένη πραγματικότητα, δυνητικά, μπορεί να βελτιώσει και τις πέντε αισθήσεις, και κυρίως – εκτός από την όραση – την ακοή και την αφή, αν και σήμερα η κυρίαρχη χρήση της είναι η προσθήκη οπτικών πληροφοριών στον πραγματικό κόσμο. Επίσης, εκτός από την πρόσθεση πληροφοριών, η επαυξημένη πραγματικότητα μπορεί να χρησιμοποιηθεί για να αφαιρέσει πληροφορίες του πραγματικού κόσμου. Μολονότι η αφαίρεση πραγματικών αντικειμένων καλείται από κάποιους ερευνητές μειωμένη πραγματικότητα (diminished reality), ουσιαστικά αποτελεί ένα υποσύνολο της επαυξημένης πραγματικότητας [3], [4].

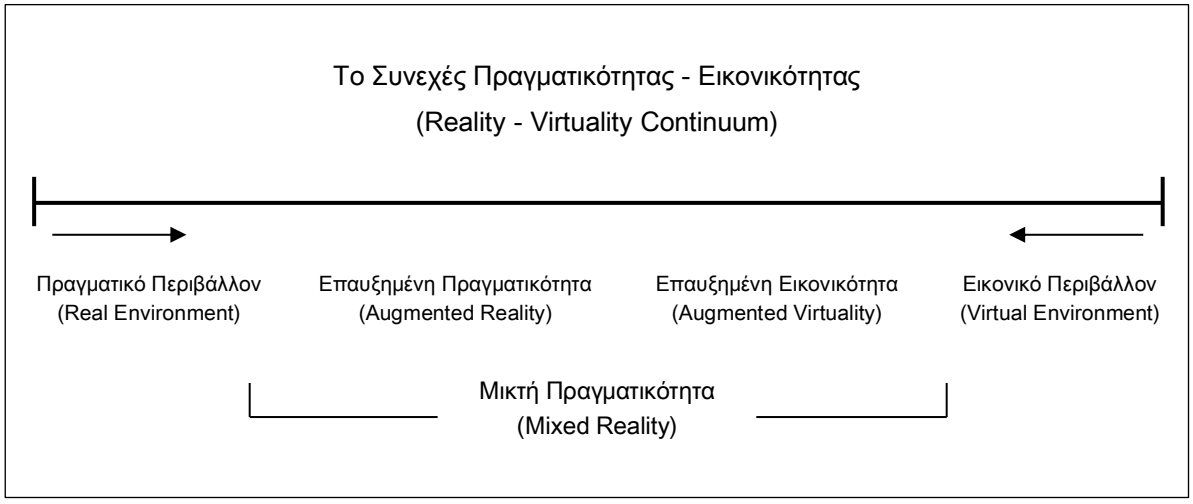
Υπάρχουν διάφοροι ορισμοί της επαυξημένης πραγματικότητας, εκ των οποίων ευρέως αποδεκτός είναι ο ορισμός που έδωσε ο Ronald Azuma το 1997 [5]. Σύμφωνα με αυτόν, η επαυξημένη πραγματικότητα είναι μία παραλλαγή της εικονικής πραγματικότητας (virtual reality), αλλά διαφέρει από την τελευταία, διότι συμπληρώνει τον πραγματικό κόσμο και δεν τον υποκαθιστά. Στοχεύει στην ενσωμάτωση συνθετικών πληροφοριών στο πραγματικό περιβάλλον – ή σε ένα βίντεο του πραγματικού περιβάλλοντος – με αποτέλεσμα αυτό να μην αποκρύπτεται εντελώς, αλλά αντίθετα να έχει κυρίαρχο ρόλο [6]. Από την άλλη πλευρά, στόχος της εικονικής πραγματικότητας είναι η πλήρης «εμβύθιση» του χρήστη σε ένα περιβάλλον πλήρως ελεγχόμενο από υπολογιστή, το οποίο δεν του επιτρέπει να δει τον πραγματικό κόσμο γύρω του [7]. Ο συνθετικός αυτός κόσμος μπορεί να μιμείται τις ιδιότητες ενός πραγματικού κόσμου, είτε αυτός υπάρχει είτε όχι, ή και να διαφέρει σημαντικά από την πραγματικότητα, δημιουργώντας ένα φανταστικό κόσμο, στον οποίο οι φυσικοί νόμοι δεν ισχύουν πλέον [8].

Τα συστήματα επαυξημένης πραγματικότητας σύμφωνα με τους Azuma et al. (2001) έχουν τα ακόλουθα τρία χαρακτηριστικά [3]:

- Συνδυάζουν τα πραγματικά και τα εικονικά αντικείμενα σε πραγματικό περιβάλλον
- Επιτρέπουν την αλληλεπίδραση σε πραγματικό χρόνο
- Εγγράφουν τα εικονικά αντικείμενα στον τρισδιάστατο χώρο

Τα παραπάνω χαρακτηριστικά καθίστανται σημαντικά στην αξιολόγηση αν ένα σύστημα είναι σύστημα επαυξημένης πραγματικότητας ή όχι. Έτσι, για παράδειγμα, κάποιες ταινίες, μολονότι μπορεί να περιλαμβάνουν φωτορεαλιστικά εικονικά αντικείμενα τα οποία συνδυάζονται ομαλά με το τρισδιάστατο πραγματικό περιβάλλον, δεν είναι διαδραστικές σε πραγματικό χρόνο και επομένως δεν αποτελούν συστήματα επαυξημένης πραγματικότητας [5]. Αντίθετα, η εισαγωγή των εικονικών δεδομένων οφείλεται σε προεπεξεργασία, η οποία αντιτίθεται με το δεύτερο χαρακτηριστικό των συστημάτων επαυξημένης πραγματικότητας.

Σύμφωνα με τους Paul Milgram et al. (1994), η επαυξημένη πραγματικότητα ανήκει στην τεχνολογία της μικτής πραγματικότητας (mixed reality) [8]. Στη δημοσίευσή τους, ορίζεται το «συνεχές πραγματικότητας - εικονικότητας» (reality - virtuality continuum), όπως φαίνεται στην Εικόνα 1-1. Το πραγματικό και το εικονικό περιβάλλον βρίσκονται στα δύο απέναντι άκρα του συνεχούς, ενώ στο περιβάλλον μικτής πραγματικότητας, αντικείμενα του πραγματικού και του εικονικού κόσμου – δηλαδή οτιδήποτε μεταξύ των άκρων του συνεχούς – παρουσιάζονται μαζί σε μία συσκευή θέασης. Η μικτή πραγματικότητα περιλαμβάνει τόσο την επαυξημένη πραγματικότητα, όσο και την επαυξημένη εικονικότητα (augmented virtuality). Ενώ στην πρώτη, το περιβάλλον που παρατηρείται είναι κυρίως αληθινό και έχουν προστεθεί βελτιώσεις που παράγονται από υπολογιστή, στην επαυξημένη εικονικότητα το παρατηρούμενο περιβάλλον είναι κυρίως εικονικό και επαυξάνεται με τη χρήση πραγματικών δεδομένων.



Εικόνα 1-1. Το Συνεχές Πραγματικότητας - Εικονικότητας (Reality - Virtuality Continuum).

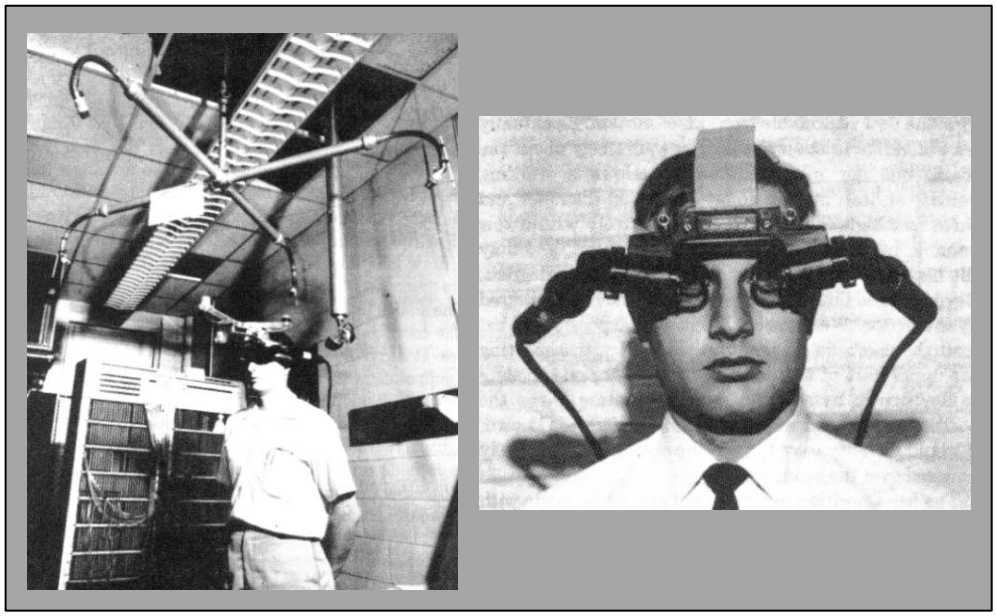
Πηγή: Milgram et al., 1994

1.2 Ιστορική Αναδρομή

Η πρώτη ιδέα που υπαινίσσεται την τεχνολογία της επαυξημένης πραγματικότητας χρονολογείται το 1901 και ανήκει στο συγγραφέα Frank Baum. Ο τελευταίος, στο μυθιστόρημά του «The Master Key», περιγράφει ένα ζευγάρι ειδικών γυαλιών, το οποίο υπερθέτει στο μέτωπο του κάθε ανθρώπου, που βρίσκεται στο οπτικό πεδίο εκείνου που φοράει τα γυαλιά, ένα γράμμα που περιγράφει το χαρακτήρα του (για παράδειγμα αν είναι καλός, κακός, σοφός κ.λπ.) [1], [9].

Το πρώτο σύστημα επαυξημένης πραγματικότητας κατασκευάζεται αρκετά αργότερα, από τον Ivan Sutherland το 1968 (βλ. Εικόνα 1-2). Πρόκειται για μία τρισδιάστατη συσκευή απεικόνισης που τοποθετείται στο κεφάλι του χρήστη («head-mounted three dimensional

display») και κρέμεται από το ταβάνι, με σκοπό να του προσθέσει τρισδιάστατες πληροφορίες [10]. Η βασική ιδέα πίσω από το σύστημα αυτό, το όνομα του οποίου είναι «The Sword of Damocles» (Δαμόκλειος Σπάθη) [11], είναι ότι η εικόνα ενός πραγματικού τρισδιάστατου αντικειμένου που κάποιος παρατηρεί, καταλήγει στον αμφιβληστροειδή χιτώνα του ματιού ως εικόνα δύο διαστάσεων. Συνεπώς, η τοποθέτηση κατάλληλων δισδιάστατων εικόνων μπροστά από τον αμφιβληστροειδή χιτώνα του ματιού του παρατηρητή μπορεί να δημιουργήσει σε αυτόν την ψευδαίσθηση της τρίτης διάστασης. Με βάση την παραπάνω λογική, η τρισδιάστατη συσκευή απεικόνισης του Sutherland παρουσιάζει στο χρήστη μία προοπτική εικόνα, η οποία αλλάζει καθώς αυτός κουνά το κεφάλι του ή κινείται σε μία περιορισμένη περιοχή. Η θέση και ο προσανατολισμός του οπτικού συστήματος το οποίο είναι στερεωμένο στο κεφάλι του χρήστη – που μετριούνται από το μηχανικό ή τον υπερηχητικό ανιχνευτή έξι βαθμών ελευθερίας (6DOF tracker) που κατασκεύασε ο Sutherland – καθορίζουν ποιά εικόνα είναι κάθε φορά η κατάλληλη.



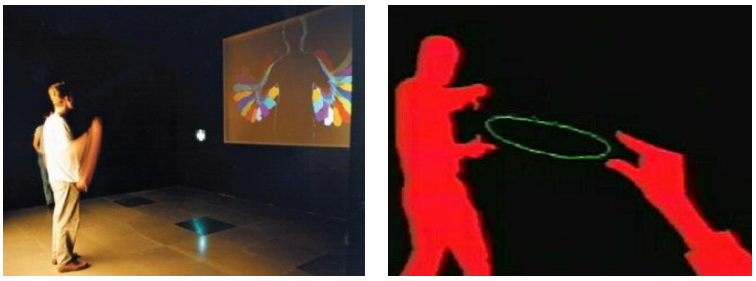
Εικόνα 1-2. Αριστερά: Το πρώτο σύστημα επαυξημένης πραγματικότητας από τον Ivan Sutherland (The Sword of Damocles). Δεξιά: Η συσκευή απεικόνισης του Sutherland που τοποθετείται στο κεφάλι του χρήστη.

Πηγή: Sutherland, 1968

Μπροστά από το κάθε μάτι του χρήστη παρουσιάζεται μία συνθετική εικόνα. Λόγω των ειδικών κατόπτρων στα πρίσματα μέσω των οποίων ο χρήστης κοιτάζει, μπορεί να βλέπει ταυτόχρονα τόσο τις εικόνες αυτές όσο και αντικείμενα από τον πραγματικό χώρο. Ωστόσο, οι παραπάνω εικόνες είναι εικόνες διάφανων τρισδιάστατων αντικειμένων, τα οποία αποτελούνται μόνο από τις κύριες ακμές τους («wire frame» line drawings) και όχι συμπαγών τρισδιάστατων αντικειμένων, λόγω των τεχνικών περιορισμών της εποχής, που καθιστούν δύσκολο τον υπολογισμό των μερών ενός αντικειμένου που κρύβεται από κάποιο άλλο. Αν και η επαυξημένη πραγματικότητα ως όρος επινοείται αρκετά αργότερα, το

σύστημα του Sutherland θεωρείται το πρώτο σύστημα επαυξημένης πραγματικότητας και είναι πράγματι για την εποχή του πρωτοπόρο.

Το 1975 ο Myron Krueger δημιουργεί το Videoplace, ένα εργαστήριο τεχνητής πραγματικότητας (artificial reality), όπως ο ίδιος το ονομάζει, στο Πανεπιστήμιο του Κονέκτικατ στις Ηνωμένες Πολιτείες. Μέσω της τεχνολογίας του Videoplace, το οποίο αποτελείται από δύο ή περισσότερα δωμάτια που μπορεί να απέχουν οποιαδήποτε απόσταση μεταξύ τους, μία οθόνη προβολής σε κάθε δωμάτιο και ειδικό hardware, οι χρήστες βλέπουν την εικόνα τους ως σιλουέτα στην οθόνη προβολής να κάνει τις ίδιες κινήσεις με αυτούς. Ταυτόχρονα, αλληλεπιδρούν με τις σιλουέτες των άλλων χρηστών στα συνδεδεμένα δωμάτια αλλά και με γραφικά αντικείμενα της οθόνης. Όλοι οι χρήστες σε όλα τα δωμάτια βλέπουν την ίδια εικόνα στην οθόνη. Επίσης, η εικόνα κάθε χρήστη μπορεί να χρωματιστεί, να σμικρυνθεί ή να περιστραφεί, ενώ γενικά το Videoplace προσφέρει στους χρήστες πολλούς τρόπους αλληλεπίδρασης, ανάμεσα στους οποίους και ζωγραφική με τα δάκτυλα [12], [13], [14]. Στην Εικόνα 1-3 φαίνεται το περιβάλλον τεχνητής πραγματικότητας του Krueger και σκηνή από την οθόνη προβολής, όπου δύο χρήστες αλληλεπιδρούν μεταξύ τους και με ένα γραφικό αντικείμενο.



Εικόνα 1-3. Σκηνές από το Videoplace του Krueger.

Πηγές: http://thedigitalage.pbworks.com/w/page/22039083/Myron%20Krueger, http://www.flong.com/texts/essays/essay_cvad

Το 1992 επινοείται ο όρος «επαυξημένη πραγματικότητα» από τον Tom Caudell, τρία χρόνια μετά την εισαγωγή του όρου «εικονική πραγματικότητα» από τον Jaron Lanier, το 1989. Ο Caudell μαζί με το συνάδελφό του David Mizell εργάζονται στην εταιρία Boeing πάνω σε ένα πρόγραμμα βελτίωσης της κατασκευής αεροσκαφών και συγκεκριμένα υποβοήθησης των εργατών στη συναρμολόγηση των καλωδιώσεων. Έτσι, κατασκευάζουν ένα σύστημα επαυξημένης πραγματικότητας με συσκευή θέασης που προσαρτάται στο κεφάλι των εργατών (Head-Mounted Display, HMD) και μπορεί να υπερθέσει στον πίνακα, στον οποίο αυτοί συνδέουν τις καλωδιώσεις, διαγράμματα που δείχνουν τις θέσεις στις οποίες τα καλώδια πρέπει να τοποθετηθούν [4], [15], [16], [17].

Την ίδια χρονιά, ο Louis Rosenberg αναπτύσσει ένα από τα πρώτα συστήματα επαυξημένης πραγματικότητας που τέθηκαν σε λειτουργία, με την ονομασία «Virtual Fixtures» (Εικονικά Εξαρτήματα) [1]. Σύμφωνα με δημοσίευσή του το 1992 [18], τα εικονικά εξαρτήματα

χρησιμοποιούνται σε συστήματα τηλεπαρουσίας (telepresence systems) και στόχος τους είναι η επέκταση των ευκολιών ενός χειριστή και των ικανοτήτων του στη λύση προβλημάτων σε ένα μακρινό περιβάλλον, στο οποίο αυτός αισθάνεται παρόν. Τα εικονικά εξαρτήματα είναι στην πραγματικότητα αφηρημένες αισθητήριες πληροφορίες (αισθήσεις παραγόμενες από υπολογιστή) που υπερτίθενται στο μακρινό χώρο εργασίας, στον οποίο ο χειριστής αισθάνεται ότι βρίσκεται, και στοχεύουν κυρίως στη βελτίωση της ακρίβειας και της επίδοσης στην εκτέλεση μίας εργασίας. Στη δημοσίευσή του, ο Rosenberg αναφέρει τη χρήση μόνο απτικών και ακουστικών εικονικών εξαρτημάτων.

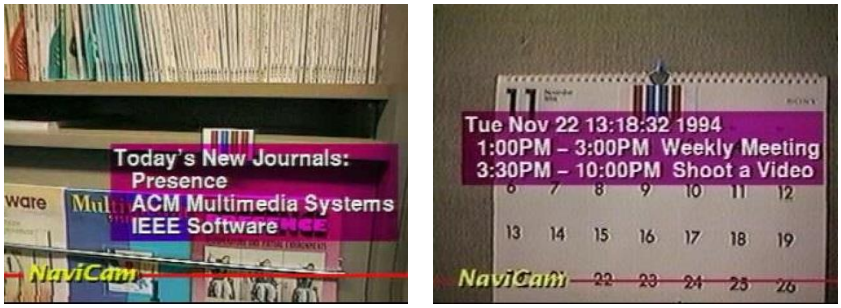
Το ίδιο έτος, οι Steven Feiner, Blair MacIntyre και Doree Seligmann παρουσιάζουν σε συνέδριο το πρώτο κύριο άρθρο πάνω σε ένα πρωτότυπο σύστημα επαυξημένης πραγματικότητας (βλ. Εικόνα 1-4). Το σύστημα αυτό, με την ονομασία KARMA («Knowledge-based Augmented Reality Maintenance Assistance»), χρησιμοποιεί ένα σύστημα HMD για την υποβοήθηση του τελικού χρήστη κατά τη συντήρηση ενός εκτυπωτή laser [1], [19].

Εικόνα 1-4. Σύστημα επαυξημένης πραγματικότητας KARMA [αριστερά] και εικόνα που βλέπει ο χρήστης μέσω της συσκευής θέασης [δεξιά].

Πηγή: http://monet.cs.columbia.edu/projects/karma/karma.html

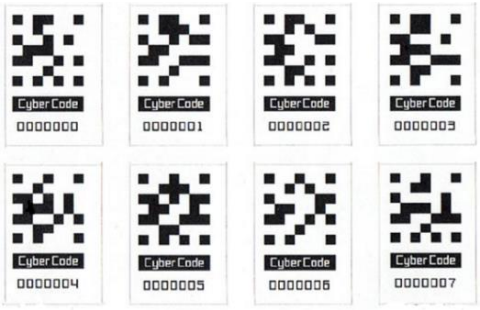
Το 1994 η επαυξημένη πραγματικότητα εισάγεται στο χώρο της τέχνης και της διασκέδασης, όταν η Julie Martin δημιουργεί την πρώτη θεατρική παραγωγή επαυξημένης πραγματικότητας με την ονομασία «Dancing in Cyberspace», στην οποία χορευτές και ακροβάτες αλληλεπιδρούν με εικονικά αντικείμενα σε πραγματικό χρόνο, τα οποία προβάλλονται στη σκηνή. Η παράσταση χρηματοδοτείται από το Συμβούλιο Τέχνης της Αυστραλίας και αναμεταδίδεται από την τηλεόραση [1], [17], [20].

Το 1995 ο Jun Rekimoto αναπτύσσει το πρώτο φορητό σύστημα χειρός επαυξημένης πραγματικότητας που βασίζεται σε έγχρωμους επίπεδους στόχους, ευρέως γνωστούς μέχρι σήμερα ως markers, το επονομαζόμενο NaviCam (βλ. Εικόνα 1-5). Το 1996 ο ίδιος δημιουργεί ένα 2D barcode σύστημα για εφαρμογές επαυξημένης πραγματικότητας (βλ. Εικόνα 1-6), που βασίζεται σε ασπρόμαυρα τετράγωνα markers δύο διαστάσεων, το CyberCode. Αυτός ο τύπος marker χρησιμοποιείται ακόμα και σήμερα [4], [21], [22], [23], [24].



Εικόνα 1-5. Παραδείγματα χρήσεων του συστήματος NaviCam.

Πηγές: http://www.sonycsl.co.jp/person/rekimoto/navi.html, http://www.sonycsl.co.jp/person/rekimoto/uist95/uist95.html



Εικόνα 1-6. Παραδείγματα επίπεδων στόχων του 2D barcode συστήματος του Rekimoto.

Πηγή: http://www.sonyinsider.com/wp-content/uploads/2011/05/cybercode.jpg

Το 1998 οι Bruce Thomas et al. [25] επιχειρούν να επεκτείνουν την επαυξημένη πραγματικότητα σε εξωτερικούς χώρους, όταν αναπτύσσουν ένα οπτικό σύστημα πλοήγησης για πεζούς, το «Map-in-hat». Πρόκειται για ένα φορητό υπολογιστικό σύστημα, εφοδιασμένο με λογισμικό πλοήγησης, με μία διάφανη συσκευή θέασης, ψηφιακή πυξίδα και GPS (βλ. Εικόνα 1-7). Το σύστημα αυτό, ενώ αρχικά περιορίζεται μόνο σε εφαρμογές πλοήγησης, αργότερα εξελίσσεται στο σύστημα Tinmith, που χρησιμοποιείται σε πληθώρα εφαρμογών επαυξημένης πραγματικότητας [21].

Εικόνα 1-7. Σύστημα «Map-in-hat».

Πηγή: Thomas et al., 1998

Το 1999 ιδρύεται η εταιρία Total Immersion και εισέρχεται στην αγορά ως η πρώτη εταιρία παροχής λύσεων επαυξημένης πραγματικότητας. Η Total Immersion με το προϊόν της, D'Fusion, παρέχει σήμερα μία πλούσια παλέτα εργαλείων για τη δημιουργία εφαρμογών επαυξημένης πραγματικότητας [4], [26].

Την ίδια χρονιά, ο Hirokazu Kato αναπτύσσει τη βιβλιοθήκη για εφαρμογές επαυξημένης πραγματικότητας ARToolKit, η οποία παρέχεται στο κοινό ως βιβλιοθήκη ανοικτού κώδικα από το εργαστήριο HIT Lab (Human Interface Technology Laboratory) του Πανεπιστημίου της Ουάσιγκτον υπό την άδεια GPL (GNU General Public License). Η βιβλιοθήκη αυτή επιτρέπει τη λήψη βίντεο και την τοποθέτηση – σε πραγματικό χρόνο – εικονικών μοντέλων πάνω σε ασπρόμαυρους επίπεδους στόχους που εντοπίζονται στην εκάστοτε σκηνή, έτσι ώστε να ακολουθούν την κίνηση της κάμερας. Η βιβλιοθήκη ARToolKit αποτελεί τη βάση για πολλές εφαρμογές επαυξημένης πραγματικότητας που ακολουθούν και χρησιμοποιείται ευρέως ακόμα και σήμερα [4], [17], [27], [28].

Το 2000 οι Bruce Thomas et al. [29] δημιουργούν το πρώτο παιχνίδι επαυξημένης πραγματικότητας, το ARQuake, που αποτελεί μία παραλλαγή του διάσημου παιχνιδιού Quake για ηλεκτρονικό υπολογιστή (βλ. Εικόνα 1-8). Το παιχνίδι βασίζεται σε ένα σύστημα ανίχνευσης 6 βαθμών ελευθερίας και χρησιμοποιεί GPS, μία ψηφιακή πυξίδα και οπτική ανίχνευση των επίπεδων στόχων. Οι χρήστες εφοδιάζονται με ένα φορητό υπολογιστή στην πλάτη τους, ένα HMD και μία απλή συσκευή εισόδου [4], [21].

Εικόνα 1-8. Παράδειγμα εικόνας που προβάλλεται στο χρήστη του παιχνιδιού ARQuake από το HMD, όπου φαίνεται τόσο ο πραγματικός κόσμος όσο και ο κόσμος που παράγεται από τον υπολογιστή.

Πηγή: http://en.wikipedia.org/wiki/File:Example_of_ARQuake.jpg

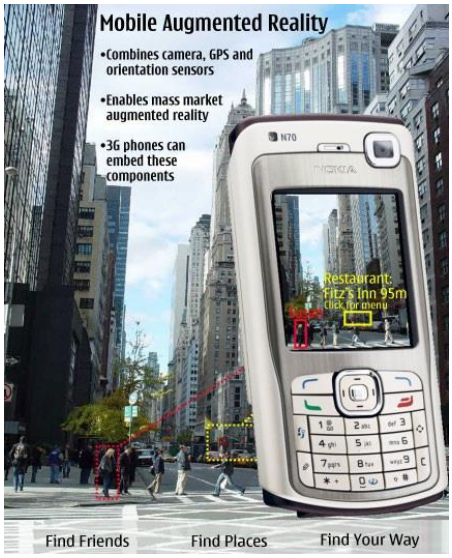
Έκτοτε σημειώνονται σημαντικά βήματα προόδου στον τομέα της επαυξημένης πραγματικότητας και όλο και περισσότεροι επιστήμονες, αλλά και εταιρίες δείχνουν ενδιαφέρον για την τεχνολογία αυτή, με αποτέλεσμα να αναπτύσσεται πληθώρα εφαρμογών. Από τα επιτεύγματα των ετών αυτών, έως το 2013, κρίνεται σκόπιμη η παράθεση κάποιων σταθμών - σημαντικών γεγονότων στην ιστορία της επαυξημένης πραγματικότητας.

Το 2001 δημιουργείται το πρώτο πρόγραμμα περιήγησης επαυξημένης πραγματικότητας, Real-World Wide Web (RWWW), από τους Bob Kooper και Blair MacIntyre [30]. Πρόκειται για μία εφαρμογή επαυξημένης πραγματικότητας που υπερθέτει δεδομένα από τον Παγκόσμιο Ιστό (World Wide Web) στον πραγματικό κόσμο, μέσω μίας συσκευής HMD. Οι υπερτιθέμενες πληροφορίες ανανεώνονται ανάλογα με τη θέση και τον προσανατολισμό του χρήστη.

Το 2002 διοργανώνεται το πρώτο διεθνές συνέδριο μικτής και επαυξημένης πραγματικότητας, International Symposium on Mixed Reality and Augmented Reality (ISMAR), το οποίο μέχρι σήμερα διοργανώνεται σε ετήσια βάση και δίνει τη δυνατότητα στους ερευνητές να ανταλλάσσουν ιδέες, σκέψεις και συμπεράσματα σχετικά με τους τομείς αυτούς [31]. Πέρα όμως από το ISMAR, αξίζει να αναφερθούν και οι πρόγονοί του: το IWAR (International Workshop on Augmented Reality), που πραγματοποιείται το 1998 και το 1999, το ISMR (International Symposium on Mixed Reality), που πραγματοποιείται το 1999 και το 2001, και το ISAR (International Symposium on Augmented Reality), που λαμβάνει χώρα το 2000 και το 2001 [32].

Το 2003 παρουσιάζεται το iLamps [33], το πρώτο σύστημα παρουσίασης ψηφιακής πληροφορίας πάνω σε πραγματικά μη επίπεδα αντικείμενα, με χρήση μικρού φορητού προβολέα. Την ίδια χρονιά, δημιουργείται το πρώτο εμπορικό παιχνίδι επαυξημένης πραγματικότητας για κινητά τηλέφωνα, το επονομαζόμενο Mozzies, το οποίο αναπτύσσεται για το κινητό τηλέφωνο Siemens SX1 [21].

Το 2004 παρουσιάζεται από τους Mathias Möhring et al. το πρώτο σύστημα ανίχνευσης τρισδιάστατων στόχων, σε κινητό τηλέφωνο, που υποστηρίζει τη διάκριση διαφορετικών τρισδιάστατων markers, και τη σωστή ενσωμάτωση τρισδιάστατων γραφικών σε ζωντανό βίντεο του πραγματικού κόσμου [34]. Το ίδιο έτος, παρουσιάζεται και το πρώτο παιχνίδι επαυξημένης πραγματικότητας για πολλούς χρήστες σε συσκευές χειρός, με το όνομα «The invisible train» [35].



Εικόνα 1-9. Εφαρμογή MARA από τη Nokia.

Πηγή: http://techblog.gr/techno/nokia-mara/

Το 2006 η Nokia εισάγει το πρόγραμμα MARA (Mobile Augmented Reality Applications project). Η πρωτότυπη αυτή εφαρμογή (βλ. Εικόνα 1-9) χρησιμοποιεί επιταχυνσιόμετρο, πυξίδα και GPS για τον υπολογισμό της (προσεγγιστικής) θέσης και του προσανατολισμού

του κινητού τηλεφώνου, με αποτέλεσμα να υπερθέτει σε ζωντανό βίντεο πληροφορίες για τα πραγματικά αντικείμενα, εκμεταλλευόμενη τις δυνατότητας του κινητού τηλεφώνου για ασύρματη σύνδεση στο διαδίκτυο, αφού αντλεί πληροφορίες από μία βάση δεδομένων [36], [37].

Το 2007 δημιουργείται η πρώτη διαφήμιση επαυξημένης πραγματικότητας σε κινητά τηλέφωνα για τη Wellington Zoo [38], ενώ το 2008 δημιουργείται το πρόγραμμα περιήγησης επαυξημένης πραγματικότητας Wikitude από την Αυστριακή εταιρία Mobilizy GmbH (βλ. Εικόνα 1-10). Πρόκειται για μία εφαρμογή επαυξημένης πραγματικότητας βάσει θέσης, που χρησιμοποιεί δεδομένα GPS, επιταχυνσιόμετρου και πυξίδας και την κάμερα του κινητού τηλεφώνου για να υπερθέσει στο βίντεο πληροφορίες από το διαδίκτυο για τον περιβάλλοντα χώρο, σε πραγματικό χρόνο. Ενώ δημιουργήθηκε για κινητά τηλέφωνα με λειτουργικό σύστημα Android, σήμερα υποστηρίζει πληθώρα λειτουργικών συστημάτων, ενώ χρησιμοποιεί ταυτόχρονα και τεχνολογία αναγνώρισης εικόνας για την υπέρθεση των κατάλληλων πληροφοριών στην κατάλληλη θέση [39].

Εικόνα 1-10. Εφαρμογή Wikitude.

Πηγή: http://www.wikitude.com/app/how-to-use-wikitude

Το 2009 η εταιρία SPRXmobile δημιουργεί το πρόγραμμα περιήγησης επαυξημένης πραγματικότητας βάσει θέσης Layar, μία προηγμένη παραλλαγή του Wikitude [21]. Σήμερα, αποτελεί μία ευρέως διαδεδομένη εφαρμογή που εκτός από την ικανότητά της να παρέχει πληροφορίες στους χρήστες για το μέρος στο οποίο βρίσκονται και τη γύρω περιοχή, επιχειρεί *να δώσει στον εκτυπωμένο κόσμο τη δυνατότητα του «κλικ»,* όπως επισημαίνεται στον επίσημο ιστότοπο του προϊόντος [40]. Με την εφαρμογή αυτή, δείχνοντας με το κινητό τηλέφωνο το περιεχόμενο περιοδικών, εφημερίδων και άλλων προϊόντων που συνοδεύονται από το σήμα του Layar, οι χρήστες βλέπουν επιπρόσθετο ψηφιακό υλικό, όπως για παράδειγμα βίντεο σχετικό με τη στατική εικόνα. Επίσης, παρέχεται και η δυνατότητα σε όποιον θέλει – επί πληρωμή – να προσθέσει ψηφιακό υλικό σε στατικές εικόνες, με το προϊόν Layar Creator.

Τέλος, πρέπει να σημειωθεί ότι από τη δεκαετία του 70 και μέχρι σήμερα, στρατιωτικές υπηρεσίες σε συνεργασία με διάφορα πανεπιστημιακά ιδρύματα και ερευνητικά κέντρα ασχολούνται με την αξιοποίηση των δυνατοτήτων της επαυξημένης πραγματικότητας, για εφαρμογές σχετικές με τις ανάγκες του στρατού, του ναυτικού και της αεροπορίας, σε θέματα διεξαγωγής πολεμικών επιχειρήσεων, εκπαίδευσης προσωπικού, επισκευής μηχανημάτων κ.ά.

1.3 Εφαρμογές

Στην ενότητα αυτή καταγράφονται οι κυριότεροι τομείς στους οποίους η επαυξημένη πραγματικότητα έχει εφαρμογή και αναφέρονται συνοπτικά διάφορα αντιπροσωπευτικά είδη εφαρμογών σε κάθε τομέα. Οι εφαρμογές αυτές μπορούν να χαρακτηριστούν ως εσωτερικού ή εξωτερικού χώρου. Επίσης, μπορούν να διακριθούν σε αυτές που δίνουν στο χρήστη ελευθερία κίνησης και σε εκείνες που προϋποθέτουν ότι αυτός παραμένει στην ίδια θέση. Ακόμη, μπορούν να διακριθούν σε εφαρμογές που δίνουν στο χρήστη μία επαυξημένη αντίληψη της πραγματικότητας και σε εκείνες που δημιουργούν ένα φανταστικό περιβάλλον. Σύμφωνα με την τελευταία αυτή διάκριση [4], οι πρώτες προσθέτουν πληροφορίες που του επιτρέπουν να έχει μία καλύτερη κατανόηση του περιβάλλοντος στο οποίο βρίσκεται, ενώ εκείνες που δημιουργούν ένα φανταστικό περιβάλλον του επιτρέπουν να δει πράγματα που δεν υπάρχουν στον πραγματικό κόσμο. Συναφής με τη διάκριση αυτή είναι και μία άλλη κατηγοριοποίηση των εφαρμογών [41], σύμφωνα με την οποία αυτές χωρίζονται σε εφαρμογές που υπερθέτουν πληροφορίες που δεν αποτελούν μέρος του φυσικού κόσμου, σε αυτές που απεικονίζουν επιπρόσθετες πληροφορίες με τέτοιο τρόπο ώστε να μην ξεχωρίζουν από το πραγματικό περιβάλλον και σε αυτές που απεικονίζουν πληροφορίες που βρίσκονται στο φυσικό περιβάλλον αλλά δεν είναι άμεσα παρατηρήσιμες.

1.3.1 Ιατρική

Ένα από τα πιο σημαντικά πεδία εφαρμογών επαυξημένης πραγματικότητας είναι εκείνο που σχετίζεται με τον τομέα της ιατρικής. Η τεχνολογία αυτή μπορεί να χρησιμοποιηθεί στη διάγνωση, στη θεραπεία μίας νόσου και στην εκτέλεση μίας χειρουργικής επέμβασης. Όσον αφορά στη διάγνωση, δεδομένα που συλλέγονται, για παράδειγμα, από μαγνητική τομογραφία, αξονική τομογραφία ή απεικονίσεις υπερήχων μπορούν να συνδυαστούν έτσι ώστε να υπερτεθούν σε πραγματικό χρόνο σε έναν πραγματικό ασθενή συγκεκριμένες ανατομικές πληροφορίες. Το γεγονός αυτό δίνει στο γιατρό τη δυνατότητα να δει στο εσωτερικό του ασθενούς κατά την ώρα της εξέτασης και συμβάλλει στην ολοκλήρωση της διάγνωσης με ένα τρόπο περισσότερο καινοτόμο σε σχέση με την απλή εξέταση ξεχωριστών εικονικών δεδομένων (βλ. Εικόνα 1-11) [1], [4], [42], [43]. Ακόμη, τέτοιου είδους

πληροφορίες είναι πολύ σημαντικές στον τομέα της ελάχιστα επεμβατικής χειρουργικής, διότι παρέχουν μία εικόνα του εσωτερικού του ασθενούς, χωρίς την ανάγκη για μεγαλύτερες τομές [5].

Εικόνα 1-11. Παράδειγμα επαυξημένης πραγματικότητας στην ιατρική.

Πηγή: http://wg11.sc29.org/augmentedReality/?page_id=1154

Κατά τη διάρκεια μίας επέμβασης, η επαυξημένη πραγματικότητα θα μπορούσε να χρησιμοποιηθεί για την αναγνώριση και επισήμανση οργάνων του ασθενούς και μελών του σώματός του στα οποία θα πρέπει – ή θα απαγορεύεται – να επέμβουν οι γιατροί. Επίσης, ο γιατρός μπορεί να παρατηρεί την κοιλιά της εγκύου και να βλέπει την τρισδιάστατη αναπαράσταση του εμβρύου. Μία τέτοια προσπάθεια έγινε στο Πανεπιστήμιο της Νότιας Καρολίνας στο Chapel Hill από μία ερευνητική ομάδα, η οποία διεξήγαγε δοκιμαστικές σαρώσεις της μήτρας εγκύου με αισθητήρα υπερήχων, δημιουργώντας μία τρισδιάστατη αναπαράσταση του εμβρύου και εμφανίζοντάς τη σε μία συσκευή HMD [4], [5]. Παράλληλα, εικονικές οδηγίες θα μπορούσαν να υπενθυμίζουν στο χειρούργο τα βήματα μίας εγχείρισης.

1.3.2 Εκπαίδευση

Η επαυξημένη πραγματικότητα μπορεί να εφαρμοστεί στον τομέα της εκπαίδευσης, δίνοντας την ευκαιρία στους μαθητές να αλληλεπιδρούν σε πραγματικό χρόνο με το πραγματικό και το επαυξημένο περιβάλλον (βλ. Εικόνα 1-12). Για παράδειγμα, κάθε σχολικό εγχειρίδιο και γενικότερα εκπαιδευτικό υλικό ανάγνωσης μπορεί να περιλαμβάνει ενσωματωμένους επίπεδους στόχους, τέτοιους ώστε όταν αυτοί εντοπίζονται από μία συσκευή επαυξημένης πραγματικότητας, να υπερτίθενται στο πραγματικό περιβάλλον του μαθητή κείμενο, γραφικά στοιχεία, βίντεο και ήχος. Έτσι, το διάβασμα γίνεται πιο ενδιαφέρον και περισσότερο αποτελεσματικό [1], [4].

Επίσης, μπορεί να εφαρμοστεί και σε πιο εξειδικευμένη εκπαίδευση, ανώτερου επιπέδου, ή και στην απόκτηση δεξιοτήτων. Για παράδειγμα, φοιτητές ιατρικής μπορούν να επωφεληθούν από την τεχνολογία αυτή, παρατηρώντας μέρη της ανθρώπινης ανατομίας με τρόπο που ήταν αδύνατος στο παρελθόν (βλ. Εικόνα 1-13). Ακόμη, οι πυροσβέστες θα

μπορούσαν να μάθουν πώς να συμπεριφέρονται σε ορισμένου είδους οπτικοποιημένες φωτιές [1], [4].

Εικόνα 1-12. Η επαυξημένη πραγματικότητα δίνει τη δυνατότητα στο μαθητή να αλληλεπιδράσει με τη δομή ενός μορίου, το οποίο εμφανίζεται σε αυτόν σε μία εικόνα κάμερας όταν ειδικό marker που κρατά στο χέρι του εντοπίζεται από την κάμερα.

Πηγή: http://www.arlab.com/blog/augmented-reality-in-education

Εικόνα 1-13. Παράδειγμα εφαρμογής επαυξημένης πραγματικότητας στον τομέα της εκπαίδευσης στην ιατρική.

Πηγή: http://technoccult.net/archives/2010/01/11/augmented-reality-medical-app

1.3.3 Στρατός

Η επαυξημένη πραγματικότητα εδώ και πολλά χρόνια έχει στρατιωτικές εφαρμογές. Τα Head-Up Displays (HUDs) και τα Helmet-Mounted Displays (HMDs) στα μαχητικά αεροσκάφη παρέχουν στους πιλότους οπτικές ενδείξεις για τους στόχους, τις απειλές και άλλες πληροφορίες – όπως την ταχύτητα ή το υψόμετρο – ενσωματωμένες στον πραγματικό κόσμο. Όπως υπονοεί και το όνομά τους, τα πρώτα αποτελούν μία διάφανη συσκευή που τοποθετείται στο οπτικό πεδίο των πιλότων και δείχνει τις επιπρόσθετες πληροφορίες, επιτρέποντάς τους να έχουν «το κεφάλι ψηλά» (βλ. Εικόνα 1-14), ενώ τα HMDs προσαρμόζονται στο κράνος τους και προβάλλουν πληροφορίες μπροστά στα μάτια τους. Έτσι, αυτοί δεν αναγκάζονται να στρέψουν το βλέμμα τους για να δουν τις πληροφορίες που χρειάζονται σε άλλα όργανα, αποφεύγοντας έτσι τον κίνδυνο απόσπασης της προσοχής τους και συνακόλουθου ενδεχόμενου λανθασμένου χειρισμού [44], [45], [46].

Εικόνα 1-14. Παράδειγμα Head-Up Display (HUD).

Πηγή: https://en.wikipedia.org/wiki/File:HUD_view.jpg

Επίσης, σε περίπτωση μάχης, ένα σύστημα επαυξημένης πραγματικότητας θα μπορούσε να προβάλλει χρήσιμα δεδομένα σε ειδικά γυαλιά των στρατιωτών σε πραγματικό χρόνο. Για παράδειγμα, άνθρωποι ή διάφορα αντικείμενα θα μπορούσαν να επισημαίνονται με κατάλληλα σήματα, τα οποία να δείχνουν αν πρόκειται για πιθανούς εχθρούς ή κινδύνους [1].

Τα τελευταία χρόνια, μάλιστα, έχοντας καταστεί κατανοητή η μεγάλη συμβολή της επαυξημένης πραγματικότητας σε στρατιωτικές εφαρμογές, ερευνητικά προγράμματα έχουν χρηματοδοτηθεί από την υπηρεσία DARPA (Defense Advanced Research Projects Agency) του Υπουργείου Άμυνας των Ηνωμένων Πολιτειών, το ερευνητικό εργαστήριο του Στρατού των Ηνωμένων Πολιτειών ARL (Army Research Laboratory) και το Γραφείο Ναυτικών Ερευνών των Ηνωμένων Πολιτειών ONR (Office of Naval Research) για την ανάπτυξη της τεχνολογίας αυτής. Ο στόχος των παραπάνω προγραμμάτων είναι η αύξηση της επίγνωσης της κατάστασης από την πλευρά των στρατιωτών και των διοικητών, τόσο στην ύπαιθρο όσο και κάτω από το νερό, μέσω της τεχνολογίας της επαυξημένης πραγματικότητας [44].

1.3.4 Ψυχαγωγία - ενημέρωση - αθλητισμός

Η επαυξημένη πραγματικότητα έχει διεισδύσει στον τομέα της ψυχαγωγίας, δείχνοντας τις συναρπαστικές δυνατότητές της. Κυκλοφορούν πολλά παιχνίδια επαυξημένης πραγματικότητας, τα οποία προσελκύουν το κοινό λόγω της δυνατότητας που του προσφέρουν να αλληλεπιδρά σε πραγματικό χρόνο τόσο με πραγματικά όσο και με εικονικά αντικείμενα. Μπορεί να απευθύνονται σε έναν ή περισσότερους χρήστες και να χρησιμοποιούν κινητά τηλέφωνα, υπολογιστές, φορητούς ή μη, ή HMDs, προσφέροντας μοναδικές εμπειρίες στους χρήστες.

Πέρα όμως από τα παιχνίδια, υπάρχουν και λογοτεχνικά βιβλία επαυξημένης πραγματικότητας για παιδιά αλλά και κόμικς [47], όπου τρισδιάστατες σκηνές της ιστορίας

παρουσιάζονται πάνω από τις πραγματικές σελίδες και καθιστούν πιο ενδιαφέρουσα την εμπειρία του διαβάσματος.

Κοινή είναι πλέον η χρήση της στα καθημερινά δελτία καιρού. Ο τηλεθεατής. μέσω του δέκτη του, βλέπει τον παρουσιαστή μπροστά από το χάρτη πρόγνωσης του καιρού, ο οποίος στην πραγματικότητα δεν υπάρχει. Με χρήση της επαυξημένης πραγματικότητας, το φόντο μπροστά από το οποίο βρίσκεται ο παρουσιαστής, αντικαθίσταται σε πραγματικό χρόνο από το χάρτη, ο οποίος μάλιστα περιλαμβάνει κινούμενα γραφικά στοιχεία.

Η επαυξημένη πραγματικότητα χρησιμοποιείται και στον αθλητισμό και η τεχνολογία αυτή εφαρμόζεται ήδη κατά την αναμετάδοση αγώνων. Ένα παράδειγμα αποτελεί η παραγόμενη από υπολογιστή κίτρινη γραμμή που φαίνεται στα γήπεδα των ποδοσφαιρικών αγώνων στους τηλεοπτικούς δέκτες (βλ. Εικόνα 1-15) και είναι ορατή μόνο από τους τηλεθεατές, διευκολύνοντάς τους στην κατανόηση τυχούσας παράβασης των κανονισμών του παιχνιδιού [1], [48]. Επίσης, η επαυξημένη πραγματικότητα χρησιμοποιείται σε πληθώρα αθλητικών γεγονότων για να δείξει διαφημίσεις να υπερτίθενται σε συγκεκριμένες περιοχές της μεταδιδόμενης εικόνας [1], [7].

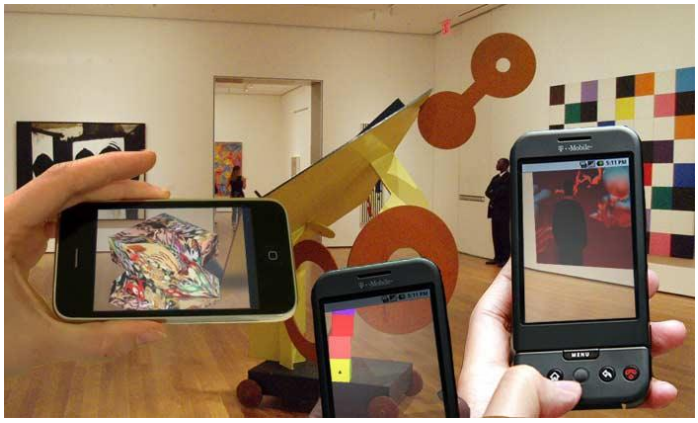
Εικόνα 1-15. Η τεχνολογία της επαυξημένης πραγματικότητας δημιουργεί την απεικονιζόμενη κίτρινη γραμμή σε γήπεδα ποδοσφαίρων, που φαίνεται μόνο στους τηλεοπτικούς δέκτες.

Πηγή: http://www.pcmag.com/encyclopedia/term/38186/augmented-reality

1.3.5 Τέχνη - πολιτισμός - τουρισμός

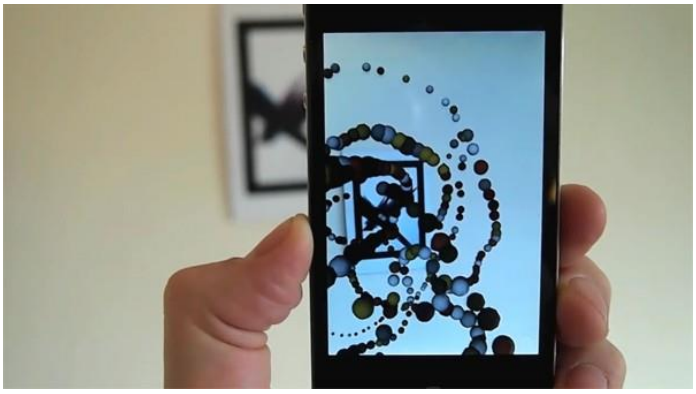
Η τεχνολογία της επαυξημένης πραγματικότητας μπορεί να διεισδύσει και στον κόσμο της τέχνης. Έτσι, η ιδέα της επαύξησης του πραγματικού κόσμου δύναται να εφαρμοστεί σε μουσεία και αίθουσες τέχνης, όπου οι επισκέπτες μπορούν να βλέπουν «κρυμμένα» έργα τέχνης μέσω ενός κινητού τηλεφώνου. Μία τέτοια έκθεση πραγματοποιήθηκε στο Μουσείο Μοντέρνας Τέχνης της Νέας Υόρκης το 2010 (βλ. Εικόνα 1-16) [49]. Ακόμη, είναι δυνατή η προβολή πληροφοριών – οπτικών ή ακουστικών – στους επισκέπτες των μουσείων, σχετικών με τα εκθέματα και τα έργα τέχνης [44].

Επιπλέον, εφαρμογές κινητών τηλεφώνων, όπως η υπάρχουσα εφαρμογή Konstruct (βλ. Εικόνα 1-17), βοηθούν το χρήστη να δημιουργήσει ένα έργο τέχνης, το οποίο είναι ορατό σε περιβάλλον επαυξημένης πραγματικότητας από την οθόνη της συσκευής, μόνο όταν αυτή εντοπίσει το αντίστοιχο marker [50].



Εικόνα 1-16. Η επαυξημένη πραγματικότητα στο Μουσείο Μοντέρνας Τέχνης της Νέας Υόρκης.

Πηγή: http://www.sndrv.nl/moma



Εικόνα 1-17. Ο χρήστης της εφαρμογής Konstruct δημιουργεί και βλέπει ένα έργο τέχνης στην οθόνη της κινητής συσκευής του σε περιβάλλον επαυξημένης πραγματικότητας.

Πηγή: http://downloadsquad.switched.com/2011/04/05/konstruct-for-ios-creates-generative-augmented-reality-art-with

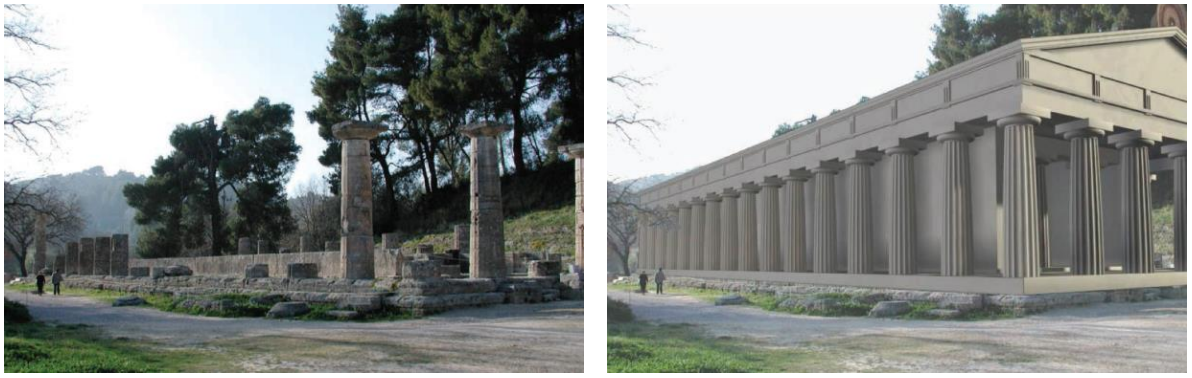
Επίσης, η επαυξημένη πραγματικότητα μπορεί να χρησιμοποιηθεί σε θεατρικές παραστάσεις, συναυλίες και κονσέρτα, όπως έχει ήδη γίνει στο πλαίσιο του «Duran Duran Project», όπου το συγκρότημα δημιούργησε εφέ επαυξημένης πραγματικότητας για το κοινό με χρήση του ARToolKit [51].

Σημαντικές είναι και οι δυνατότητές της στον τομέα του τουρισμού, όπου μπορεί να χρησιμοποιηθεί για τη βελτίωση της παρατήρησης αξιοθέατων, τοπίων, αλλά και αρχαιολογικών χώρων. «Ξεκλειδώνοντας» σημαντικές πληροφορίες για το εκάστοτε μέρος, ο τουρίστας έχει πλέον την ευκαιρία μέσω της τεχνολογίας αυτής να εξερευνήσει σημαντικές λεπτομέρειες για αυτό [4].

Επιπλέον, πολύ σημαντική πτυχή της είναι η δυνατότητα που δίνει για εικονική ανακατασκευή των αρχαιολογικών χώρων όπως ακριβώς υπήρχαν στην αρχαιότητα. Μία αξιοσημείωτη τέτοια δουλειά πραγματοποιήθηκε από Έλληνες και ξένους επιστήμονες [52] στον αρχαιολογικό χώρο της Ολυμπίας στην Πελοπόννησο. Οι επισκέπτες εξοπλίζονται με μία συσκευή HMD και ένα φορητό υπολογιστή στην πλάτη τους, ή με ένα pen-tablet PC ή ακόμα και με μία κινητή συσκευή χειρός τύπου palmtop και έχουν τη δυνατότητα να δουν τα

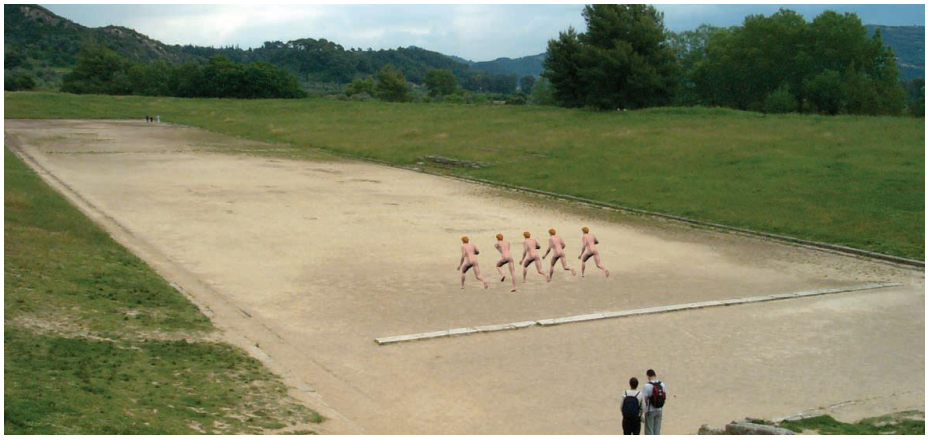
τρισδιάστατα μνημεία όπως υπήρχαν στο παρελθόν, αντί των σημερινών ερειπίων τους (βλ. Εικόνα 1-18), αλλά και εικονικούς αθλητές («avatars») που επιδίδονται σε αθλήματα (βλ. Εικόνα 1-19).

Πέρα από την εικονική ανακατασκευή αρχαιολογικών χώρων, η επαυξημένη πραγματικότητα μπορεί να συντελέσει στην εικονική ανακατασκευή κτηρίων και στην αναπαράσταση της όψης τους όπως ήταν στο παρελθόν, γεγονός το οποίο μπορεί να κινήσει το ενδιαφέρον τουριστών αλλά και άλλων ανθρώπων.



Εικόνα 1-18. Ερείπια του Ναού της Ήρας στη σημερινή τους κατάσταση [αριστερά] και επαυξημένος Ναός με το τρισδιάστατο μοντέλο του να υπερτίθεται σε ζωντανό βίντεο [δεξιά].

Πηγή: Vlahakis et al., 2002



Εικόνα 1-19. Εικονικοί αθλητές συναγωνίζονται στο στάδιο της Αρχαίας Ολυμπίας.

Πηγή: Vlahakis et al., 2002

1.3.6 Πλοήγηση

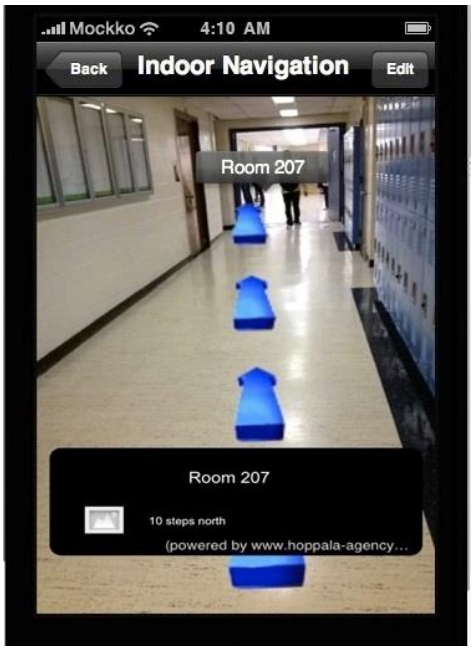
Η επαυξημένη πραγματικότητα μπορεί να χρησιμοποιηθεί στην πλοήγηση και να βελτιώσει σημαντικά την αποτελεσματικότητα απλών εφαρμογών της κατηγορίας αυτής. Για παράδειγμα, σε εφαρμογές πλοήγησης που απευθύνονται σε οδηγούς αυτοκινήτων και αξιοποιούν τις δυνατότητες της επαυξημένης πραγματικότητας, οι τελευταίοι ακολουθούν την πορεία που απεικονίζεται στην εικόνα του πραγματικού κόσμου μέσω κατάλληλων γραφικών και δεν απαιτείται να συμβουλεύονται κάποιο χάρτη, ο οποίος ενδεχομένως να τους

δυσκολεύει, αποσπώντας τους περισσότερο την προσοχή από την οδήγηση (βλ. Εικόνα 1-20) [45]. Βέβαια, ο κίνδυνος που προκύπτει από το γεγονός ότι ο οδηγός μπορεί να κοιτάει την οθόνη του κινητού του όταν οδηγεί παραμένει και στις δύο περιπτώσεις.

Εικόνα 1-20. Εικόνες από την εφαρμογή Wikitude Drive.

Πηγές: http://eurodroid.com/2010/05/20/google-io-2010-wikitude-drive-opening-to-2000-us-test-drivers, http://android.appstorm.net/roundups/utilities-roundups/30-apps-to-digitize-your-car-journey

Ο συνδυασμός της επαυξημένης πραγματικότητας και των υπηρεσιών βάσει θέσης (Location Based Services - LBS) μπορεί να έχει ως αποτέλεσμα σημαντικές εφαρμογές, που απευθύνονται όχι μόνο σε οδηγούς, αλλά και σε πεζούς. Αυτές, υπερθέτουν στο πραγματικό περιβάλλον οδηγίες για την εύρεση σημείων ενδιαφέροντος, όπως είναι μέρη για φαγητό, ποτό, κ.λπ. [4]. Επίσης, η επαυξημένη πραγματικότητα μπορεί να χρησιμοποιηθεί και σε περιπτώσεις πλοήγησης σε εσωτερικό χώρο, όπως φαίνεται και στην Εικόνα 1-21.



Εικόνα 1-21. Η επαυξημένη πραγματικότητα σε εφαρμογή πλοήγησης σε εσωτερικό χώρο.

Πηγή: http://www.behance.net/gallery/NYU-Wayfarer/4528801

1.3.7 Εμπόριο και διαφήμιση

Πολλές εταιρίες εκμεταλλεύονται τις δυνατότητες της επαυξημένης πραγματικότητας για να προωθήσουν τα προϊόντα τους. Ήδη σήμερα, υπάρχουν αρκετές εταιρίες που επιτρέπουν σε υποψήφιους πελάτες να «δοκιμάζουν» σε πραγματικό χρόνο τα προϊόντα τους μέσω του Διαδικτύου. Αυτοί κάθονται μπροστά από την κάμερα του υπολογιστή τους και βλέπουν τον εαυτό τους στην οθόνη να φορά το ρολόι ή τα γυαλιά, για παράδειγμα, που επέλεξαν.

Άλλες εταιρίες χρησιμοποιούν διαφημίσεις περιοδικών ή εφημερίδων σε συνδυασμό με την τεχνολογία της επαυξημένης πραγματικότητας για να παρέχουν στο χρήστη μία τρισδιάστατη – ή επαυξημένη με διάφορες πληροφορίες και άλλα γραφικά στοιχεία – θέα του προϊόντος που διαφημίζεται ή ακόμη και βίντεο, μέσω εφαρμογής σε κινητό τηλέφωνο. Μία τέτοια διαφημιστική καμπάνια έκανε και η Nissan [53], με χρήση της εφαρμογής Layar, η οποία είχε μεγάλη ανταπόκριση από το κοινό.

Άλλες χρήσεις της επαυξημένης πραγματικότητας στον τομέα του εμπορίου και της διαφήμισης είναι η δυνατότητα που δίνει στους πελάτες να δουν ένα προϊόν μέσα από το πακέτο συσκευασίας του, όπως επιχείρησε με μεγάλη επιτυχία η Lego. Οι πελάτες – και κυρίως τα μικρά παιδιά που συναρπάζονται από τις δυνατότητες της νέας τεχνολογίας – στέκονται μπροστά από το επονομαζόμενο «Digital Box», όπως φαίνεται και στην Εικόνα 1-22, και μπορούν να δουν το τρισδιάστατο μοντέλο του προϊόντος που κρατούν στα χέρια τους από όλες τις γωνίες, με κάθε λεπτομέρεια.

Εικόνα 1-22. Επαυξημένη εικόνα από το «Digital Box» της Lego.

Πηγή: http://emalliab.wordpress.com/tag/lego

1.3.8 Αρχιτεκτονική - Σχεδιασμός εσωτερικού χώρου

Η επαυξημένη πραγματικότητα μπορεί να βοηθήσει στην οπτικοποίηση κατασκευαστικών έργων κατά το σχεδιασμό τους, καθώς μέσω της τεχνολογίας αυτής τρισδιάστατα μοντέλα

προγραμματισμένων κτηρίων ή άλλων έργων μπορούν να υπερτεθούν στο μέρος όπου σχεδιάζεται να κατασκευαστούν. Το γεγονός αυτό επιτρέπει να εξεταστεί ο αντίκτυπος που θα έχει το νέο έργο στην περιοχή, δηλαδή να διερευνηθεί αν αυτό ταιριάζει με το γύρω χώρο και εντάσσεται ομαλά σε αυτόν.

Επίσης, η επαυξημένη πραγματικότητα μπορεί να έχει εφαρμογή στο σχεδιασμό και στη διακόσμηση εσωτερικών χώρων, για παράδειγμα με την τοποθέτηση εικονικών επίπλων σε ένα πραγματικό δωμάτιο, με σκοπό τη μετέπειτα πραγματική τοποθέτησή τους στην κατάλληλη θέση. Στην περίπτωση αυτή, δεν απαιτείται να καταβληθεί προσπάθεια για τη μετακίνησή τους κατά την πραγματοποίηση των πιθανών συνδυασμών [4].

1.3.9 Υποστήριξη εργασιών

Η υποστήριξη πολύπλοκων εργασιών του ανθρώπου, όπως της κατασκευής, συντήρησης και επισκευής, είναι τομέας στον οποίο η επαυξημένη πραγματικότητα παρουσιάζει μεγάλες δυνατότητες. Η τεχνολογία αυτή μπορεί να συμβάλλει στη μείωση του χρόνου εκπαίδευσης των εργαζομένων, στην επιτάχυνση της ολοκλήρωσης των εργασιών και στην ελάττωση των λαθών, οδηγώντας έτσι σε εξοικονόμηση χρημάτων [44]. Η λογική τέτοιων εφαρμογών είναι η υπέρθεση στο οπτικό πεδίο των εργαζομένων, και γενικότερα αυτών που χρειάζονται βοήθεια στην εκτέλεση ενός καθήκοντος, πληροφοριών, οδηγιών και γραφικών στοιχείων. Παραδείγματα τέτοιων εφαρμογών, που ήδη αναφέρθηκαν, είναι το σύστημα των Caudell και Mizell για την εταιρία Boeing και το σύστημα KARMA των Feiner, MacIntyre και Seligmann.

1.3.10 Άλλες εφαρμογές

Εκτός από τους παραπάνω τομείς, στους οποίους η επαυξημένη πραγματικότητα μπορεί να χρησιμοποιηθεί, συνδέεται και με πληθώρα άλλων εφαρμογών. Για παράδειγμα, μπορεί να χρησιμοποιηθεί στην αυτόματη μετάφραση, κατά την οποία η μεταφρασμένη φράση αντικαθιστά την πρωτότυπη στο πραγματικό περιβάλλον όπου αυτή συναντάται (βλ. Εικόνα 1-23). Επίσης, μπορεί να βοηθήσει στη διευκόλυνση της συνεργασίας μεταξύ απομακρυσμένων μελών μίας ομάδας, μέσω συνεδριών με πραγματικούς και εικονικούς συμμετέχοντες [1]. Ακόμη, μπορεί να χρησιμοποιηθεί και σε πιο εξειδικευμένους τομείς, όπως στη ρομποτική, για προγραμματισμό της κίνησης ενός ρομπότ [5]. Σε μία τέτοια εφαρμογή, είναι προτιμότερο η κίνηση μίας εικονικής έκδοσης του ρομπότ να προηγείται από την κίνηση του πραγματικού ρομπότ, για την εκ των προτέρων γνώση των συνεπειών που θα προκαλέσει στο γύρω περιβάλλον και την ενδεχόμενη αλλαγή του προγραμματισμού.

Εικόνα 1-23. Εφαρμογή μετάφρασης επαυξημένης πραγματικότητας Word Lens. Αριστερά: πρωτότυπη εικόνα. Δεξιά: εικόνα επαυξημένης πραγματικότητας.

Πηγή: http://laughingsquid.com/word-lens-translate-printed-words-using-an-iphones-video-camera

1.4 Τεχνολογίες Θέασης της Επαυξημένης Σκηνής

Η τεχνολογία θέασης αποτελεί ένα από τα βασικότερα συστατικά στοιχεία των συστημάτων επαυξημένης πραγματικότητας, τα υπόλοιπα εκ των οποίων είναι το υπολογιστικό σύστημα, η κάμερα ή κάποιο άλλο οπτικό μέσο, κατάλληλο λογισμικό, η πραγματική σκηνή και τα εικονικά αντικείμενα και – ανάλογα με την εφαρμογή – διάφοροι αισθητήρες (GPS, πυξίδα, επιταχυνσιόμετρα, γυροσκόπια), markers ή πρότυπα προς αναγνώριση, διακομιστής περιεχομένου, υπηρεσίες διαδικτύου και πιθανώς και άλλα. Οι τεχνολογίες θέασης των συστημάτων επαυξημένης πραγματικότητας μπορούν να χωριστούν σε τρεις κύριες κατηγορίες [4], [45]. Αυτές είναι οι «φορετές» συσκευές στο κεφάλι ή σε μέρος του κεφαλιού του χρήστη, οι φορητές συσκευές χειρός και οι συσκευές απεικόνισης των συστημάτων χωρικής επαυξημένης πραγματικότητας (Spatial Augmented Reality - SAR). Στην ενότητα αυτή παρουσιάζονται τα τρία αυτά είδη τεχνολογιών θέασης της επαυξημένης σκηνής.

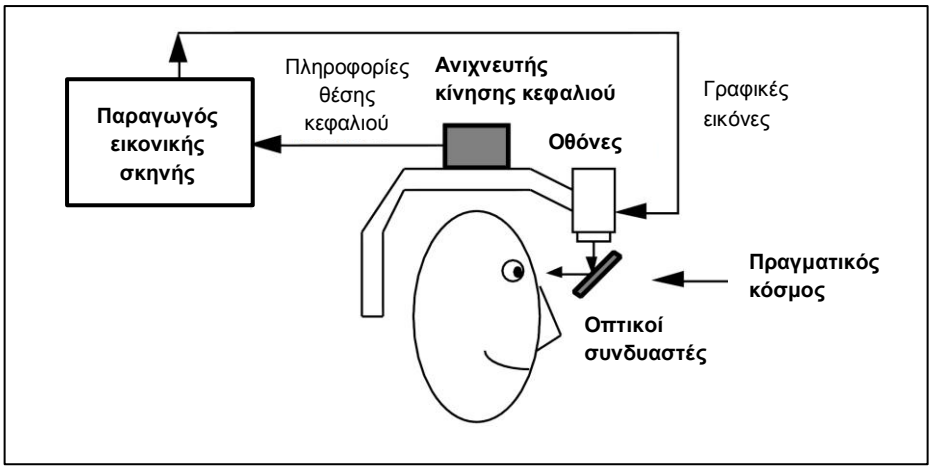
1.4.1 Φορετές συσκευές

Οι συσκευές αυτές, γνωστές ως Head-Worn Displays (HWD) ή και Head-Attached Displays, φοριούνται στο κεφάλι του χρήστη ή σε μέρος του (π.χ. οι ειδικοί φακοί επαφής που επαυξάνουν την πραγματικότητα). Στις παραγράφους που ακολουθούν περιγράφονται οι κυριότερες κατηγορίες των συσκευών αυτών.

1.4.1.1 Προσαρτημένα επί κεφαλής (ή επί κράνους) συστήματα παρουσίασης

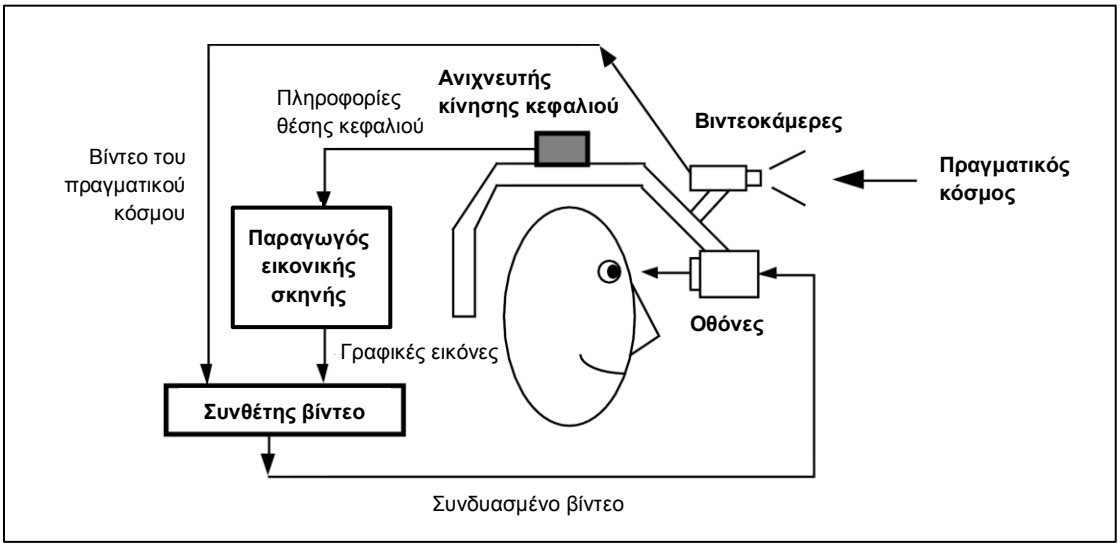
Ένα σύστημα Head-Mounted Display (HMD) είναι μία συσκευή προσαρτημένη στο κεφάλι του χρήστη (βλ. Εικόνα 1-26, αριστερά) που συνδυάζει την εικόνα του πραγματικού κόσμου με εικονικά αντικείμενα - πληροφορίες, μέσω χρήσης οπτικής ή βίντεο τεχνολογίας [5], [54].

Τα οπτικά συστήματα παρουσίασης προσαρτημένα στο κεφάλι του χρήστη (optical see-through HMDs) τοποθετούν μπροστά από τα μάτια του «οπτικούς συνδυαστές» (συνήθως ημιδιαπερατά και ημιανακλαστικά κάτοπτρα), που του επιτρέπουν να δει τόσο το πραγματικό περιβάλλον γύρω του, όσο και τις παραγόμενες από υπολογιστή εικόνες που ανακλούνται μέσω των κατόπτρων από τις οθόνες που είναι προσαρτημένες στο κεφάλι του (βλ. Εικόνα 1-24).



Εικόνα 1-24. Εννοιολογικό διάγραμμα οπτικού συστήματος HMD.

Πηγή: Azuma, 1997



Εικόνα 1-25. Εννοιολογικό διάγραμμα βίντεο συστήματος HMD.

Πηγή: Azuma, 1997

Τα HMDs που χρησιμοποιούν βίντεο τεχνολογία (video see-through HMDs), συνδυάζουν ένα «κλειστό» HMD (closed-view HMD[1]) και μία ή δύο βιντεοκάμερες προσαρτημένες στο

[1] Τα κλασικά closed-view HMDs δεν επιτρέπουν την άμεση παρατήρηση του πραγματικού κόσμου, σε αντίθεση με τα see-through HMDs, τα οποία επιτρέπουν στο χρήστη να δει τον πραγματικό κόσμο, συνδυαζόμενο με εικονικά αντικείμενα παραγόμενα από υπολογιστή.

κεφάλι του χρήστη, οι οποίες του παρέχουν μία θέα του πραγματικού κόσμου. Το βίντεο από τις κάμερες αυτές και οι παραγόμενες εικονικές σκηνές συνδυάζονται μεταξύ τους και το αποτέλεσμα στέλνεται στις οθόνες που βρίσκονται μπροστά από τα μάτια του χρήστη στο closed-view HMD (βλ. Εικόνα 1-25).

Το HMD αναφέρεται και ως Helmet-Mounted Display (βλ. Εικόνα 1-26, κέντρο) και προσαρμόζεται στο κράνος του χρήστη κυρίως για εφαρμογές στην αεροπορία και ιδιαίτερα σε μαχητικά αεροσκάφη [55].

Νεότερα από τα HMDs είναι τα προσαρτημένα επί κεφαλής συστήματα παρουσίασης μέσω προβολής, γνωστά ως Head-Mounted Projective Displays (HMPDs) (βλ. Εικόνα 1-26, δεξιά). Τα κυριότερα μέρη των συστημάτων αυτών είναι ένας μικροσκοπικός προβολικός φακός στερεωμένος στο κεφάλι του χρήστη και οπισθοανακλαστικές (retroreflective) επιφάνειες που τοποθετούνται μπροστά από αυτόν. Οι ακτίνες φωτός που προσπίπτουν στις επιφάνειες αυτές εκτρέπονται προς την κατεύθυνση πρόσπτωσής τους και δεν διαχέονται, όπως συμβαίνει με τα συμβατικά προβολικά συστήματα, με αποτέλεσμα να παράγουν φωτεινότερες εικόνες [56], [57], [58].

Εικόνα 1-26. Head-Mounted Display [αριστερά], Helmet-Mounted Display [κέντρο] και Head-Mounted Projective Display [δεξιά].

Πηγές: https://en.wikipedia.org/wiki/Head-mounted_display, http://www.baesystems.com/product/BAES_055039, http://3dvis.optics.arizona.edu/research/research.html

1.4.1.2 Συσκευές απεικόνισης στον αμφιβληστροειδή του ματιού (VRDs)

Οι συσκευές αυτές, γνωστές ως Virtual Retinal Displays (VRDs) ή και Retinal Scan Displays (RSDs), σχηματίζουν εικόνες απευθείας στον αμφιβληστροειδή χιτώνα του ματιού, με τη χρήση συνήθως πηγής φωτός λέιζερ χαμηλής ισχύος. Με τη βοήθειά τους, οι χρήστες βλέπουν τον πραγματικό κόσμο επαυξημένο με επιπρόσθετες πληροφορίες, οι οποίες «αιωρούνται» μπροστά τους [3], [59], [60].

1.4.1.3 Ειδικά γυαλιά

Οι επαυξημένες σκηνές μπορούν να θεαθούν και με ειδικές συσκευές που θυμίζουν γυαλιά οράσεως ή ηλίου (βλ. Εικόνα 1-27). Κάποιες συσκευές χρησιμοποιούν κάμερες που «αιχμαλωτίζουν» τη θέα του πραγματικού κόσμου και εμφανίζουν αντί αυτής την

επαυξημένη θέα μέσω των προσοφθάλμιων φακών. Σε άλλα γυαλιά επαυξημένης πραγματικότητας, οι πληροφορίες που συμπληρώνουν το φυσικό περιβάλλον προβάλλονται μέσω των επιφανειών των φακών των γυαλιών ή ανακλώνται από τις επιφάνειές τους. Ειδικά «ψηφιακά γυαλιά» που σχεδίασε ο Steve Mann το 1981 και βελτίωσε το 1999, με το όνομα «EyeTap», γνωστά και ως γυαλιά 4ης γενιάς, έχουν ως αποτέλεσμα το μάτι να λειτουργεί τόσο ως κάμερα, για να αιχμαλωτίζει την πραγματική σκηνή που είναι διαθέσιμη γύρω του, όσο και ως οθόνη, για να υπερθέτει μία εικόνα παραγόμενη από υπολογιστή στην πραγματική σκηνή που είναι διαθέσιμη στο μάτι [1], [61], [62].

Η τεχνολογία των γυαλιών επαυξημένης πραγματικότητας αναπτύσσεται συνεχώς τα τελευταία χρόνια για να παρέχει στο χρήστη μία συνεχή εμπειρία επαυξημένης πραγματικότητας [1]. Στο πλαίσιο του προγράμματος «Project Glass», η Google αναπτύσσει την επονομαζόμενη συσκευή «Google Glass», η οποία υπολογίζεται ότι θα κάνει την είσοδό της στην αγορά το 2014 και αποτελεί μία πολλά υποσχόμενη συσκευή όχι μόνο επαυξημένης πραγματικότητας, αλλά με πολλές ακόμα δυνατότητες [63]. Ωστόσο, υπάρχουν ήδη γυαλιά επαυξημένης πραγματικότητας, όπως εκείνα της εταιρίας Vuzix ήδη από το φθινόπωρο του 2012 [64].

Εικόνα 1-27. Παραδείγματα γυαλιών επαυξημένης πραγματικότητας. Αριστερά: EyeTap (2η έκδοση, 1999). Κέντρο: Vuzix SMART Glasses. Δεξιά: Google Glass.

Πηγές: http://www.glogger.mobi/userpub2.php?username=mann&i=2012_04_16_14_45_33_24654000, http://www.vuzix.com/augmented-reality/docs/SMART-Glasses-Brochure.pdf, http://www.buygoogleglassonline.com/google-glasses-faq/

1.4.1.4 Ειδικοί φακοί επαφής

Τα τελευταία χρόνια, αναπτύσσονται βιονικοί φακοί επαφής που δίνουν τη δυνατότητα στο χρήστη να δει όχι μόνο τον πραγματικό κόσμο, αλλά και επιπρόσθετες πληροφορίες. Τα στοιχεία που θα επαυξήσουν την πραγματικότητα του χρήστη μπορεί να περιλαμβάνονται ενσωματωμένα στο φακό, ο οποίος δύναται να περιέχει ηλεκτρονικά κυκλώματα, μία κεραία για ασύρματη επικοινωνία και πολλές διόδους εκπομπής φωτός LEDs, για το σχηματισμό των εικόνων μπροστά από το μάτι του χρήστη, όπως για παράδειγμα λέξεων, διαγραμμάτων, ή φωτογραφιών [1], [65], [66]. Εκτός από τους βιονικούς αυτούς φακούς επαφής με τεχνολογία LED, αναπτύσσονται και φακοί βασισμένοι στην τεχνολογία υγρών κρυστάλλων LCD [67].

1.4.2 Φορητές συσκευές χειρός

Οι συσκευές αυτές, ευρέως γνωστές ως Handheld Displays (HDs), περιλαμβάνουν μία μικρή οθόνη, μέσω της οποίας είναι δυνατή η θέαση της επαυξημένης σκηνής, και είναι φορητές, μικρού μεγέθους, έτσι ώστε ο χρήστης να μπορεί να τις κρατήσει στο χέρι του και να τις μεταφέρει μαζί του. Τέτοιες συσκευές είναι τα κινητά τηλέφωνα – και συγκεκριμένα τα «έξυπνα» κινητά τηλέφωνα, γνωστά ως smartphones – τα tablet PCs και οι – λιγότερο χρησιμοποιούμενοι πλέον – προσωπικοί ψηφιακοί οδηγοί PDAs (Personal Digital Assistants). Στη μεγάλη πλειονότητα των εφαρμογών επαυξημένης πραγματικότητας που προορίζονται για φορητές συσκευές χειρός, χρησιμοποιείται η βίντεο τεχνολογία, όπου η ενσωματωμένη στις συσκευές αυτές βιντεοκάμερα «αιχμαλωτίζει» τη ζωντανή εικόνα του πραγματικού περιβάλλοντος, η οποία επαυξάνεται με επιπρόσθετες γραφικές πληροφορίες πριν την απεικόνισή της στην οθόνη. Ωστόσο, έχουν υλοποιηθεί και εφαρμογές επαυξημένης πραγματικότητας για φορητές συσκευές χειρός που χρησιμοποιούν την οπτική τεχνολογία [56].

Μέσω των συσκευών αυτών, σε αντίθεση με τις φορετές συσκευές στο κεφάλι – ή σε τμήμα του κεφαλιού – του χρήστη, που παρουσιάστηκαν στην υποενότητα 1.4.1, ο χρήστης δεν «βυθίζεται» στο επαυξημένο περιβάλλον, αλλά το παρακολουθεί μέσω της οθόνης της συσκευής του, γεγονός το οποίο θα μπορούσε να θεωρηθεί ως μειονέκτημα των HDs [7]. Επίσης, οι χαμηλές δυνατότητες των συσκευών αυτών – κυρίως των κινητών τηλεφώνων και των PDAs – σε επεξεργαστική ισχύ και μνήμη, συγκρινόμενες με τις αντίστοιχες ενός ηλεκτρονικού υπολογιστή, μπορεί να έχουν ως αποτέλεσμα καθυστέρηση κατά την εκτέλεση της εφαρμογής και συνακόλουθο χαμηλό ρυθμό ανανέωσης των αλλεπάλληλων στιγμιοτύπων του βίντεο (frame rate). Τα tablet PCs έχουν μεγαλύτερη επεξεργαστική ισχύ, αλλά παρουσιάζουν το μειονέκτημα του μεγαλύτερου μεγέθους και βάρους [56], [68].

Πέρα από τα μειονεκτήματα των συσκευών αυτών για εφαρμογές επαυξημένης πραγματικότητας, τα κυριότερα εκ των οποίων αναφέρθηκαν παραπάνω, αυτές παρουσιάζουν το πλεονέκτημα ότι μεταφέρονται εύκολα, ενώ μεγάλο ποσοστό του πληθυσμού της γης έχει στην κατοχή του μία τέτοια συσκευή, γεγονός το οποίο συντελεί στην εμπορική επιτυχία εφαρμογών επαυξημένης πραγματικότητας, που μπορούν να απευθυνθούν μέσω των συσκευών αυτών στο μέσο πολίτη. Ακόμη, οι περισσότερες από τις εν λόγω συσκευές διαθέτουν πρόσβαση στο GPS, καθώς επίσης και επιταχυνσιόμετρα, γυροσκόπια και πυξίδα, με αποτέλεσμα οι εφαρμογές επαυξημένης πραγματικότητας που αναπτύσσονται για αυτές να έχουν τη δυνατότητα να υιοθετήσουν μία προσέγγιση επαύξησης βάσει θέσης και όχι να περιορίζονται στη σύνθεση του εικονικού με το πραγματικό με τη βοήθεια μόνο αλγορίθμων της φωτογραμμετρίας και της όρασης υπολογιστών.

1.4.3 Χωρικές συσκευές απεικόνισης

Σε αντίθεση με τις συσκευές απεικόνισης που προσαρτώνται στο κεφάλι του χρήστη (HWDs) ή κρατούνται από αυτόν (HDs), οι χωρικές συσκευές απεικόνισης επαυξημένων σκηνών (spatial displays) αποσπούν το μεγαλύτερο μέρος της τεχνολογίας από το χρήστη και το ενσωματώνουν στο περιβάλλον. Υπάρχουν τρεις διαφορετικές προσεγγίσεις ανάλογα με τον τρόπο που επαυξάνουν το περιβάλλον: χρησιμοποιώντας βίντεο τεχνολογία (video see-through), οπτική τεχνολογία (optical see-through) ή άμεση επαύξηση [6].

1.4.3.1 Συσκευές απεικόνισης σε οθόνη με χρήση βίντεο τεχνολογίας

Οι συσκευές απεικόνισης που χρησιμοποιούν βίντεο τεχνολογία (βλ. Εικόνα 1-28), κάνοντας χρήση της επεξεργασίας εικόνας και της μίξης του εικονικού με το πραγματικό, απεικονίζουν τις επαυξημένες σκηνές σε μία κοινή οθόνη (screen-based video see-through displays). Το είδος αυτό της επαυξημένης πραγματικότητας αναφέρεται συχνά και ως «παράθυρο στον κόσμο» (window on the world) και αποτελεί μία από τις πιο αποδοτικές από πλευρά κόστους λύσεις, καθώς απαιτεί έναν τυπικό ηλεκτρονικό υπολογιστή και έτοιμο hardware. Ωστόσο, παρουσιάζει κάποια μειονεκτήματα, όπως τη μη πλήρη «εμβύθιση» του χρήστη στο περιβάλλον επαυξημένης πραγματικότητας, λόγω του σχετικά μικρού μεγέθους της οθόνης και του συνακόλουθου μικρού οπτικού πεδίου που παρέχεται σε αυτόν. Ακόμη, δευτερεύον μειονέκτημα αποτελεί η συνήθως περιορισμένη ανάλυση της εικόνας του πραγματικού περιβάλλοντος.

Εικόνα 1-28. Παράδειγμα εφαρμογής επαυξημένης πραγματικότητας μέσω συσκευής απεικόνισης σε οθόνη με χρήση βίντεο τεχνολογίας. Στην εικόνα απεικονίζεται το αποτύπωμα δεινοσαύρου το οποίο επαυξάνεται με το πόδι του.

Πηγή: Bimber & Raskar, 2005

1.4.3.2 Χωρικές συσκευές απεικόνισης με χρήση οπτικής τεχνολογίας

Οι χωρικές συσκευές απεικόνισης οπτικής τεχνολογίας (spatial optical see-through displays) (βλ. Εικόνα 1-29) παράγουν εικόνες που τοποθετούνται σε σωστή θέση εντός του φυσικού

περιβάλλοντος και χρησιμοποιούν χωρικούς «οπτικούς συνδυαστές», όπως επίπεδα ή καμπυλωμένα κάτοπτρα - διαχωριστές δέσμης, διάφανες οθόνες ή οπτικά ολογραφήματα[2]. Πλεονεκτήματα αυτής της τεχνολογίας σε σχέση με την προηγούμενη είναι η συνήθως καλύτερη ανάλυση και το μεγαλύτερο οπτικό πεδίο που παρέχουν στο χρήστη.

Εικόνα 1-29. Παράδειγμα εφαρμογής επαυξημένης πραγματικότητας με χρήση οπτικής τεχνολογίας - διάφανη οθόνη.

Πηγή: Bimber & Raskar, 2005

1.4.3.3 Χωρικές συσκευές απεικόνισης μέσω προβολής

Οι χωρικές συσκευές απεικόνισης μέσω προβολής (projection-based spatial displays) προβάλλουν άμεσα τις εικόνες στις φυσικές επιφάνειες των αντικειμένων, μέσω εμπρόσθιας προβολής. Δεν είναι αναγκαία μία στερεοσκοπική προβολή αν λόγω της υπέρθεσης των εικόνων αλλάζουν μόνο οι ιδιότητες των πραγματικών αντικειμένων (για παράδειγμα χρώμα, φωτισμός ή υφή). Ωστόσο, για την παρουσίαση τρισδιάστατων γραφικών μπροστά ή πίσω από τις επιφάνειες των αντικειμένων απαιτείται στερεοσκοπική προβολή. Ένα παράδειγμα χωρικής επαυξημένης πραγματικότητας με χρήση συσκευής απεικόνισης μέσω προβολής παρουσιάζεται στην Εικόνα 1-30.

Εικόνα 1-30. Παράδειγμα χωρικής επαυξημένης πραγματικότητας με χρήση συσκευής απεικόνισης μέσω προβολής. Αριστερά: το πραγματικό περιβάλλον. Κέντρο και δεξιά: το επαυξημένο περιβάλλον.

Πηγή: Bimber & Raskar, 2005

[2] Πρόκειται για τρισδιάστατες απεικονίσεις που δημιουργούνται με τη βοήθεια ενός λέιζερ και μπορούν να παρατηρηθούν από διάφορες γωνίες.

1.5 Μέθοδοι Επαύξησης της Πραγματικότητας

Απαραίτητο στοιχείο για την ολοκλήρωση μίας εφαρμογής επαυξημένης πραγματικότητας είναι η γνώση της θέσης του πραγματικού αντικειμένου, στο οποίο θα τοποθετηθεί το επιπρόσθετο εικονικό αντικείμενο. Ανάλογα με τη μέθοδο που χρησιμοποιείται για την εύρεση της θέσης του αληθινού αντικειμένου που θα επαυξηθεί, μπορούν να καταγραφούν τέσσερα κύρια είδη επαυξημένης πραγματικότητας [4]. Αυτά είναι η επαυξημένη πραγματικότητα βάσει προτύπου, βάσει περιγράμματος, βάσει θέσης και βάσει επιφανείας και περιγράφονται στη συνέχεια.

1.5.1 Επαυξημένη πραγματικότητα βάσει προτύπου

Πρόκειται για μέθοδο επαύξησης της πραγματικότητας που στηρίζεται στην αναγνώριση προτύπου, το οποίο μπορεί να είναι είτε ένας στόχος που τοποθετείται στην πραγματική σκηνή (marker-based augmented reality), είτε μία άλλη εικόνα της πραγματικής σκηνής, που έχει οριστεί στο σύστημα ως πρότυπο (markerless augmented reality), όπως φαίνεται στην Εικόνα 1-31. Τα markers είναι στόχοι, συνήθως επίπεδοι, τετράγωνοι και ασπρόμαυροι, οι οποίοι εισάγονται στο φυσικό περιβάλλον και αναγνωρίζονται από το σύστημα για την τοποθέτηση των εικονικών αντικειμένων πάνω σε αυτούς. Αντίθετα, η δεύτερη τεχνική (markerless augmented reality), βασίζεται στην αναγνώριση αντικειμένων που υπάρχουν στον πραγματικό κόσμο. Το σύστημα, στην περίπτωση αυτή, συγκρίνει την εικόνα του πραγματικού κόσμου, που έχει οριστεί ως πρότυπο, με το βίντεο του πραγματικού κόσμου, για την αναγνώριση του προτύπου στο βίντεο και την επαύξηση της σκηνής.

Εικόνα 1-31. Παράδειγμα επαυξημένης πραγματικότητας βασισμένης σε markers (marker-based AR) [αριστερά] και σε φυσικά χαρακτηριστικά του αντικειμένου (markerless AR) [δεξιά].

Πηγή: http://www.arlab.com/blog/markerless-augmented-reality

Ενώ στην περίπτωση της επαυξημένης πραγματικότητας βάσει στόχου, ο αλγόριθμος ανίχνευσης του marker είναι πιο γρήγορος και εύρωστος σε αλλαγές φωτισμού, δεν αποδίδει αποτελέσματα αν ο στόχος έχει μερικώς καλυφθεί από άλλο αντικείμενο. Επίσης, είναι δεσμευτικό και ίσως αντιαισθητικό το γεγονός ότι πρέπει να τυπωθεί ένα marker και να

τοποθετηθεί στον πραγματικό κόσμο [69]. Και στις δύο περιπτώσεις (marker-based και markerless augmented reality), η αναγνώριση του προτύπου είναι καθοριστικής σημασίας για τη μετέπειτα επαύξηση της σκηνής με ένα τρισδιάστατο μοντέλο ή με άλλη γραφική πληροφορία, ή ακόμα και με βίντεο ή ήχο.

1.5.2 Επαυξημένη πραγματικότητα βάσει περιγράμματος

Σύμφωνα με τη μέθοδο αυτή, αναγνωρίζεται ένα μέρος του σώματος, όπως για παράδειγμα τα χέρια ή το πρόσωπο, ή και ολόκληρο το σώμα, και στη συνέχεια επαυξάνεται με κάποιο εικονικό αντικείμενο. Επιπλέον, δίνεται η δυνατότητα στο χρήστη μίας τέτοιας εφαρμογής να αλληλεπιδράσει με το εικονικό αντικείμενο με φυσικές κινήσεις του, όπως για παράδειγμα να το κουνήσει ή να το κρατήσει στο χέρι του [4]. Η μέθοδος αυτή ακολουθείται από πολλές διαφημιστικές καμπάνιες (εικονική δοκιμή ενός ρολογιού ή ενός ζευγαριού γυαλιών ενώ ο χρήστης κάθεται μπροστά από την οθόνη του υπολογιστή του, κ.λπ.).

1.5.3 Επαυξημένη πραγματικότητα βάσει θέσης

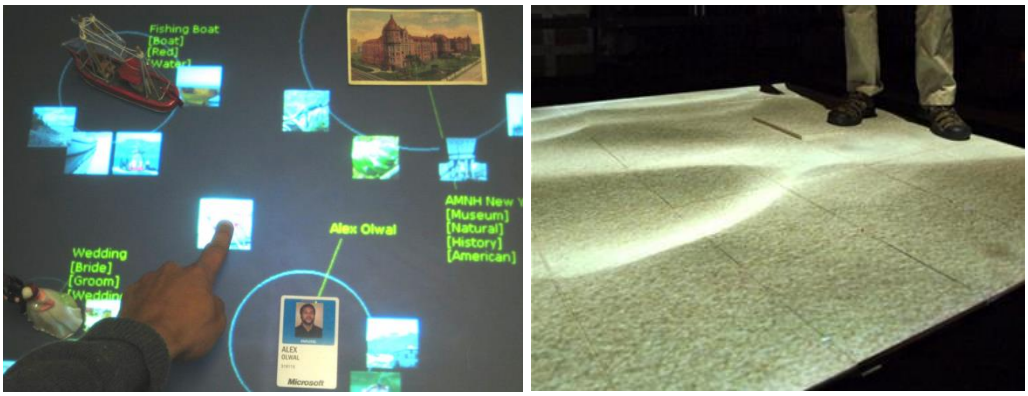
Η μέθοδος αυτή βασίζεται σε πληροφορίες GPS ή τριγωνισμού (για παράδειγμα μέσω του συνδυασμού σημάτων WiFi και των κεραιών κινητής τηλεφωνίας για την εύρεση της θέσης της κινητής συσκευής, η οποία διαθέτει σύνδεση στο Internet μέσω δικτύου κινητής τηλεφωνίας) για την υπέρθεση των συνθετικών αντικειμένων στην εικόνα του πραγματικού κόσμου. Λόγω των πληροφοριών αυτών, σε συνδυασμό με πληροφορίες από άλλους αισθητήρες, όπως επιταχυνσιόμετρο και γυροσκόπιο, είναι δυνατή η γνώση του εξωτερικού προσανατολισμού κάθε στιγμιότυπου (βλ. ενότητα 3.6), η οποία επιτρέπει την προβολή της εικονικής πληροφορίας στη σωστή θέση της συσκευής θέασης. Συνήθως, η μέθοδος αυτή υιοθετείται από κινητά τηλέφωνα για εφαρμογές επαυξημένης πραγματικότητας σε εξωτερικούς χώρους (ακόμα και σε εσωτερικούς με χρήση των δεδομένων τριγωνισμού) και ιδιαίτερα για τα προγράμματα περιήγησης επαυξημένης πραγματικότητας.

1.5.4 Επαυξημένη πραγματικότητα σε επιφάνεια

Η επαυξημένη πραγματικότητα σε επιφάνεια (surface augmented reality) επιτυγχάνεται με τη βοήθεια οθονών, πατωμάτων ή τοίχων που αποκρίνονται στο άγγιγμα από ανθρώπους ή αντικείμενα και παρέχουν εικονικές πληροφορίες σε πραγματικό χρόνο, ενώ δίνουν τη δυνατότητα στους χρήστες να αλληλεπιδρούν με αυτές [4]. Ένα τέτοιο παράδειγμα είναι το SurfaceFusion (βλ. Εικόνα 1-32, αριστερά), που κατασκευάστηκε από τη Microsoft και δίνει τη δυνατότητα στο χρήστη να αλληλεπιδράσει με φυσικά αντικείμενα και με τις ψηφιακές πληροφορίες που τα συνοδεύουν [70]. Συγγενές ερευνητικό πρόγραμμα της Microsoft είναι το LightSpace, που συνδυάζει τις «υπολογιστικές επιφάνειες» με στοιχεία επαυξημένης

πραγματικότητας για να δημιουργήσει ένα περιβάλλον, κάθε επιφάνεια του οποίου είναι πλήρως διαδραστική, όπως είναι και ο χώρος μεταξύ των επιφανειών [71].

Άλλο παράδειγμα της επαυξημένης πραγματικότητας επιφανείας είναι το πάτωμα επαυξημένης πραγματικότητας (βλ. Εικόνα 1-32, δεξιά), που χρησιμοποιεί ειδικά πλακίδια, τα οποία με χρήση ειδικών βαθμονομημένων δονήσεων μπορούν να προσομοιάσουν την άμμο, το χιόνι, το γρασίδι και ποικιλία άλλων επιφανειών [4].



Εικόνα 1-32. Παραδείγματα επαυξημένης πραγματικότητας σε επιφάνεια. Αριστερά: SurfaceFusion. Δεξιά: πάτωμα επαυξημένης πραγματικότητας που μιμείται την εμφάνιση και την αίσθηση της άμμου.

Πηγές: Olwal & Wilson, 2008,
http://www.technologyreview.com/view/418718/augmented-reality-floor-tiling

1.6 Προκλήσεις και Προβλήματα Επαύξησης της Πραγματικότητας

Μία από τις μεγαλύτερες προκλήσεις με τις οποίες έρχονται αντιμέτωποι οι δημιουργοί εφαρμογών επαυξημένης πραγματικότητας είναι η σωστή τοποθέτηση του εικονικού αντικειμένου εντός του πραγματικού περιβάλλοντος, έτσι ώστε η συνθετική πληροφορία να δίνει την εντύπωση ότι ανήκει σε αυτό. Η διαδικασία αυτή είναι γνωστή ως registration (γεωαναφορά). Η σωστή συγχώνευση και ευθυγράμμιση των δύο κόσμων – του φυσικού και του παραγόμενου από υπολογιστή – είναι κύρια προϋπόθεση για την εκπλήρωση του στόχου των εφαρμογών επαυξημένης πραγματικότητας, ενώ λάθη ή ανακρίβειες στην τοποθέτηση του εικονικού αντικειμένου θα έχουν ως αποτέλεσμα να χαθεί η ψευδαίσθηση ότι οι δύο κόσμοι συνυπάρχουν [3]. Πολλές εφαρμογές, μάλιστα, όπως για παράδειγμα στην ιατρική, απαιτούν ακριβή γεωαναφορά και δεν είναι επιτρεπτά λάθη και αστοχίες.

Για τη σωστή γεωαναφορά, απαραίτητη προϋπόθεση είναι η πρότερη ανίχνευση της θέσης και του προσανατολισμού της κάμερας – και γενικά της συσκευής μέσω της οποίας επαυξάνεται η πραγματικότητα – ή του κεφαλιού του χρήστη (π.χ. σε εφαρμογή με HMD). Η διαδικασία αυτή είναι γνωστή ως tracking (ανίχνευση), απαντά στα ερωτήματα: πού βρίσκεται ο χρήστης, πού εστιάζεται το ενδιαφέρον του και πού πρέπει να παρουσιαστεί το εικονικό αντικείμενο [72] και συνεπώς η σωστή και ακριβής, ανάλογα με την εφαρμογή,

διεκπεραίωσή της είναι κρίσιμη για τη δημιουργία πειστικών εφαρμογών επαυξημένης πραγματικότητας. Για την επίτευξη των τελευταίων, η ανίχνευση πρέπει πρακτικά να διεξάγεται σε πραγματικό χρόνο, δηλαδή η εκτίμηση της θέσης να γίνεται σε χιλιοστά του δευτερολέπτου, καθώς επίσης και να είναι εύρωστη, δηλαδή να δίνει ικανοποιητικά αποτελέσματα κάτω από ποικίλες συνθήκες, όπως για παράδειγμα σε μεταβαλλόμενο φωτισμό [68].

Υπάρχουν πολλές μέθοδοι ανίχνευσης 6 βαθμών ελευθερίας (6DOF tracking), όπως το μηχανικό tracking, τεχνική που υιοθετήθηκε και από το πρώτο σύστημα επαυξημένης πραγματικότητας του Sutherland, το υπερηχητικό tracking και το οπτικό tracking. Υπάρχουν και άλλες που υπολογίζουν μόνο θέση ή προσανατολισμό, όπως για παράδειγμα το tracking που βασίζεται σε πληροφορίες μόνο από GPS ή μόνο από γυροσκόπια [15]. Για εφαρμογές που απαιτούν ακριβή γεωαναφορά, ωστόσο, απαιτείται ακριβές στιγμιαίο 6DOF tracking υπό οποιεσδήποτε συνθήκες. Επειδή η τέλεια ανίχνευση είναι – τουλάχιστον προς το παρόν – ανέφικτη, εξαιτίας των χρονικών καθυστερήσεων ή και των περιορισμών λόγω ακριβείας, κύρια πρόκληση αποτελεί η εύρεση της μεθόδου ανίχνευσης που είναι ιδανική για τη συγκεκριμένη κάθε φορά εφαρμογή [73].

Υπάρχει ένας ακόμη αριθμός προκλήσεων που συνδέονται με το πρόβλημα της ανίχνευσης και τοποθέτησης του εικονικού αντικειμένου εντός του φυσικού περιβάλλοντος [4], η κυριότερη από τις οποίες είναι οι αποκρύψεις (occlusion). Σύμφωνα με την τελευταία, πρέπει να επιτυγχάνεται η πλήρης ή μερική απόκρυψη του εικονικού αντικειμένου όταν κάποιο άλλο αντικείμενο του πραγματικού περιβάλλοντος τοποθετείται μπροστά από αυτό και το κρύβει, πλήρως ή μερικώς. Άλλη δυσκολία στις εφαρμογές επαυξημένης πραγματικότητας, σχετική με την οπτική ανίχνευση, είναι η μη εστιασμένη κάμερα στο marker ή στο πρότυπο που πρέπει να αναγνωριστεί για την επαύξηση της πραγματικότητας, γεγονός το οποίο μπορεί να οδηγήσει στη μη αναγνώρισή του, ή σε λάθη στην τοποθέτηση του εικονικού αντικειμένου, λόγω της χαμηλότερης ακρίβειας με την οποία αποδίδεται. Μία ακόμη πρόκληση, συγγενική με την οπτική ανίχνευση (visual tracking), είναι ο μη ομοιόμορφος φωτισμός, λόγω του οποίου ένα marker μπορεί να συσκοτιστεί σε κάποια τμήματά του και να μην αναγνωρίζεται από το πρόγραμμα ή να αναγνωρίζεται ως διαφορετικό marker. Όμοια, λόγω μεταβαλλόμενου φωτισμού, υπάρχει η πιθανότητα μη αναγνώρισης της εικόνας που έχει οριστεί ως πρότυπο. Τέλος, η θαμπάδα που μπορεί να προκληθεί λόγω γρήγορης κίνησης της κάμερας – κυρίως μίας κινητής συσκευής – είναι ένας ακόμα παράγοντας που δύναται να δυσκολέψει τη σωστή επαύξηση της πραγματικής σκηνής.

Ένα κύριο στοιχείο των εφαρμογών επαυξημένης πραγματικότητας είναι η απεικόνιση του εικονικού αντικειμένου στην πραγματική σκηνή, δηλαδή η δημιουργία της συνθετικής επαυξημένης σκηνής, σε πραγματικό χρόνο (real-time rendering). Η φύση κάποιων εφαρμογών απαιτεί οι γραφικές πληροφορίες να ενσωματώνονται στο φυσικό περιβάλλον με

τέτοιο τρόπο ώστε ο παρατηρητής να μην μπορεί να ξεχωρίσει ποιο είναι το πραγματικό και ποιο το εικονικό. Στις εφαρμογές αυτές, εκτός από το σωστό και ακριβές tracking και registration, απαιτείται ταυτόχρονα και φωτορεαλιστικό rendering, με σωστή σκίαση και φωτισμό του εικονικού αντικειμένου και οποιαδήποτε άλλη αυτόματη από το λογισμικό επεξεργασία πραγματικού χρόνου αυτό συνεπάγεται [6]. Καθοριστικό στοιχείο στη βελτίωση της ποιότητας του rendering αποτελεί η ικανότητα της εφαρμογής να λαμβάνει και να αξιοποιεί πληροφορία για το φωτισμό του περιβάλλοντος και την ανάκλαση [3].

Ένα ακόμη βασικό στοιχείο των εφαρμογών επαυξημένης πραγματικότητας είναι, όπως έχει ήδη αναφερθεί, η τεχνολογία θέασης, η οποία ταυτόχρονα αποτελεί και μία πρόκληση, καθώς η βελτίωσή της, ανάλογα με τις απαιτήσεις της εκάστοτε εφαρμογής, μπορεί να συντελέσει σε μεγαλύτερη αποδοχή της τεχνολογίας αυτής από το κοινό. Πράγματι, είναι πολύ πιο βολικό και «γνώριμο» στον άνθρωπο να φορέσει γυαλιά ή φακούς που θα επαυξήσουν την πραγματικότητά του σε σχέση με κάποια βαριά και μεγάλη συσκευή που προσαρτάται στο κεφάλι του (HMD ή HMPD). Εκτός από τέτοιου είδους περιορισμούς, που οφείλονται στον ανθρώπινο παράγοντα και για τους οποίους γίνονται σήμερα πολλές προσπάθειες βελτίωσης, υφίστανται και άλλες προκλήσεις σχετικές με την τεχνολογία θέασης της επαυξημένης πραγματικότητας [6], όπως είναι για παράδειγμα οι οπτικοί περιορισμοί λόγω του περιορισμένου οπτικού πεδίου του χρήστη, καθώς και οι τεχνικοί περιορισμοί, όπως η περιορισμένη ανάλυση και διάφοροι άλλοι παράγοντες.

Εκτός των παραπάνω σημαντικών προκλήσεων με τις οποίες έρχονται αντιμέτωπες οι εφαρμογές επαυξημένης πραγματικότητας, υπάρχουν και άλλα τεχνικά ζητήματα σχετικά με αυτές, όπως είναι η αλληλεπίδραση του χρήστη με την εικονική πληροφορία, γεγονός το οποίο θα του δώσει την αίσθηση της πλήρους ενσωμάτωσης στο συνθετικό αυτό κόσμο, που συνδυάζει το πραγματικό με το εικονικό. Τέλος, οι δημιουργοί των εφαρμογών επαυξημένης πραγματικότητας πρέπει να δίνουν σημασία και στο πλήθος των εικονικών πληροφοριών που υπερτίθενται στο πραγματικό περιβάλλον του χρήστη, έτσι ώστε αυτό να μην εμποδίζει το χρήστη από τη θέαση του φυσικού κόσμου, αλλά ούτε και να είναι ανεπαρκές.

ΚΕΦΑΛΑΙΟ 2

ΨΗΦΙΑΚΗ ΣΥΝΤΑΥΤΙΣΗ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ ΣΗΜΕΙΩΝ

Ενώ στο πρώτο κεφάλαιο έγινε μία εισαγωγή στην έννοια της επαυξημένης πραγματικότητας, που αποτελεί αντικείμενο των εφαρμογών που προγραμματίστηκαν για τις ανάγκες της παρούσας διπλωματικής εργασίας, στο κεφάλαιο αυτό αλλά και στο επόμενο που ακολουθεί παρουσιάζεται το θεωρητικό υπόβαθρο των αλγορίθμων και των σχέσεων που χρησιμοποιήθηκαν στις εφαρμογές.

Με τον όρο συνταύτιση εικόνας εννοείται η εφαρμογή κατάλληλων αλγορίθμων για την αυτόματη αποκατάσταση ομολογιών μεταξύ σημείων, τόνων του γκρι, χαρακτηριστικών, σχέσεων ή και άλλων οντοτήτων σε μία ή, κυρίως, σε επικαλυπτόμενες ψηφιακές εικόνες [74]. Αποτελεί σημαντική διαδικασία για πληθώρα εφαρμογών φωτογραμμετρίας, όπως για την αυτόματη εύρεση σημείων σύνδεσης σε πολλές εικόνες στο πλαίσιο του αεροτριγωνισμού, την αυτόματη εύρεση ομόλογων σημείων σε προσανατολισμένο στερεοζεύγος για το μετέπειτα προσδιορισμό των τρισδιάστατων συντεταγμένων τους και τη συλλογή του Ψηφιακού Μοντέλου Εδάφους, την αυτόματη μέτρηση εικονοσημάτων κ.ά. Η ψηφιακή συνταύτιση αποτελεί σημαντικό βήμα για την ανάπτυξη των εφαρμογών επαυξημένης πραγματικότητας.

Το κεφάλαιο αυτό αποτελεί μία εισαγωγή στη συνταύτιση χαρακτηριστικών σημείων, που εφαρμόστηκε στο πλαίσιο των εφαρμογών. Για λόγους πληρότητας, ωστόσο, στην αρχή του κεφαλαίου γίνεται μία σύντομη αναφορά και στα τρία κύρια είδη της ψηφιακής συνταύτισης, τα οποία είναι, εκτός από τη συνταύτιση χαρακτηριστικών, η επιφανειακή και η σχεσιακή συνταύτιση. Στη συνέχεια, περιγράφεται η ανίχνευση και περιγραφή των σημείων ενδιαφέροντος, διαδικασία απαραίτητη για τη διενέργεια της συνταύτισης χαρακτηριστικών. Στο πλαίσιο αυτής, παρουσιάζεται αναλυτικά ο αλγόριθμος SURF (Speeded-Up Robust Features), ο οποίος χρησιμοποιήθηκε στις εφαρμογές. Τέλος, περιγράφεται η κυρίως διαδικασία της συνταύτισης, δηλαδή η εύρεση των ομολογιών μεταξύ των χαρακτηριστικών σημείων, ενώ κρίνεται σκόπιμη και η παράθεση κάποιων μεθόδων απαλοιφής των μη

έγκυρων αντιστοιχίσεων, έτσι ώστε το τελικό αποτέλεσμα της συνταύτισης να θεωρείται αποδεκτό.

2.1 Κατηγορίες Ψηφιακής Συνταύτισης

Υπάρχουν τρεις γενικές κατηγορίες ψηφιακής συνταύτισης. Αυτές διακρίνονται ανάλογα με την οντότητα η οποία ταυτίζεται και είναι η επιφανειακή συνταύτιση (area-based matching), η συνταύτιση χαρακτηριστικών (feature-based matching) και η σχεσιακή συνταύτιση (relational matching) [74]. Προτού παρουσιαστεί αναλυτικά η συνταύτιση χαρακτηριστικών, που πραγματοποιήθηκε στο πλαίσιο των εφαρμογών, κρίνεται σκόπιμο να γίνει μία σύντομη αναφορά στα τρία αυτά είδη, στην ενότητα αυτή.

Στο πλαίσιο της επιφανειακής συνταύτισης, τα ομόλογα εικονοσημεία καθορίζονται με βάση την ομοιότητα των ψηφιακών τιμών τους και έτσι οι οντότητες που συγκρίνονται είναι οι τιμές χρώματος ή γκρίζου τόνου. Η λογική των αλγορίθμων επιφανειακής συνταύτισης έγκειται στον ορισμό γύρω από το εικονοστοιχείο, το οποίο έχει εντοπιστεί ως σημείο ενδιαφέροντος, ενός «παραθύρου αναφοράς» ψηφιακών τιμών, που συγκρίνεται με όλα τα αντίστοιχα «παράθυρα σύγκρισης» σε ένα ευρύτερο «παράθυρο αναζήτησης» της εικόνας στην οποία επιδιώκεται η εύρεση του ομολόγου του. Τα κριτήρια ομοιότητας με βάση τα οποία μπορούν να καθοριστούν οι ομολογίες είναι ο συντελεστής συσχέτισης και το τυπικό σφάλμα σ_0, ή ισοδύναμα η σύγκλιση της ελαχιστοτετραγωνικής επίλυσης (ελαχιστοτετραγωνική συνταύτιση, LSM). Όσον αφορά στον αλγόριθμο επιφανειακής συνταύτισης με βάση το πρώτο μέτρο ομοιότητας, υπολογίζεται ο συντελεστής συσχέτισης μεταξύ των παραθύρων σύγκρισης και αναφοράς. Το εικονοστοιχείο στο κέντρο του παραθύρου σύγκρισης, για το οποίο παρατηρείται ο μέγιστος συντελεστής συσχέτισης με το παράθυρο αναφοράς, είναι το ζητούμενο ομόλογο εικονοστοιχείο εκείνου που βρίσκεται στο κέντρο του παραθύρου αναφοράς. Βασική αρχή της ελαχιστοτετραγωνικής συνταύτισης είναι η ελαχιστοποίηση των διαφορών των ψηφιακών τιμών μεταξύ του παραθύρου σύγκρισης και του παραθύρου αναφοράς, καθώς το παράθυρο σύγκρισης έχει τη δυνατότητα να μεταβάλλεται κατά θέση, προσανατολισμό και σχήμα [74].

Σύμφωνα λοιπόν με την επιφανειακή συνταύτιση, εντοπίζονται σε μία εικόνα σημεία ενδιαφέροντος και μέσω μίας τεχνικής τοπικής έρευνας, όπως είναι η συσχέτιση ή τα ελάχιστα τετράγωνα, ανιχνεύονται τα ίδια σε άλλες επικαλυπτόμενες εικόνες. Στο πλαίσιο της συνταύτισης χαρακτηριστικών, εντοπίζονται σημεία ενδιαφέροντος ανεξάρτητα σε κάθε μία εικόνα και στη συνέχεια αυτά ταυτίζονται, με βάση τοπικές πληροφορίες από τη γειτονιά τους. Ενώ η πρώτη προσέγγιση είναι περισσότερο κατάλληλη για εφαρμογή σε εικόνες που λήφθηκαν από κοντινά σημεία λήψης, ή που αντιστοιχούν σε μικρή χρονική διαφορά μεταξύ τους, η δεύτερη ενδείκνυται σε περιπτώσεις εικόνων μεταξύ της λήψης των οποίων μεσολάβησε μεγάλο χρονικό διάστημα ή κίνηση, με αποτέλεσμα να μη λήφθηκαν από

κοντινά σημεία. Συγκεκριμένα, στο πλαίσιο της συνταύτισης χαρακτηριστικών, οι οντότητες οι οποίες ταυτίζονται μπορεί να είναι όχι μόνο σημεία, αλλά και ακμές και γενικότερα γραμμές, ή περιοχές, ενώ μέτρο ομοιότητας αποτελεί μία συνάρτηση κόστους [74]. Διεξάγεται σε δύο βήματα, εκ των οποίων το πρώτο αφορά στην ανίχνευση των χαρακτηριστικών και το δεύτερο στην αντιστοίχισή τους.

Τα χαρακτηριστικά σημεία περιγράφονται από μεγέθη που χαρακτηρίζουν τα γειτονικά τους εικονοστοιχεία και ταυτίζονται με βάση αυτά. Κατά παράδοση, αναφέρονται και ως «γωνίες» (corners), μολονότι μπορεί να μην πρόκειται για τέτοιου είδους σημεία. Ο χαρακτηρισμός τους ως «γωνίες» οφείλεται στο γεγονός ότι αρχικά αναπτύχθηκαν αλγόριθμοι ανίχνευσης ακμών και εν συνεχεία αναλύονταν οι ακμές για την εύρεση των γωνιών. Ωστόσο, τα σημεία αυτά μπορεί να μην είναι γωνίες, με την αυστηρή έννοια. Συχνά, ως σημεία ενδιαφέροντος αναφέρονται και κάποιες δομές - περιοχές ενδιαφέροντος, γνωστές ως «blobs», για τις οποίες θα γίνει λόγος στο κεφάλαιο αυτό. Οι αλγόριθμοι ανίχνευσης και ταύτισης των blobs συγκαταλέγονται συχνά στην κατηγορία των αλγορίθμων ανίχνευσης και ταύτισης σημείων ενδιαφέροντος, καθώς οι δομές αυτές συνήθως περιέχουν ένα προτιμώμενο χαρακτηριστικό σημείο [75], [76]. Επίσης, συχνή σε φωτογραμμετρικές εφαρμογές είναι η ανίχνευση ακμών, οι οποίες αποτελούν όρια μεταξύ περιοχών διαφορετικού χρώματος, έντασης ή υφής και είναι γραμμές εκατέρωθεν των οποίων παρατηρείται μία σημαντική αλλαγή της έντασης ή του χρώματος της εικόνας. Καθώς στις περιοχές αυτές εμφανίζει μέγιστο η πρώτη προσανατολισμένη παράγωγος της συνάρτησης έντασης της εικόνας, δηλαδή γίνεται μέγιστο το διάνυσμα κλίσης, η ανίχνευσή τους μπορεί να γίνει με βάση αυτή και με ένα κατώφλι, άνω του οποίου το σημείο μπορεί να θεωρηθεί ως σημείο ακμής. Επίσης, η ανίχνευση ακμών μπορεί να πραγματοποιηθεί και με τη δεύτερη παράγωγο, μέσω του υπολογισμού της Λαπλασιανής, καθώς η δεύτερη παράγωγος, αναπαριστώντας τη μεταβολή της κλίσης, μηδενίζεται επί της ακμής, παρουσιάζοντας τοπικό μέγιστο και ελάχιστο εκατέρωθεν αυτής. Μία άλλη προσέγγιση, γνωστή ως «Laplacian of Gaussian» (LoG), είναι ο συνδυασμός δύο τελεστών, της Γκαουσιανής, για μείωση του «θορύβου» και της Λαπλασιανής, για την εύρεση των ακμών, σε ένα βήμα [74], [77], [78]. Υπάρχουν βέβαια και άλλες τεχνικές ανίνχευσης ακμών, ενώ σημαντικός αλγόριθμος είναι ο ανιχνευτής ακμών Canny, που αναπτύχθηκε το 1986 από τον John Canny και χρησιμοποιείται ευρέως μέχρι σήμερα [79]. Οι ακμές μπορούν να ταυτιστούν με βάση τον προσανατολισμό και την τοπική τους εμφάνιση [75]. Ανήκουν στη γενικότερη κατηγορία των γραμμών, οι οποίες μπορεί να αποτελέσουν, όπως ήδη αναφέρθηκε, οντότητα προς ταύτιση στο πλαίσιο της συνταύτισης χαρακτηριστικών και ανιχνεύονται με διάφορες τεχνικές, όπως για παράδειγμα με το μετασχηματισμό Hough [80].

Τρίτη κατηγορία αποτελεί η σχεσιακή συνταύτιση, η οποία συνταυτίζει σχεσιακές (συμβολικές) περιγραφές πολύπλοκων ή μη χαρακτηριστικών δομών. Μέτρο ομοιότητας και στην περίπτωση αυτή αποτελεί μία συνάρτηση κόστους. Οι δομές αυτές περιγράφονται από

τις παραμέτρους των επιμέρους στοιχείων τους και από τις σχέσεις, γεωμετρικές, ραδιομετρικές και τοπολογικές, μεταξύ των επιμέρους χαρακτηριστικών [74].

2.2 Ανίχνευση και Περιγραφή Σημείων Ενδιαφέροντος

Η συνταύτιση χαρακτηριστικών, όπως ήδη αναφέρθηκε, στηρίζεται στην αρχική ανίχνευση των σημείων ενδιαφέροντος. Προκειμένου να χαρακτηριστεί ένα σημείο ως σημείο ενδιαφέροντος για τις ανάγκες τις συνταύτισης, πρέπει να διαθέτει κάποιες ιδιότητες. Αυτές, σύμφωνα με το Wolfgang Förstner, είναι οι εξής [81]:

- Η διακριτότητα (distinctness), σύμφωνα με την οποία τα σημεία ενδιαφέροντος πρέπει να διαφέρουν από τα σημεία της γειτονιάς τους.
- Η αμεταβλητότητα (invariance), σύμφωνα με την οποία τα χαρακτηριστικά σημεία οφείλουν να είναι ανεξάρτητα από γεωμετρικές και ραδιομετρικές παραμορφώσεις.
- Η σταθερότητα (stability), κατά την οποία τα χαρακτηριστικά σημεία πρέπει να υπάρχουν και στις άλλες εικόνες που δείχνουν την ίδια περιοχή. Η ιδιότητα αυτή υποδεικνύει, εκτός των άλλων, ότι τα χαρακτηριστικά που ανιχνεύονται δεν πρέπει να αποτελούν «θόρυβο».
- Η σπανιότητα (seldomness), η οποία υπονοεί τη μοναδικότητα των σημείων ενδιαφέροντος στην εικόνα και εξασφαλίζεται με το μη εντοπισμό σημείων ενός επαναλαμβανόμενου προτύπου ως χαρακτηριστικών σημείων τα οποία εν συνεχεία θα ταυτιστούν. Διαφέρει από τη διακριτότητα, υπό την έννοια ότι η τελευταία αναφέρεται στη δυνατότητα διαχωρισμού των σημείων στη γειτονιά τους και όχι σε ολόκληρη την εικόνα.
- Η ερμηνευσιμότητα (interpretability), σύμφωνα με την οποία τα σημεία πρέπει να επιδέχονται φυσική ερμηνεία. Η απαίτηση για το χαρακτηριστικό αυτό, βέβαια, δεν υφίσταται πάντα, αλλά συνήθως σε περιπτώσεις που τα σημεία αυτά χρησιμοποιούνται για ανάλυση εικόνας.

Την προηγούμενη δεκαετία έγιναν αξιόλογες προσπάθειες για την ανάπτυξη αλγορίθμων ανίχνευσης και περιγραφής τοπικών χαρακτηριστικών, τα οποία μπορούν να εντοπιστούν και εν συνεχεία να ταυτιστούν, ανεξάρτητα από διάφορους μετασχηματισμούς της εικόνας, όπως η μετατόπιση, η στροφή, η αλλαγή κλίμακας και αφινικές παραμορφώσεις. Η ταύτιση των χαρακτηριστικών αυτών, που είναι γνωστά ως «local invariant features», ανεξάρτητα από τους παραπάνω μετασχηματισμούς, επιτυγχάνεται μέσω της περιγραφής τους [82]. Σύμφωνα με τα παραπάνω, σημαντικό χαρακτηριστικό των αλγορίθμων ανίχνευσης και περιγραφής των σημείων ενδιαφέροντος, πέραν της ακρίβειάς τους, αποτελεί η επαναληψιμότητά τους (repeatability), σύμφωνα με την οποία πρέπει να εντοπίζονται τα ίδια

χαρακτηριστικά σε εικόνες που δείχνουν το ίδιο αντικείμενο, ανεξάρτητα από την κλίμακά τους, τον προσανατολισμό τους κ.λπ.

2.2.1 Γενικές πληροφορίες και διάφοροι αλγόριθμοι

Μία κοινή διαδικασία ανίχνευσης και περιγραφής των σημείων ενδιαφέροντος θα μπορούσε να συνοψισθεί όπως περιγράφεται παρακάτω [82]. Αρχικά εντοπίζονται στην εικόνα σημεία ενδιαφέροντος, τα οποία διαφέρουν από τα γειτονικά τους. Για κάθε σημείο ενδιαφέροντος ορίζεται μία περιοχή, με τρόπο ανεξάρτητο κλίμακας ή ανεξάρτητο αφινικών μετασχηματισμών (affine invariant). Ακολούθως γίνεται «κανονικοποίηση» της περιοχής αυτής, ώστε να είναι ανεξάρτητη από στροφή της εικόνας, και υπολογίζεται ο «περιγραφέας» της, γνωστός ως «descriptor», με τη μορφή κάποιου πίνακα περιγραφικών χαρακτηριστικών. Αφού ολοκληρωθεί η παραπάνω διαδικασία ανίχνευσης και περιγραφής των σημείων ενδιαφέροντος, είναι δυνατή η διεξαγωγή της συνταύτισης, μέσω της ταύτισης των περιγραφέων των σημείων ενδιαφέροντος.

Για την ανίχνευση των σημείων ενδιαφέροντος μπορούν να χρησιμοποιηθούν διάφοροι τελεστές, όπως οι ανιχνευτές Hessian και Harris. Ο πρώτος βασίζεται στον Εσσιανό Πίνακα – ο οποίος θα περιγραφεί στη συνέχεια – και εντοπίζει μέρη της εικόνας που παρουσιάζουν μεγάλες τιμές παραγώγων σε δύο ορθογώνιες διευθύνσεις, δηλαδή παρουσιάζουν μεγάλη αλλαγή στην κλίση και στις δύο διευθύνσεις. Ο ανιχνευτής αυτός εντοπίζει γωνίες, αλλά και περιοχές έντονης υφής. Ο τελεστής Harris βασίζεται στον υπολογισμό της μέγιστης και ελάχιστης τιμής της συνάρτησης μεταβολής της έντασης σε κάθε εικονοστοιχείο, για τον εντοπισμό των σημείων ενδιαφέροντος, τα οποία στην περίπτωση αυτή αποτελούν γωνίες. Το μειονέκτημά των τελεστών αυτών είναι η μη εύρεση των ίδιων σημείων ενδιαφέροντος σε εικόνες που απεικονίζουν την ίδια περιοχή αλλά παρουσιάζουν μεγάλες διαφορές κλίμακας μεταξύ τους. Ωστόσο, μπορεί να είναι εύρωστοι σε μεταβολές φωτισμού, σε «θόρυβο» και σε μικρές αλλαγές της κλίμακας [82], [83], [84].

Για την ανίχνευση χαρακτηριστικών αμετάβλητων λόγω κλίμακας μπορούν να χρησιμοποιηθούν οι τελεστές Laplacian of Gaussian (LoG) και Difference of Gaussian (DoG), οι οποίοι εντοπίζουν δομές blobs και εισάγουν την έννοια του «χώρου κλίμακας». Εκτός αυτών, μπορούν να συνδυαστούν οι τελεστές Hessian και Harris με την ιδέα επιλογής κλίμακας μέσω της Λαπλασιανής. Οι παραλλαγές των τελεστών αυτών είναι γνωστές ως Hessian-Laplacian και Harris-Laplacian αντίστοιχα [82], [84].

Για πολλά πρακτικά προβλήματα είναι σημαντική η εύρεση χαρακτηριστικών, ανθεκτικών σε αφινικούς μετασχηματισμούς της εικόνας, με αποτέλεσμα να μπορούν να εντοπίζονται ακόμα και σε μεγάλες αλλαγές της οπτικής γωνίας θέασης της σκηνής. Στην περίπτωση εξαγωγής χαρακτηριστικών ανθεκτικών σε αφινικούς μετασχηματισμούς (γνωστής ως «affine covariant extraction»), το σχήμα των χαρακτηριστικών αυτών ποικίλει, ανάλογα με

την υποκείμενη αφινική παραμόρφωση, έτσι ώστε το περιεχόμενο του χαρακτηριστικού που εντοπίζεται να παραμένει αμετάβλητο. Ένας τέτοιος αλγόριθμος είναι ο ανιχνευτής MSER, ενώ μπορούν να χρησιμοποιηθούν και παραλλαγές των ανιχνευτών Hessian και Harris, γνωστές ως Hessian-Affine και Harris-Affine αντίστοιχα [82].

Μετά από την ανίχνευση της περιοχής γίνεται ένα είδος «κανονικοποίησης» του περιεχομένου της, ώστε να είναι ανεξάρτητη από τυχόν στροφή της εικόνας. Τυπικά, αυτό επιτυγχάνεται με την εύρεση του κυρίαρχου προσανατολισμού της περιοχής και την περιστροφή του περιεχομένου της σύμφωνα με τη γωνία που βρέθηκε [82].

Επόμενο βήμα αποτελεί η περιγραφή των σημείων ενδιαφέροντος. Πολύ γνωστός περιγραφέας χαρακτηριστικών είναι εκείνος που υλοποιείται από τον αλγόριθμο SIFT (Scale Invariant Feature Transform), ενώ ο SURF, ο οποίος παρουσιάζεται αναλυτικά στην υποενότητα 2.2.2, είναι εξίσου γνωστός και αποτελεσματικός αλγόριθμος ανίχνευσης και περιγραφής σημείων ενδιαφέροντος[3].

2.2.2 Αλγόριθμος SURF

Ο αλγόριθμος SURF (Speeded-Up Robust Features) παρουσιάστηκε το 2006 από τους Herbert Bay, Tinne Tuytelaars και Luc Van Gool [85]. Όπως φανερώνει και το όνομά του, είναι ένας αλγόριθμος «ταχύτερα υπολογιζόμενων εύρωστων χαρακτηριστικών», ο οποίος, σύμφωνα με τους συγγραφείς του, προσεγγίζει αλγορίθμους που προτάθηκαν πριν από αυτόν και μάλιστα παρουσιάζει καλύτερες επιδόσεις από αυτούς. Στηρίζεται στα δυνατά σημεία παλαιότερων αλγορίθμων, σε απλοποίηση διαφόρων μεθόδων τους και σε μία ειδική μορφή αναπαράστασης της εικόνας, τη λεγόμενη «αθροιστική εικόνα» (integral image), με αποτέλεσμα τη μείωση του χρόνου ανίχνευσης και περιγραφής των σημείων ενδιαφέροντος και τη μετέπειτα μείωση του χρόνου εύρεσης των ομόλογων σημείων, για την οποία θα γίνει λόγος στην επόμενη ενότητα. Ειδικότερα, έχει αρκετές ομοιότητες με τον αλγόριθμο SIFT, ενώ οι συγγραφείς του έδειξαν πειραματικά ότι έχει καλύτερες επιδόσεις από αυτόν τόσο από άποψη υπολογιστικού χρόνου, όσο και από άποψη ακρίβειας.

Ο SURF στηρίζεται σε έναν ανιχνευτή και σε έναν περιγραφέα. Ο πρώτος εντοπίζει σημεία ενδιαφέροντος, ενώ ο δεύτερος περιγράφει τη γειτονιά τους, χωρίς να χρησιμοποιούν πληροφορίες χρώματος. Τόσο ο ανιχνευτής όσο και ο περιγραφέας είναι ανεξάρτητοι από περιστροφή και κλίμακα. Αυτό σημαίνει ότι τα σημεία ενδιαφέροντος τα οποία εντοπίζονται είναι αναλλοίωτα σε μετασχηματισμούς περιστροφής και κλίμακας, με συνέπεια να εντοπίζονται τα ίδια σημεία σε εικόνες που απεικονίζουν κάποιο κοινό τμήμα και έχουν

[3] Οι αλγόριθμοι και τελεστές που αναφέρθηκαν δεν είναι οι μοναδικοί και υπάρχει πληθώρα άλλων για τον ίδιο σκοπό. Ωστόσο, σκοπός της παρούσας διπλωματικής εργασίας δεν είναι η παρουσίαση των διαφόρων αλγορίθμων ανίχνευσης και περιγραφής σημείων ενδιαφέροντος και έτσι αναφέρθηκαν ενδεικτικά κάποιοι από αυτούς.

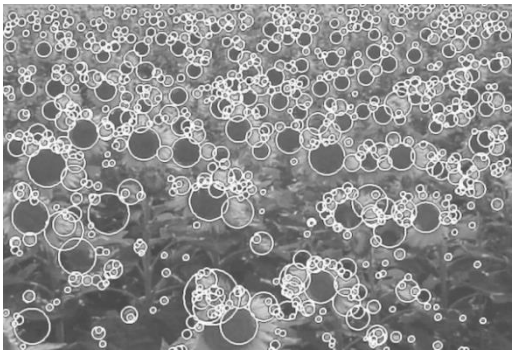
ληφθεί με διαφορετικές γωνίες στροφής ή / και έχουν διαφορετική κλίμακα και στη συνέχεια να μπορούν να ταυτιστούν με βάση τις πληροφορίες του περιγραφέα, κατά τη διαδικασία της εύρεσης των ομολογιών. Επίσης, ο περιγραφέας είναι ανεξάρτητος από μεταβολές φωτισμού και αντίθεσης, ενώ λόγω της ευρωστίας του, καλύπτονται ως ένα βαθμό η μη ορθογωνικότητα των πλευρών των εικονοστοιχείων, η ανισοτροπική κλίμακα και διάφορες προοπτικές επιδράσεις [86], [87].

Στόχος της ανάπτυξης του αλγορίθμου αυτού ήταν η δημιουργία ενός ανιχνευτή και ενός περιγραφέα οι οποίοι να είναι γρήγοροι και ταυτόχρονα να έχουν καλή επίδοση. Η διατήρηση της ισορροπίας ανάμεσα στις παραπάνω απαιτήσεις πραγματοποιείται μέσω της απλοποίησης της διαδικασίας ανίχνευσης και της ταυτόχρονης διατήρησης της ακρίβειάς της καθώς επίσης και μέσω της μείωσης του μεγέθους του διανύσματος περιγραφικών χαρακτηριστικών της γειτονιάς του σημείου ενδιαφέροντος και της παράλληλης διατήρησης της ικανότητας περιγραφής κάθε σημείου με ξεχωριστό τρόπο [86]. Στο πλαίσιο της υποενότητας αυτής γίνεται μία παρουσίαση του αλγορίθμου SURF για χρήση του στην ανίχνευση και περιγραφή των σημείων ενδιαφέροντος.

2.2.2.1 Δομές blobs

Τα blobs ορίστηκαν από τους Chad Carson et al. [88] και αποτελούν σύνθετες δομές που μπορούν να περιγραφούν με βάση πληροφορίες χρώματος, υφής, χωρικού μέσου και πίνακα διασποράς. Αντιστοιχούν σε περιοχές της εικόνας που έχουν διαφορετικές ιδιότητες, όπως φωτεινότητα ή χρώμα, σε σχέση με τις γύρω τους περιοχές, ενώ τα σημεία εντός μίας περιοχής blob μπορούν να θεωρηθούν ότι έχουν παρόμοιες ιδιότητες το ένα με το άλλο [89].

Ο αλγόριθμος SURF εντοπίζει τέτοιου είδους δομές, οι οποίες συναντώνται σε γωνίες των αντικειμένων και σε κηλίδες [87] (βλ. Εικόνα 2-1). Μολονότι πρόκειται για περιοχές της εικόνας, ο SURF θεωρείται ως αλγόριθμος ανίχνευσης και περιγραφής σημείων ενδιαφέροντος, καθώς θεωρεί ως χαρακτηριστικό το κεντρικό σημείο των blobs, το οποίο περιγράφει και εν συνεχεία ταυτίζει με τα υπόλοιπα.



Εικόνα 2-1. Δομές blobs.

Πηγή: Bay et al., 2008

2.2.2.2 Αθροιστικές εικόνες

Πρώτο βήμα του αλγορίθμου είναι η μετατροπή της εικόνας σε μία ειδική μορφή, γνωστή ως integral image ή αθροιστική εικόνα. Πρόκειται για μία αναπαράσταση της εικόνας που επινοήθηκε το 2001 από τους Paul Viola και Michael Jones [90], η οποία επιτρέπει τη γρήγορη διεξαγωγή της συνέλιξης με φίλτρα τύπου «κουτιού» (box-type filters ή box filters)[4] και την επακόλουθη γρήγορη ανίχνευση σημείων ενδιαφέροντος σε διάφορες κλίμακες [85]. Η αθροιστική εικόνα στη θέση (x,y) περιέχει το άθροισμα των τιμών των εικονοστοιχείων εντός μίας ορθογώνιας περιοχής που δημιουργείται από την αρχή της εικόνας (το επάνω αριστερά pixel) και τη θέση (x,y). Δηλαδή, η τιμή του κάθε pixel της αθροιστικής εικόνας είναι το άθροισμα των επάνω και αριστερά εικονοστοιχείων της αρχικής εικόνας, συμπεριλαμβανομένου και του εικονοστοιχείου στη συγκεκριμένη θέση [85], [90], όπως γίνεται κατανοητό και από την επόμενη σχέση:

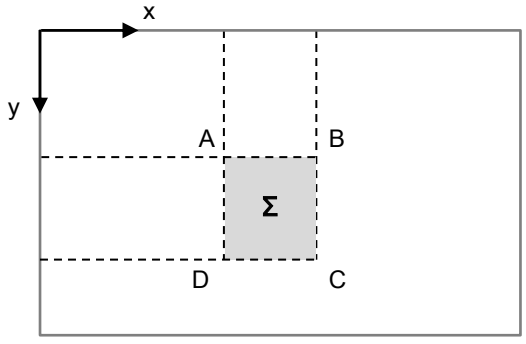
$$I_{\Sigma}(x,y) = \sum_{i=0}^{i\le x}\sum_{j=0}^{j\le y} I(x,y) \tag{2.1}$$

όπου I η αρχική εικόνα και I_{Σ} η αθροιστική εικόνα. Η τελευταία μπορεί να υπολογιστεί σε ένα πέρασμα, με βάση τις ακόλουθες σχέσεις:

$$S(x,y) = S(x,y-1) + I(x,y) \tag{2.2}$$

$$I_{\Sigma}(x,y) = I_{\Sigma}(x-1,y) + S(x,y) \tag{2.3}$$

όπου S(x,y) είναι το μερικό άθροισμα κατά στήλη, ενώ θεωρείται ότι S(x,-1)=0 και I_{Σ}(-1,y)=0, το οποίο σύμφωνα με τις σχέσεις (2.2) και (2.3) συνεπάγεται ότι: S(x,0)=I(x,0) και I_{Σ}(0,y)=S(0,y).



Εικόνα 2-2. Το άθροισμα των τιμών των εικονοστοιχείων του ορθογωνίου εντός της εικόνας μπορεί να υπολογιστεί μέσω της απλής πράξης: Σ=A+C-B-D, όπου A, B, C και D είναι οι τιμές της αθροιστικής εικόνας στις αντίστοιχες θέσεις.

Πηγή: Viola & Jones, 2001

Αφού υπολογιστεί η αθροιστική εικόνα, το άθροισμα των τιμών των εικονοστοιχείων εντός μίας ορθογώνιας περιοχής της εικόνας, μπορεί να υπολογιστεί με πρόσθεση και αφαίρεση

[4] Πρόκειται για φίλτρα που συντίθενται από ορθογώνιες περιοχές, όπως θα γίνει κατανοητό στη συνέχεια.

μεταξύ τεσσάρων μόνο στοιχείων, των τιμών της αθροιστικής εικόνας (A, B, C, D) στις γωνίες του ορθογωνίου. Το άθροισμα αυτό είναι ίσο με: A+C-B-D (βλ. Εικόνα 2-2) [90]. Το γεγονός αυτό οδηγεί σε ανεξαρτησία του χρόνου υπολογισμού του αθροίσματος από το μέγεθος του φίλτρου, κάτι το οποίο είναι σημαντικό κατά την εκτέλεση του αλγορίθμου, κυρίως όταν χρησιμοποιούνται φίλτρα μεγαλύτερων μεγεθών [86]

2.2.2.3 Εσσιανός Πίνακας και παραλλαγή του που χρησιμοποιείται από τον αλγόριθμο

Ο Εσσιανός Πίνακας (Hessian Matrix) είναι ένας τετραγωνικός πίνακας, τα στοιχεία του οποίου είναι οι μερικές παράγωγοι δεύτερης τάξης μίας συνάρτησης. Αναπτύχθηκε από το Γερμανό μαθηματικό Ludwig Otto Hesse το 19ο αιώνα και περιγράφει την τοπική καμπυλότητα μίας συνάρτησης πολλών μεταβλητών [91].

Ο Εσσιανός Πίνακας στη θέση (x,y) της εικόνας και για συντελεστή κλίμακας σ, ορίζεται ως εξής [86]:

$$H(x,y,\sigma)=\begin{bmatrix} L_{xx}(x,y,\sigma) & L_{xy}(x,y,\sigma) \\ L_{xy}(x,y,\sigma) & L_{yy}(x,y,\sigma) \end{bmatrix} \tag{2.4}$$

όπου L_{xx}, L_{yy} και L_{xy} τα αποτελέσματα των συνελίξεων της δεύτερης τάξης παραγώγου της Γκαουσιανής συνάρτησης δύο διαστάσεων κατά τις διευθύνσεις x, y και xy αντίστοιχα με την εικόνα, στο σημείο (x,y). Ο συντελεστής κλίμακας σ ταυτίζεται με την τυπική απόκλιση της Γκαουσιανής κατανομής. Μαθηματικά, τα στοιχεία του Εσσιανού Πίνακα συμβολίζονται ως εξής:

$$L_{xx}(x,y,\sigma)=\frac{\partial^2}{\partial x^2}G(x,y,\sigma)*I(x,y) \tag{2.5}$$

$$L_{yy}(x,y,\sigma)=\frac{\partial^2}{\partial y^2}G(x,y,\sigma)*I(x,y) \tag{2.6}$$

$$L_{xy}(x,y,\sigma)=\frac{\partial^2}{\partial x\partial y}G(x,y,\sigma)*I(x,y) \tag{2.7}$$

όπου το σύμβολο * δηλώνει την πράξη της συνέλιξης[5]. Η Γκαουσιανή συνάρτηση δύο διαστάσεων δίνεται από την ακόλουθη σχέση [92]:

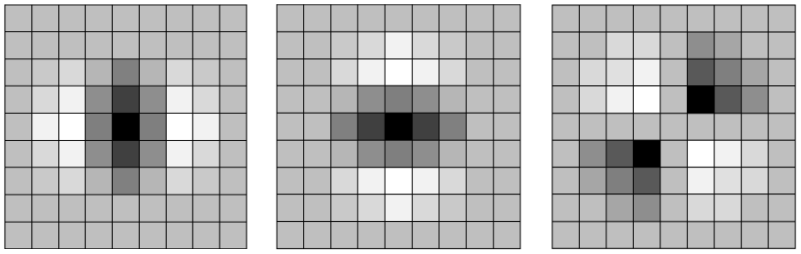
$$G(x,y,\sigma)=\frac{1}{2\pi\sigma^2}e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{2.8}$$

Οι δεύτερης τάξης παράγωγοι της συνάρτησης Gauss πρέπει να διακριτοποιηθούν και να κοπούν [86], με αποτέλεσμα την παραγωγή φίλτρων, ανάλογων με αυτών που φαίνονται

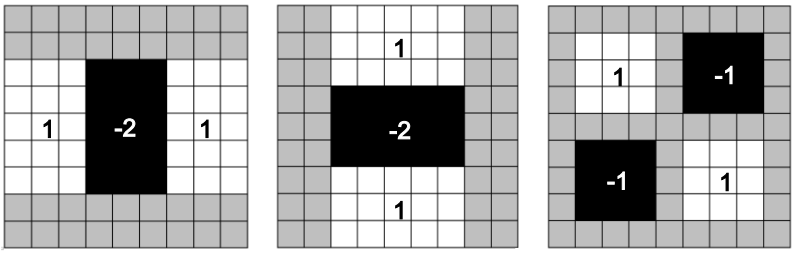
[5] Σύμφωνα με αυτήν, πολλαπλασιάζεται κάθε στοιχείο ενός φίλτρου (μάσκα) με το αντίστοιχο στοιχείο του πίνακα ίδιων διαστάσεων με το φίλτρο, ο οποίος σχηματίζεται με κέντρο το pixel στη θέση (x,y) της εικόνας, και τα γινόμενα αυτά αθροίζονται. Το άθροισμα των γινομένων είναι το αποτέλεσμα της συνέλιξης.

στην Εικόνα 2-3. Λαμβάνοντας υπ' όψιν, για παράδειγμα, τη σχέση (2.5), η τιμή $L_{xx}(x,y,\sigma)$ του Εσσιανού Πίνακα είναι το άθροισμα των γινομένων κάθε στοιχείου του φίλτρου $\partial^2 G(x,y,\sigma)/\partial x^2$, που προκύπτει από τη διακριτοποίηση και κοπή των παραγώγων δεύτερης τάξης της Γκαουσιανής συνάρτησης, με το αντίστοιχο στοιχείο του πίνακα της εικόνας που έχει ως κέντρο το pixel στη θέση (x,y) και διαστάσεις ίδιες με εκείνες του φίλτρου.

Ο αλγόριθμος SURF για την ανίχνευση των σημείων ενδιαφέροντος χρησιμοποιεί μία παραλλαγή του Εσσιανού Πίνακα, τα στοιχεία της οποίας είναι τα αποτελέσματα της συνέλιξης του παραθύρου της εικόνας με κέντρο το pixel στη θέση (x,y) με φίλτρα τύπου «κουτιού». Στην Εικόνα 2-4 φαίνονται τα 9x9 box filters που προσεγγίζουν τις παραγώγους δεύτερης τάξης της συνάρτησης Gauss για σ=1.2 και αντιστοιχούν στο φιλτράρισμα μικρότερης κλίμακας, δηλαδή μεγαλύτερης χωρικής ανάλυσης, για τον υπολογισμό της απόκρισης της εικόνας σε blobs.



Εικόνα 2-3. Φίλτρα, διαστάσεων 9x9, που προκύπτουν από τη διακριτοποίηση και κοπή των παραγώγων δεύτερης τάξης της συνάρτησης Gauss δύο διαστάσεων στις διευθύνσεις x [αριστερά], y [κέντρο] και xy [δεξιά].



Εικόνα 2-4. Φίλτρα, διαστάσεων 9x9, που προσεγγίζουν τις παραγώγους δεύτερης τάξης της συνάρτησης Gauss δύο διαστάσεων για σ=1.2, στις διευθύνσεις x [αριστερά], y [κέντρο] και xy [δεξιά]. Οι γκρι περιοχές έχουν μηδενική τιμή.

Ο προσεγγιστικός Εσσιανός Πίνακας που χρησιμοποιείται από τον αλγόριθμο είναι ο εξής:

$$H_{approx.}(x,y,\sigma) = \begin{bmatrix} D_{xx}(x,y,\sigma) & D_{xy}(x,y,\sigma) \\ D_{xy}(x,y,\sigma) & D_{yy}(x,y,\sigma) \end{bmatrix} \quad (2.9)$$

όπου D_{xx}, D_{xy} και D_{yy} τα αποτελέσματα των συνελίξεων των box filters που προσεγγίζουν τις παραγώγους δεύτερης τάξης της Γκαουσιανής συνάρτησης κατά τις διευθύνσεις x, y και xy αντίστοιχα με την εικόνα, στο σημείο (x,y).

Η ορίζουσα του προσεγγιστικού Εσσιανού Πίνακα υπολογίζεται σύμφωνα με την επόμενη σχέση:

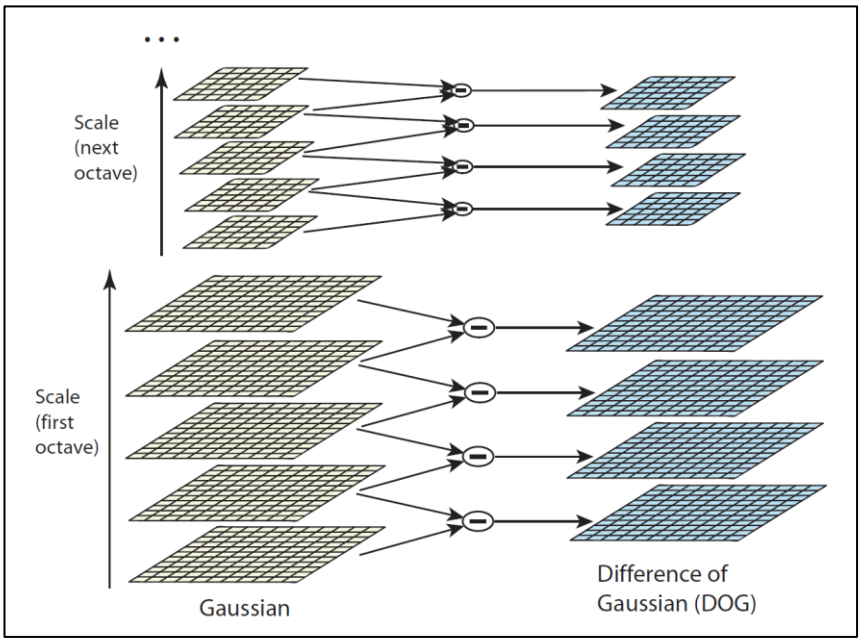
$$\det\left(H_{approx}\right) = D_{xx}\left(x,y,\sigma\right)\cdot D_{yy}\left(x,y,\sigma\right) - \left[w\cdot D_{xy}\left(x,y,\sigma\right)\right]^2 \quad (2.10)$$

Ο όρος w εισάγεται στον υπολογισμό της ορίζουσας για να αντισταθμίσει το σφάλμα που προκύπτει από την προσέγγιση των παραγώγων δεύτερης τάξης της συνάρτησης Gauss με τα φίλτρα τύπου «κουτιού». Μολονότι εξαρτάται από την κλίμακα σ, στην πράξη μπορεί να λάβει προσεγγιστικά την τιμή 0.9. Η ορίζουσα αυτή κανονικοποιείται ανάλογα με το μέγεθος του φίλτρου, δηλαδή ανάλογα με τον παράγοντα κλίμακας. Η διακρίνουσα αυτή αναπαριστά την απόκριση της εικόνας σε blobs (blob response), στη θέση (x,y) και στην κλίμακα σ που αντιστοιχεί στις διαστάσεις των χρησιμοποιούμενων φίλτρων. Οι αποκρίσεις σε blobs αποθηκεύονται σε ένα χάρτη απόκρισης σε blobs (blob response map) σε διαφορετικές κλίμακες και ανιχνεύονται τοπικά μέγιστα [86], [87].

2.2.2.4 Αναπαράσταση του χώρου κλίμακας

Ο αλγόριθμος SURF εντοπίζει σημεία ενδιαφέροντος σε διάφορες κλίμακες, έτσι ώστε να μπορούν να βρεθούν κοινά σημεία σε εικόνες που απεικονίζουν την ίδια περιοχή αλλά με διαφορετική κλίμακα και έτσι η συνταύτιση να μπορεί να διεξαχθεί ομαλά με βάση τις ομολογίες των σημείων. Η ανίχνευση σημείων ενδιαφέροντος αναλλοίωτων σε μετασχηματισμούς κλίμακας γίνεται όχι μόνο από το SURF, αλλά και από άλλους αλγορίθμους, με το συνδυασμό πληροφοριών από πολλαπλές συνελίξεις με την ίδια οικογένεια φίλτρων, σε διαφορετικές κλίμακες [87]. Προκειμένου να επιτευχθεί αυτό, πραγματοποιείται αναπαράσταση του επονομαζόμενου χώρου κλίμακας (scale space), ο οποίος αντιστοιχεί στα αποτελέσματα των συνελίξεων.

Συνήθως, ο χώρος κλίμακας αναπαρίσταται με τη βοήθεια της πυραμίδας της εικόνας. Η εικόνα εξομαλύνεται με φίλτρο Gauss και πραγματοποιείται δειγματοληψία στη νέα αυτή εικόνα, με επιλογή κάθε δεύτερης εικονοψηφίδας κατά την οριζόντια και κατακόρυφη διεύθυνση για τη δημιουργία ενός νέου επιπέδου πυραμίδας. Η διαδικασία αυτή επαναλαμβάνεται. Ειδικότερα, ο David Lowe [93] χρησιμοποιεί την παραγωγή εικόνων Difference of Gaussian (DoG), στο πλαίσιο του αλγορίθμου SIFT, όπως περιγράφεται στη συνέχεια. Πραγματοποιείται διαδοχικά συνέλιξη της αρχικής εικόνας με φίλτρα εξομάλυνσης Gauss, διαφορετικών τυπικών αποκλίσεων. Σύμφωνα με πειραματικά δεδομένα, επαρκεί η συνέλιξη της εικόνας τέσσερις φορές με φίλτρα Gauss. Οι διαδοχικές αυτές εικόνες αφαιρούνται ανά δύο, για να δημιουργήσουν τις εικόνες DoG. Ακολουθεί η επανασύσταση της εικόνας με επιλογή κάθε δεύτερης εικονοψηφίδας κατά την οριζόντια και την κατακόρυφη διεύθυνση και η επανάληψη της παραπάνω διαδικασίας, δηλαδή η συνέλιξη αυτής της εικόνας με φίλτρα Gauss διαφορετικών τυπικών αποκλίσεων και ο υπολογισμός των νέων εικόνων DoG [83], [86], [93], [94]. Η διαδικασία αυτή επαναλαμβάνεται, όπως φαίνεται στην Εικόνα 2-5.



Εικόνα 2-5. Γκαουσιανή πυραμίδα και πυραμίδα DoG (Difference of Gaussian).

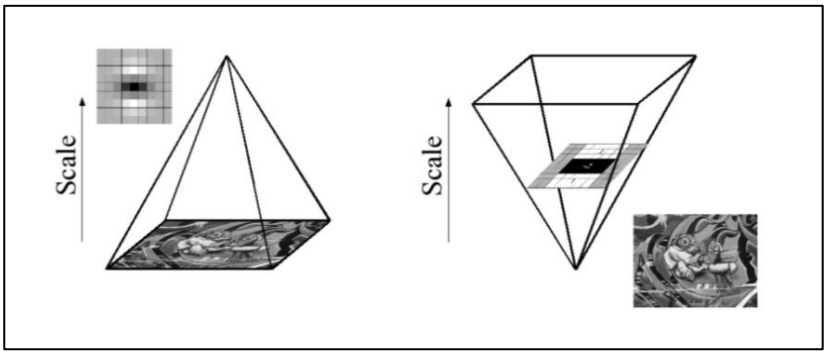
Πηγή: Lowe, 2004

Στην περίπτωση του SURF, η αναπαράσταση του χώρου κλίμακας πραγματοποιείται με την αύξηση του μεγέθους του φίλτρου και δεν απαιτείται μείωση του μεγέθους της εικόνας, όπως γίνεται συνήθως (βλ. Εικόνα 2-6). Συγκεκριμένα, πραγματοποιείται συνέλιξη της αρχικής εικόνας με φίλτρα τύπου «κουτιού» διαφόρων μεγεθών ακριβώς στην ίδια ταχύτητα – λόγω της χρήσης των αθροιστικών εικόνων και του συγκεκριμένου είδους φίλτρων – ακόμα και παράλληλα. Τα διαδοχικά επίπεδα του χώρου κλίμακας αποκτώνται με την αύξηση του μεγέθους των φίλτρων και καθώς αυτό αυξάνεται, μεγαλώνει και η τιμή της αντίστοιχης κλίμακας. Η κλίμακα αυτή μπορεί να υπολογιστεί σύμφωνα με την ακόλουθη σχέση [95]:

$$s = (\text{current filter size}) \cdot \frac{(\text{base filter scale})}{(\text{base filter size})} = (\text{current filter size}) \cdot \frac{1.2}{9} \qquad (2.11)$$

Για παράδειγμα, για το 9x9 φίλτρο, η προκύπτουσα κλίμακα είναι s=9·1.2/9=1.2 και προσεγγίζει τις Γκαουσιανές παραγώγους με σ=1.2, όπως έχει ήδη αναφερθεί. Πλεονέκτημα της αναπαράστασης του χώρου κλίμακας με τον τρόπο αυτό, πέραν της υπολογιστικής αποτελεσματικότητας, είναι η εξάλειψη του φαινομένου του aliasing, που προκύπτει από «υποδειγματοληψία», σύμφωνα με την οποία η ανάλυση δειγματοληψίας δεν είναι επαρκώς υψηλή για να συλλάβει το σύνολο της υπάρχουσας πληροφορίας [74], [86], [96].

Ο χώρος κλίμακας χωρίζεται σε οκτάβες. Μία οκτάβα αναπαριστά μία σειρά από χάρτες απόκρισης σε blobs, που αποκτήθηκαν με συνέλιξη της εικόνας με φίλτρα αυξανόμενου μεγέθους [86]. Κάθε οκτάβα κάνει δειγματοληψία, γραμμικά, σε μία μικρή περιοχή του χώρου κλίμακας. Το μέγεθος της περιοχής αυτής είναι ανάλογο με τα μεγέθη κλίμακας που χρησιμοποιούνται. Κάθε οκτάβα διπλασιάζει το μέγεθος της περιοχής και το βήμα της δειγματοληψίας που χρησιμοποιήθηκε στην αμέσως προηγούμενη οκτάβα [87].

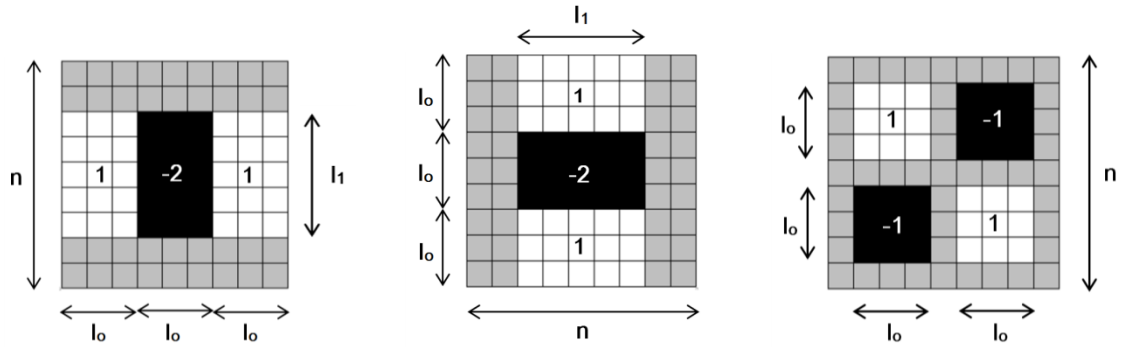


Εικόνα 2-6. Αντί για την επαναληπτική μείωση του μεγέθους της εικόνας [αριστερά], ο αλγόριθμος SURF χρησιμοποιεί την αύξηση του μεγέθους του φίλτρου για την αναπαράσταση του χώρου κλίμακας [δεξιά].

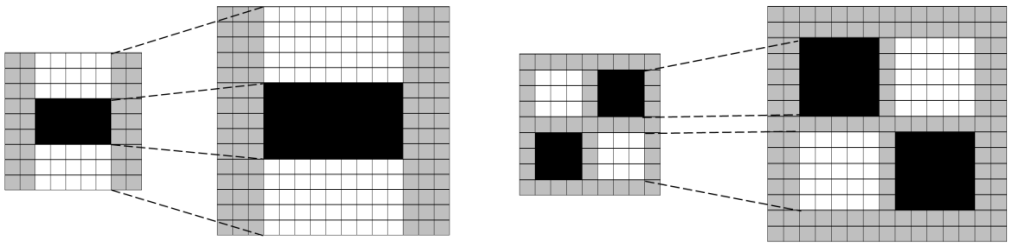
Πηγή: Bay et al., 2008

Κάθε οκτάβα υποδιαιρείται σε ένα σταθερό αριθμό επιπέδων κλίμακας. Η μικρότερη διαφορά κλίμακας μεταξύ δύο διαδοχικών κλιμάκων εξαρτάται από το μέγεθος l_o, το οποίο λαμβάνεται κατά την κατεύθυνση της παραγώγισης και φαίνεται στην Εικόνα 2-7. Το μέγεθος αυτό τίθεται ίσο με το ένα τρίτο του μεγέθους του φίλτρου. Για παράδειγμα, για δύο διαδοχικά επίπεδα της πρώτης οκτάβας, το l_o πρέπει να αυξηθεί κατά τουλάχιστον 2 pixels (ένα pixel στην κάθε πλευρά), προκειμένου το μέγεθος του φίλτρου να παραμείνει περιττός αριθμός και έτσι να εξασφαλιστεί η παρουσία του κεντρικού pixel. Αυτό έχει ως αποτέλεσμα μία συνολική αύξηση του μεγέθους του φίλτρου μεταξύ των διαδοχικών επιπέδων κλίμακας της πρώτης οκτάβας κατά 6 pixels, όπως φαίνεται και στην Εικόνα 2-8. Για μεγέθη διαφορετικά από το l_o, όπως είναι το μέγεθος l_1 που λαμβάνεται κατά την κάθετη στην παραγώγιση κατεύθυνση και φαίνεται στην Εικόνα 2-7, η αλλαγή κλίμακας του φίλτρου εισάγει ένα σφάλμα λόγω στρογγυλοποίησης, το οποίο ωστόσο είναι επιτρεπτό, επειδή είναι αρκετά μικρότερο από το μέγεθος l_o. Εξάλλου, το μέγεθος l_1 μπορεί να αυξηθεί μόνο κατά ένα ζυγό αριθμό, για να διασφαλιστεί η παρουσία του κεντρικού pixel [86].

Η κατασκευή του χώρου κλίμακας ξεκινά με το 9x9 φίλτρο τύπου «κουτιού», το οποίο υπολογίζει την απόκριση της εικόνας σε blobs για τη μικρότερη κλίμακα και αντιστοιχεί στη δεύτερη παράγωγο της Γκαουσιανής συνάρτησης με σ=1.2. Στη συνέχεια εφαρμόζονται τα φίλτρα 15x15, 21x21 και 27x27 στο πλαίσιο της πρώτης οκτάβας. Η δεύτερη οκτάβα χρησιμοποιεί φίλτρα μεγέθους 15x15, 27x27, 39x39 και 51x51, καθώς η αύξηση του μεγέθους του φίλτρου μεταξύ των διαδοχικών επιπέδων κλίμακας της οκτάβας διπλασιάζεται σε σχέση με την πρώτη οκτάβα (από 6 γίνεται 12). Στην τρίτη οκτάβα, το βήμα δειγματοληψίας στην περιοχή του χώρου κλίμακας λαμβάνει την τιμή 24 και τα μεγέθη των φίλτρων που την απαρτίζουν είναι συνεπώς 27x27, 51x51, 75x75 και 99x99. Τέλος, αν η εικόνα είναι αρκετά μεγάλη, χρησιμοποιείται και μία ακόμη οκτάβα με φίλτρα μεγέθους 51x51, 99x99, 147x147 και 195x195 [86], [87].



Εικόνα 2-7. Μεγέθη l_o και l_1 των 9x9 φίλτρων που προσεγγίζουν τις παραγώγους δεύτερης τάξης της συνάρτησης Gauss δύο διαστάσεων για σ=1.2, στις διευθύνσεις x [αριστερά], y [κέντρο] και xy [δεξιά]. Με το γράμμα n συμβολίζεται το μέγεθος του φίλτρου.

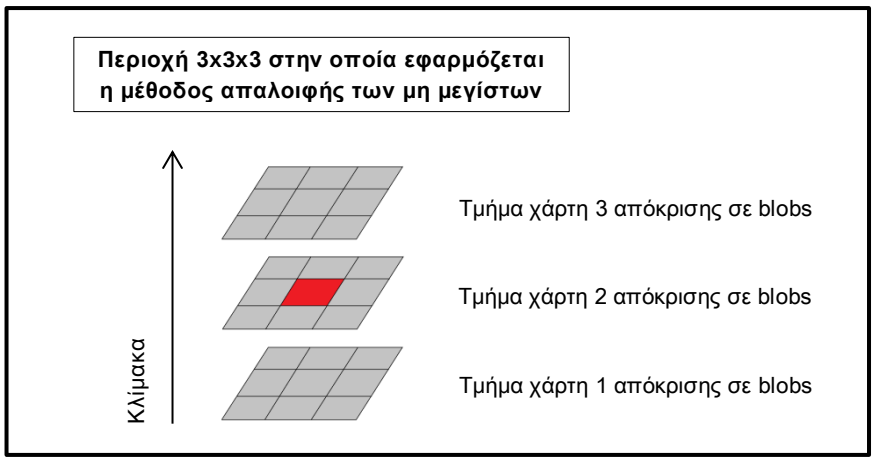


Εικόνα 2-8. Φίλτρα 9x9 και 15x15 που προσεγγίζουν τις παραγώγους δεύτερης τάξης της συνάρτησης Gauss δύο διαστάσεων στις διευθύνσεις y [αριστερά] και xy [δεξιά] για δύο διαδοχικά επίπεδα κλίμακας της πρώτης οκτάβας.

Πηγή: Bay et al., 2008

2.2.2.5 Εντοπισμός των σημείων ενδιαφέροντος

Κριτήριο για τον εντοπισμό των σημείων ενδιαφέροντος αποτελεί η ορίζουσα της παραλλαγής του Εσσιανού Πίνακα, με βάση την οποία επιλέγονται τόσο η θέση, όσο και η κλίμακά τους. Για τον προσδιορισμό των σημείων ενδιαφέροντος στην εικόνα και στο χώρο κλίμακας, δηλαδή για την εύρεση της θέσης τους (x,y) και της κλίμακας στην οποία εντοπίστηκαν, εφαρμόζεται η μέθοδος απαλοιφής των μη μεγίστων (non-maximum suppression), όπως παρουσιάστηκε από τους Alexander Neubeck και Luc Van Gool, σε μία γειτονιά 3x3x3 στο «χώρο εικόνας - κλίμακας» [86] (βλ. Εικόνα 2-9). Σύμφωνα με τη μέθοδο αυτή [97], εντοπίζονται τοπικά μέγιστα, δηλαδή pixels που έχουν τιμή μεγαλύτερη από τα γειτονικά τους. Στην προκειμένη περίπτωση, η τιμή που συγκρίνεται είναι η ορίζουσα της παραλλαγής του Εσσιανού Πίνακα, ενώ τα γειτονικά εικονοστοιχεία του συγκεκριμένου pixel είναι τα 26 pixels εντός της 3x3x3 περιοχής όπου εφαρμόζεται η μέθοδος απαλοιφής των μη μεγίστων (8 εικονοστοιχεία που το περιβάλλουν στο χάρτη απόκρισης σε blobs, 9 εικονοστοιχεία του χάρτη απόκρισης σε blobs του επόμενου επιπέδου κλίμακας και 9 του προηγούμενου επιπέδου κλίμακας). Η διαδικασία αυτή γίνεται για όλες τις οκτάβες και καθίσταται φανερό ότι από κάθε μία οκτάβα χρησιμοποιούνται μόνο οι δύο μεσαίοι χάρτες απόκρισης σε blobs για τον εντοπισμό του σημείου ενδιαφέροντος στο χώρο εικόνας - κλίμακας [98].



Εικόνα 2-9. Περιοχή 3x3x3 στην οποία εφαρμόζεται η μέθοδος απαλοιφής των μη μεγίστων.

Τα αποτελέσματα που βρίσκονται σε μεγαλύτερες κλίμακες μπορεί να έχουν μεγαλύτερο σφάλμα, λόγω του μεγαλύτερου διαστήματος μεταξύ των επιπέδων κλίμακας στις αντίστοιχες οκτάβες. Για να αντισταθμίσει το γεγονός αυτό, ο αλγόριθμος SURF παρεμβάλλει τα μέγιστα της ορίζουσας της παραλλαγής του Εσσιανού Πίνακα στο χώρο εικόνας - κλίμακας, παρέχοντας «sub-pixel» και «sub-scale» ακρίβεια [86], [87]. Η παρεμβολή αυτή γίνεται με τη μέθοδο που προτείνεται από τους Matthew Brown και David Lowe [99], ταιριάζοντας στην ορίζουσα της παραλλαγής του Εσσιανού Πίνακα μία επιφάνεια δευτέρου βαθμού (τρισδιάστατη επιφάνεια που περιγράφεται με τη χρήση εξισώσεων δευτέρου βαθμού). Συγκεκριμένα, η ορίζουσα της Εσσιανής συνάρτησης εκφράζεται ως ανάπτυγμα Taylor με όρους μέχρι και δεύτερης τάξης στη θέση και κλίμακα όπου εντοπίστηκε το σημείο ενδιαφέροντος:

$$\det\left(\mathrm{H}_{\text{approx.}}\left(\mathbf{x}\right)\right)=\det\left(\mathrm{H}_{\text{approx.}}\left(\mathbf{x_0}\right)\right)+\frac{\partial\det\left(\mathrm{H}_{\text{approx.}}\left(\mathbf{x}\right)\right)^{\mathrm{T}}}{\partial\mathbf{x}}\mathbf{x}+\frac{1}{2}\mathbf{x}^{\mathrm{T}}\frac{\partial^2\det\left(\mathrm{H}_{\text{approx.}}\left(\mathbf{x}\right)\right)}{\partial\mathrm{x}^2}\mathbf{x} \qquad (2.12)$$

$$\text{όπου } \mathbf{x}=\left[\mathrm{x}\ \ \mathrm{y}\ \ \sigma\right]^{\mathrm{T}}$$

Η παρεμβαλλόμενη θέση $\hat{\mathbf{x}}=\left[\hat{\mathrm{x}}\ \ \hat{\mathrm{y}}\ \ \hat{\sigma}\right]^{\mathrm{T}}$ του σημείου ενδιαφέροντος με υποψηφιδική ακρίβεια και ακρίβεια «sub-scale» θεωρείται στο ακρότατο της τρισδιάστατης επιφάνειας δευτέρου βαθμού και υπολογίζεται θέτοντας τις παραγώγους της συνάρτησης αυτής ίσες με μηδέν. Τελικά προκύπτει η επόμενη σχέση:

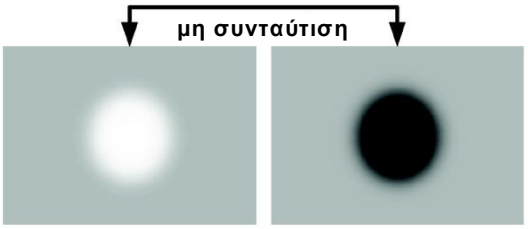
$$\hat{\mathbf{x}}=-\frac{\partial^2\det\left(\mathrm{H}_{\text{approx.}}\left(\mathbf{x}\right)\right)^{-1}}{\partial\mathbf{x}^2}\cdot\frac{\partial\det\left(\mathrm{H}_{\text{approx.}}\left(\mathbf{x}\right)\right)}{\partial\mathbf{x}}=-\begin{bmatrix}\mathrm{d}_{\mathrm{xx}} & \mathrm{d}_{\mathrm{yx}} & \mathrm{d}_{\mathrm{sx}}\\ \mathrm{d}_{\mathrm{xy}} & \mathrm{d}_{\mathrm{yy}} & \mathrm{d}_{\mathrm{sx}}\\ \mathrm{d}_{\mathrm{xs}} & \mathrm{d}_{\mathrm{ys}} & \mathrm{d}_{\mathrm{ss}}\end{bmatrix}^{-1}\cdot\begin{bmatrix}\mathrm{d}_{\mathrm{x}}\\ \mathrm{d}_{\mathrm{y}}\\ \mathrm{d}_{\mathrm{s}}\end{bmatrix} \qquad (2.13)$$

όπου $\mathrm{d}_{\mathrm{x}}=\partial\mathrm{I}/\partial\mathrm{x}$ και $\mathrm{d}_{\mathrm{xx}}=\partial^2\mathrm{I}/\partial\mathrm{x}^2$ η πρώτη και η δεύτερη παράγωγος αντίστοιχα κατά τη x διεύθυνση της εικόνας, κ.ο.κ. Οι παράγωγοι αυτές προσεγγίζονται με αφαιρέσεις των γειτονικών εικονοστοιχείων [86], [87], [95], [96], [99].

Πρέπει να σημειωθεί ότι ανάλογα με την εφαρμογή μπορεί να οριστεί ένα κατώφλι για την τιμή της ορίζουσας της παραλλαγής του Εσσιανού Πίνακα και όλα τα pixels με τιμή της ορίζουσας αυτής κάτω του κατωφλίου να απορρίπτονται, πριν την εφαρμογή της μεθόδου απαλοιφής των μη μεγίστων. Έτσι, ακόμα και αν τα εικονοστοιχεία αυτά είναι τοπικά μέγιστα στη γειτονιά τους, δεν μπορούν να θεωρηθούν ως σημεία ενδιαφέροντος αν έχουν τιμή μικρότερη του κατωφλίου. Το τελευταίο, βέβαια, εξαρτάται κάθε φορά από τη συγκεκριμένη εφαρμογή και προφανώς όσο μεγαλύτερο είναι, τόσο λιγότερα σημεία ενδιαφέροντος θα εντοπιστούν.

Επίσης, οι συγγραφείς του αλγορίθμου προτείνουν να επιστρέφεται το πρόσημο της Λαπλασιανής, δηλαδή του ίχνους του Εσσιανού Πίνακα[6], για κάθε σημείο ενδιαφέροντος, για τη μετέπειτα γρηγορότερη διεξαγωγή της κυρίως διαδικασίας της συνταύτισης. Το πρόσημο αυτό διαχωρίζει ανοικτές δομές blobs σε σκούρο περιβάλλον από την αντίθετη κατάσταση (σκούρες δομές blobs σε ανοικτό περιβάλλον), με αποτέλεσμα τη μετέπειτα σύγκριση μόνο των χαρακτηριστικών σημείων των δομών blobs που έχουν τον ίδιο τύπο αντίθεσης με το περιβάλλον (βλ. Εικόνα 2-10). Η ιδιότητα αυτή διαχωρίζεται με βάση το πρόσημο της Λαπλασιανής, όπως φαίνεται στη συνέχεια:

$$\operatorname{sgn}\left\{L_{xx}\left(x,y,\sigma\right)+L_{yy}\left(x,y,\sigma\right)\right\}=\begin{cases}+1 \;\rightarrow\; \text{ανοικτή blob σε σκούρο περιβάλλον}\\ -1 \;\rightarrow\; \text{σκούρα blob σε ανοικτό περιβάλλον}\end{cases} \qquad (2.14)$$



Εικόνα 2-10. Δομές blobs με διαφορετικό τύπο αντίθεσης. Αν αποθηκεύεται πληροφορία για τον τύπο της αντίθεσης, η μετέπειτα διεξαγωγή της συνταύτισης μπορεί να επιταχυνθεί.

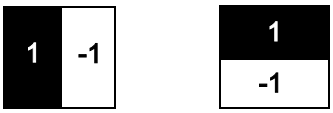
Πηγή: Bay et al., 2008

2.2.2.6 Καθορισμός του κυρίαρχου προσανατολισμού κάθε σημείου ενδιαφέροντος

Την ανίχνευση των σημείων ενδιαφέροντος ακολουθεί η περιγραφή τους, κύριο βήμα της οποίας αποτελεί ο καθορισμός ενός κυρίαρχου προσανατολισμού για κάθε σημείο που ανιχνεύτηκε. Το βήμα αυτό απαιτείται προκειμένου ο αλγόριθμος να εντοπίζει τα ίδια σημεία ενδιαφέροντος σε εικόνες που διαφέρουν κατά μία στροφή. Για το λόγο αυτό, γίνεται συνέλιξη με οριζόντια και κάθετα Haar wavelet φίλτρα σε μία ομάδα εικονοστοιχείων εντός της γειτονιάς του σημείου ενδιαφέροντος.

[6] Πρόκειται για το άθροισμα των στοιχείων της κύριας διαγωνίου του πίνακα.

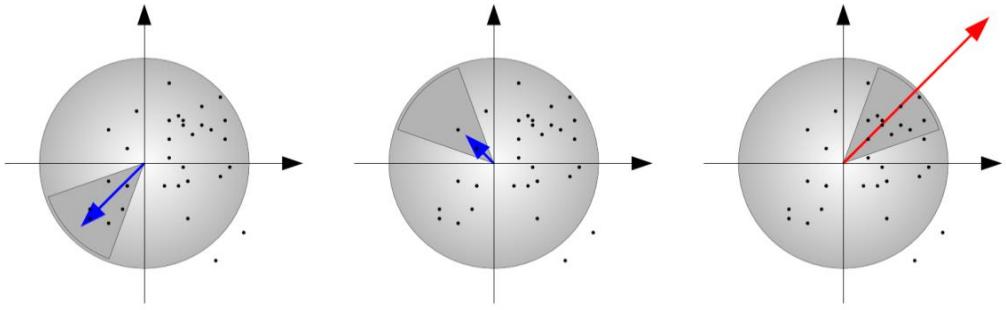
Τα Haar wavelet φίλτρα (βλ. Εικόνα 2-11) χρησιμοποιούνται για την εύρεση κλίσεων στη x και y κατεύθυνση, δηλαδή υπολογίζουν παραγώγους κατά x και y της έντασης. Καθώς αποτελούνται από δύο ορθογώνιες περιοχές, εφαρμόζονται σε αθροιστικές εικόνες για γρήγορο υπολογισμό. Συγκεκριμένα, στην περίπτωση του SURF, υπολογίζονται οι αποκρίσεις Haar wavelet μεγέθους 4s στις διευθύνσεις x και y για μία ομάδα από pixels, εντός μίας κυκλικής περιοχής ακτίνας 6s με κέντρο το σημείο ενδιαφέροντος, όπου s η κλίμακα στην οποία ανιχνεύτηκε το σημείο ενδιαφέροντος. Ωστόσο, δεν υπολογίζονται οι αποκρίσεις αυτές για όλα τα pixels εντός της κυκλικής περιοχής, αλλά γίνεται δειγματοληψία σε αυτά με βήμα s. Οι αποκρίσεις στη συνέχεια σταθμίζονται με μία Γκαουσιανή συνάρτηση τυπικής απόκλισης σ=2s με κέντρο το σημείο ενδιαφέροντος [86], [87], [95], [96].



Εικόνα 2-11. Haar wavelet φίλτρα. Το αριστερό φίλτρο υπολογίζει τη`ν απόκριση στη διεύθυνση x και το δεξί στη διεύθυνση y.

Πηγή: Bay et al., 2008

Στη συνέχεια, οι αποκρίσεις αυτές παρουσιάζονται ως σημεία σε ένα επίπεδο σύστημα συντεταγμένων και ειδικότερα η οριζόντια απόκριση αποτελεί την τετμημένη και η κατακόρυφη αποτελεί την τεταγμένη. Ο κυρίαρχος προσανατολισμός εκτιμάται με τον υπολογισμό του αθροίσματος όλων των αποκρίσεων εντός ενός κυλιόμενου κυκλικού τομέα 60°. Σε κάθε θέση του κυκλικού τομέα αθροίζονται οι αποκρίσεις dx και dy, κατά τις διευθύνσεις x και y αντίστοιχα, που βρίσκονται εντός αυτού και σχηματίζουν ένα τοπικό διάνυσμα προσανατολισμού. Το μεγαλύτερο από τα διανύσματα αυτά ορίζει τον κυρίαρχο προσανατολισμό του σημείου ενδιαφέροντος, ο οποίος αναγνωρίζεται στο κέντρο του κυκλικού τομέα στον οποίο παρατηρήθηκε η μεγαλύτερη απόκριση, όπως φαίνεται στην Εικόνα 2-12 [86], [95], [96].



Εικόνα 2-12. Διαδικασία ανάθεσης προσανατολισμού του σημείου ενδιαφέροντος.

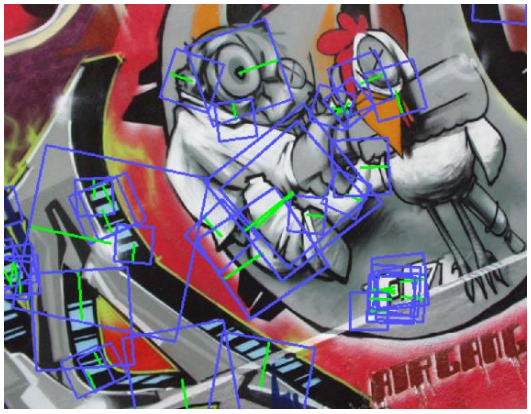
Πηγή: ftp://tracking.mat.ucsb.edu/projects/external_docs/tracking/intro_to_matching.pdf

Για κάποιες εφαρμογές, ωστόσο, δεν απαιτείται η μη μεταβλητότητα των σημείων ενδιαφέροντος σε μετασχηματισμούς στροφής. Στην περίπτωση αυτή, μπορεί να χρησιμοποιηθεί μία παραλλαγή του SURF, γνωστή ως U-SURF (Upright SURF), η οποία

παραλείπει το στάδιο ανάθεσης προσανατολισμού και με αυτόν τον τρόπο είναι πιο γρήγορη, ενώ ταυτόχρονα είναι ανθεκτική σε στροφές μέχρι 15° [86].

2.2.2.7 Εξαγωγή του διανύσματος περιγραφικών χαρακτηριστικών κάθε σημείου ενδιαφέροντος

Μετά τον καθορισμό του κυρίαρχου προσανατολισμού του σημείου ενδιαφέροντος, σχηματίζεται μία τετράγωνη περιοχή μεγέθους 20s γύρω από το σημείο αυτό, η οποία προσανατολίζεται με βάση τον παραπάνω προσανατολισμό (βλ. Εικόνα 2-13). Η περιοχή αυτή διαιρείται σε έναν κάναβο 4x4, δηλαδή αποτελείται από 16 υποπεριοχές και για κάθε μία από αυτές γίνεται υπολογισμός των Haar wavelet αποκρίσεων, με μέγεθος φίλτρου 2s, σε δείγμα 25 κανονικά κατανεμημένων σημείων της. Η διαδικασία αυτή φαίνεται στην Εικόνα 2-14. Βέβαια, για λόγους αποδοτικότητας, οι Haar wavelet αποκρίσεις υπολογίζονται στη μη στραμμένη εικόνα και στη συνέχεια παρεμβάλλονται και έτσι αποφεύγεται η περιστροφή της εικόνας. Οι αποκρίσεις dx και dy, κατά την οριζόντια και κάθετη διεύθυνση αντίστοιχα[7], σταθμίζονται με μία Γκαουσιανή συνάρτηση, τυπικής απόκλισης σ=3.3s, με κέντρο το σημείο ενδιαφέροντος [86], [95].



Εικόνα 2-13. Τετράγωνα παράθυρα που σχηματίζονται γύρω από κάθε σημείο ενδιαφέροντος, με μέγεθος εικοσαπλάσιο της κλίμακας στην οποία εντοπίστηκε και προσανατολισμένα κατά μήκος του κυρίαρχου προσανατολισμού κάθε σημείου (ο οποίος φαίνεται στην εικόνα με πράσινο χρώμα).

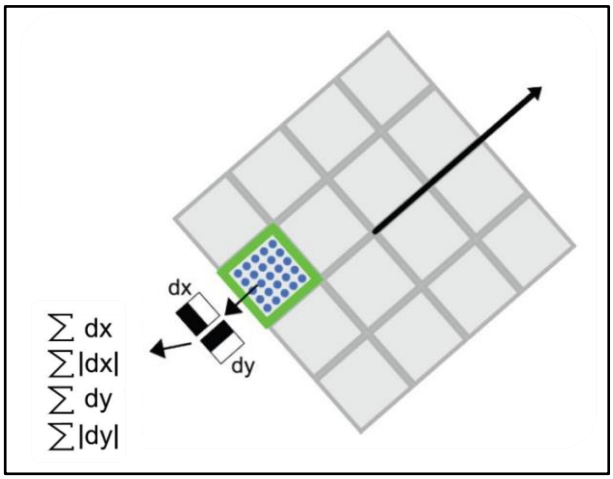
Πηγή: Evans, 2009

Στη συνέχεια, για τα 25 κανονικά κατανεμημένα σημεία του δείγματος κάθε υποπεριοχής υπολογίζονται τα αθροίσματα των αποκρίσεων dx και dy και των απόλυτων τιμών τους |dx| και |dy| και σχηματίζουν μία λίστα από εισόδους στο διάνυσμα περιγραφικών χαρακτηριστικών του σημείου. Συνεπώς, κάθε υποπεριοχή έχει ένα διάνυσμα περιγραφικών χαρακτηριστικών τεσσάρων διαστάσεων:

$$v_{\text{υποπεριοχής}} = \left[\sum dx \quad \sum dy \quad \sum |dx| \quad \sum |dy| \right] \tag{2.15}$$

[7] Η οριζόντια και κάθετη διεύθυνση ορίζονται με βάση τον επιλεγμένο προσανατολισμό του σημείου ενδιαφέροντος.

που περιγράφει την υποκείμενη δομή της έντασής της. Δεδομένου ότι ο κάναβος περιλαμβάνει 16 τέτοιες υποπεριοχές, αποτέλεσμα για κάθε σημείο ενδιαφέροντος είναι ένα διάνυσμα περιγραφικών χαρακτηριστικών 4·16=64 διαστάσεων. Αυτές οι αποκρίσεις είναι αμετάβλητες σε κάποια αύξηση ή μείωση του φωτισμού, ενώ η μη μεταβλητότητα στην αντίθεση επιτυγχάνεται με κανονικοποίηση του περιγραφέα, έτσι ώστε η ευκλείδεια νόρμα του να είναι μοναδιαία. Επομένως, το διάνυσμα περιγραφικών χαρακτηριστικών είναι αμετάβλητο σε μετασχηματισμούς στροφής, κλίμακας και φωτεινότητας, και μετά από κανονικοποίηση σε μεταβολές λόγω αντίθεσης [86], [87], [95], [100].



Εικόνα 2-14. Δημιουργία μίας τετράγωνης περιοχής γύρω από κάθε σημείο ενδιαφέροντος, προσανατολισμένης με βάση τον κυρίαρχο προσανατολισμό, και διαίρεσή της σε 16 τετράγωνες υποπεριοχές. Για κάθε μία υποπεριοχή υπολογίζεται το διάνυσμα περιγραφικών χαρακτηριστικών της από τον υπολογισμό των αποκρίσεων Haar wavelet σε δείγμα 25 κανονικά κατανεμημένων σημείων της.

Πηγή: http://www.sci.utah.edu/~fletcher/CS7960/slides/Scott.pdf

Εκτός από το παραπάνω είδος του διανύσματος περιγραφικών χαρακτηριστικών 64 διαστάσεων, δοκιμάστηκε από τους συγγραφείς του SURF και ένα ακόμα είδος διανύσματος περιγραφικών χαρακτηριστικών 128 διαστάσεων. Ο SURF-128, όπως ονομάστηκε, υπολογίζει τα αθροίσματα των dx και |dx| χωριστά για dy<0 και dy≥0, καθώς και τα αθροίσματα των dy και |dy| χωριστά για dx<0 και dx≥0. Επομένως, κάθε τοπικό διάνυσμα περιγραφικών χαρακτηριστικών είναι 8 διαστάσεων και, με βάση τις 16 υποπεριοχές, το τελικό διάνυσμα περιγραφικών χαρακτηριστικών έχει 128 διαστάσεις. Ωστόσο, η παραλλαγή αυτή του SURF είναι περισσότερο απαιτητική σε χρόνο κατά τη διαδικασία της συνταύτισης. Μία ακόμα παραλλαγή του SURF, γνωστή ως SURF-36, με 3x3=9 υποπεριοχές του παραθύρου διάστασης 20s, αντί των 16, έχει λίγο χειρότερα αποτελέσματα, σύμφωνα με πειράματα των συγγραφέων του SURF, τα οποία όμως εξακολουθούν να είναι αποδεκτά. Επιπλέον, επιταχύνεται σημαντικά η διαδικασία της εύρεσης ομολογιών [86].

2.2.2.8 Βήματα του αλγορίθμου SURF για τον προσδιορισμό των σημείων ενδιαφέροντος

Με βάση την παρουσίαση του αλγορίθμου SURF που έγινε στις προηγούμενες παραγράφους, τα βήματα που ακολουθεί για τον εντοπισμό και την περιγραφή των σημείων ενδιαφέροντος μπορούν να περιγραφούν ως ακολούθως:

1. Υπολογισμός της αθροιστικής εικόνας, με βάση την αρχική εικόνα
2. Ανίχνευση των σημείων ενδιαφέροντος
   - Υπολογισμός της ορίζουσας της παραλλαγής του Εσσιανού Πίνακα για κάθε pixel, με χρήση των φίλτρων τύπου «κουτιού» και της αθροιστικής εικόνας, και απόκτηση των χαρτών απόκρισης σε blobs για όλα τα επίπεδα κλίμακας σε όλες τις οκτάβες
   - Απόρριψη των pixels που έχουν τιμή μικρότερη ενός κατωφλίου, το οποίο μπορεί να οριστεί ανάλογα με την εκάστοτε εφαρμογή
   - Εφαρμογή της μεθόδου απαλοιφής των μη μεγίστων σε γειτονιά 3x3x3, μέσω της οποίας επιλέγονται τα εικονοστοιχεία με τιμή της ορίζουσας της παραλλαγής του Εσσιανού Πίνακα μεγαλύτερη των γειτόνων τους
   - Παρεμβολή της γειτονικής πληροφορίας για να επιτευχθεί ακρίβεια sub-pixel και sub-scale στον εντοπισμό της θέσης και της κλίμακας των σημείων ενδιαφέροντος
   - Για κάθε σημείο ενδιαφέροντος επιστροφή της θέσης του (x,y) στην εικόνα και της κλίμακας s στην οποία εντοπίστηκε και – δυνητικά – του προσήμου της Λαπλασιανής
3. Περιγραφή των σημείων ενδιαφέροντος
   - Καθορισμός του κυρίαρχου προσανατολισμού κάθε σημείου ενδιαφέροντος, με βάση πληροφορίες από μία κυκλική περιοχή γύρω από αυτό
     - Δειγματοληψία σημείων με βήμα s εντός ενός κύκλου ακτίνας 6s γύρω από κάθε σημείο ενδιαφέροντος και υπολογισμός των Haar wavelet αποκρίσεων μεγέθους 4s κατά x και y για κάθε σημείο δειγματοληψίας, με χρήση της αθροιστικής εικόνας
     - Στάθμιση των αποκρίσεων με μία Γκαουσιανή συνάρτηση τυπικής απόκλισης σ=2s με κέντρο το σημείο ενδιαφέροντος
     - Υπολογισμός του αθροίσματος των αποκρίσεων εντός ενός κυλιόμενου κυκλικού τομέα 60° με κέντρο το κάθε σημείο ενδιαφέροντος και εύρεση του τοπικού διανύσματος προσανατολισμού

  - Εύρεση του κυρίαρχου προσανατολισμού κάθε σημείου ενδιαφέροντος μέσω του εντοπισμού του μεγαλύτερου τοπικού διανύσματος προσανατολισμού
- Εξαγωγή του διανύσματος περιγραφικών χαρακτηριστικών 64 διαστάσεων κάθε σημείου ενδιαφέροντος, που αντιστοιχεί στην προσανατολισμένη και κατάλληλου μεγέθους (αντίστοιχου της κλίμακας εντοπισμού του σημείου σημείου) γειτονιά του
  - Διαίρεση μίας τετράγωνης περιοχής μεγέθους 20s, με κέντρο κάθε σημείο ενδιαφέροντος και προσανατολισμένης με βάση τον κυρίαρχο προσανατολισμό του, σε 16 τετράγωνες υποπεριοχές
  - Υπολογισμός των Haar wavelet αποκρίσεων μεγέθους 2s κατά x και y σε 25 κανονικά κατανεμημένα σημεία της κάθε υποπεριοχής για κάθε σημείο ενδιαφέροντος
  - Στάθμιση των αποκρίσεων με μία Γκαουσιανή συνάρτηση τυπικής απόκλισης σ=3.3s με κέντρο κάθε σημείο ενδιαφέροντος
  - Υπολογισμός του διανύσματος περιγραφικών χαρακτηριστικών της κάθε υποπεριοχής $\left[\sum dx \ \sum dy \ \sum |dx| \ \sum |dy|\right]$ και τελικός υπολογισμός του διανύσματος περιγραφικών χαρακτηριστικών κάθε σημείου ενδιαφέροντος, 64 διαστάσεων
  - Κανονικοποίηση του διανύσματος περιγραφικών χαρακτηριστικών κάθε σημείου ενδιαφέροντος, ώστε η νόρμα του να είναι μοναδιαία

2.3 Εύρεση των Ομόλογων Σημείων Ενδιαφέροντος

Η αντιστοίχιση των σημείων ενδιαφέροντος γίνεται με βάση τις πληροφορίες των διανυσμάτων περιγραφικών χαρακτηριστικών της γειτονιάς τους και η ολοκλήρωση της συνταύτισης επιτυγχάνεται μέσω της ελαχιστοποίησης μίας συνάρτησης κόστους, η οποία αποτελεί ένα είδος απόστασης μεταξύ των διανυσμάτων περιγραφικών χαρακτηριστικών. Συγκεκριμένα, στο τελικό αυτό στάδιο, συνήθως, κάθε διάνυσμα περιγραφικών χαρακτηριστικών της μίας εικόνας συγκρίνεται με όλους τους περιγραφείς χαρακτηριστικών τις δεύτερης εικόνας και τα σημεία για τα οποία η απόσταση των περιγραφικών χαρακτηριστικών είναι ελάχιστη, και εφόσον είναι μικρότερη από ένα κατώφλι, θεωρούνται ως ομόλογα εικονοσημεία. Ωστόσο, σε περιπτώσεις εφαρμογών, κατά τις οποίες η έρευνα για πιθανές συνταυτίσεις πρέπει να γίνει σε μία βάση δεδομένων εκατομμυρίων χαρακτηριστικών, μία τέτοια προσέγγιση δεν είναι αποδοτική και υιοθετούνται άλλες μέθοδοι, όπως η χρήση ειδικών δομών δεδομένων, όπως για παράδειγμα η ιεραρχική δομή δεδομένων k-d tree (δέντρο k-διαστάσεων), προσεγγιστικοί αλγόριθμοι εγγύτερου γείτονα

κ.ά. [82], [101]. Η ανάλυση αυτών των μεθόδων ξεφεύγει από τους σκοπούς της παρούσας εργασίας. Στο πλαίσιο των εφαρμογών που προγραμματίστηκαν, δεν είναι μεγάλο το πλήθος των σημείων ενδιαφέροντος, με αποτέλεσμα η συμβατική μέθοδος συνταύτισης να αποδίδει ικανοποιητικά αποτελέσματα.

2.3.1 Κριτήρια ταύτισης των διανυσμάτων περιγραφικών χαρακτηριστικών

Τα διανύσματα περιγραφικών χαρακτηριστικών ταυτίζονται με βάση ένα μέτρο απόστασης. Το μέτρο αυτό είναι συνήθως ένα είδος νόρμας του διανύσματος που προκύπτει ως διαφορά των διανυσμάτων των περιγραφικών χαρακτηριστικών που συγκρίνονται κάθε φορά.

Μία τέτοια νόρμα [102] μπορεί να είναι η ευκλείδεια νόρμα, γνωστή και ως νόρμα L_2, η οποία ορίζεται ως η τετραγωνική ρίζα του αθροίσματος των τετραγώνων των απόλυτων τιμών των στοιχείων του διανύσματος και χρησιμοποιείται πολύ συχνά κατά τη διαδικασία της εύρεσης ομολογιών. Έτσι, αν συμβολιστούν ως v_1 και v_2 τα διανύσματα περιγραφικών χαρακτηριστικών που συγκρίνονται, τότε η νόρμα:

$$\left\|v_1 - v_2\right\|_2 = \left(\sum_{i=1}^{n}\left|v_{1i} - v_{2i}\right|^2 = \right)^{1/2} = \left(\left|v_{11} - v_{21}\right|^2 + \left|v_{12} - v_{22}\right|^2 + ... + \left|v_{1v} - v_{2v}\right|^2\right)^{1/2} \quad (2.16)$$

υπολογίζει την ευκλείδεια απόσταση μεταξύ των δύο διανυσμάτων. Στην παραπάνω σχέση, n είναι το μέγεθος του διανύσματος περιγραφικών χαρακτηριστικών, το οποίο στην περίπτωση του SURF είναι ίσο με 64.

Άλλη νόρμα που μπορεί να χρησιμοποιηθεί είναι η αθροιστική νόρμα, ή νόρμα L1, η οποία ορίζεται ως το άθροισμα των απόλυτων τιμών των στοιχείων του διανύσματος. Ομοίως, ορίζεται το διάνυσμα $v_1 - v_2$ ως η διαφορά των διανυσμάτων περιγραφικών χαρακτηριστικών που συγκρίνονται κάθε φορά και η αθροιστική νόρμα του υπολογίζεται ως εξής:

$$\left\|v_1 - v_2\right\|_1 = \sum_{i=1}^{n}\left|v_{1i} - v_{2i}\right| = \left|v_{11} - v_{21}\right| + \left|v_{12} - v_{22}\right| + ... + \left|v_{1v} - v_{2v}\right| \quad (2.17)$$

Οι δύο αυτές νόρμες είναι συνήθεις στον υπολογισμό της απόστασης των διανυσμάτων περιγραφικών χαρακτηριστικών, αλλά όχι και οι μοναδικές που μπορούν να χρησιμοποιηθούν.

2.3.2 Εύρεση των ομόλογων σημείων ενδιαφέροντος από τον αλγόριθμο SURF

Ο αλγόριθμος SURF, στο στάδιο της εύρεσης των ομολογιών, συγκρίνει μόνο τα διανύσματα περιγραφικών χαρακτηριστικών των σημείων ενδιαφέροντος που έχουν ίδιο τύπο αντίθεσης, με βάση το ίχνος του Εσσιανού Πίνακα [86]. Μόνο εφόσον πρόκειται για σημεία ενδιαφέροντος με ίδιο τύπο αντίθεσης, προβαίνει σε σύγκριση των διανυσμάτων περιγραφικών χαρακτηριστικών τους, μέσω συνήθως της ευκλείδειας απόστασης. Σύμφωνα με τους συγγραφείς του αλγορίθμου, η τεχνική αυτή οδηγεί σε ταχύτερη εύρεση των ομολογιών και αποτελεί πλεονέκτημα στην περίπτωση χρήσης πιο προχωρημένων τεχνικών ταύτισης, όπως οι δομές δεδομένων k-d trees.

2.3.3 Απόρριψη μη έγκυρων ομολογιών

Γενικά, οι αντιστοιχίσεις σημείων που επιστρέφονται από τον αλγόριθμο συνταύτισης μπορεί να ανήκουν σε μία από τις τέσσερις ακόλουθες κατηγορίες, εκ των οποίων οι δύο πρώτες είναι επιθυμητές, ενώ οι δύο τελευταίες ανεπιθύμητες.

- «True Positive». Πρόκειται για τις ομολογίες σημείων οι οποίες επιστράφηκαν από τον αλγόριθμο και είναι σωστές.
- «True Negative». Πρόκειται για τα τις μη ομολογίες σημείων, οι οποίες σωστά απορρίφθηκαν από τον αλγόριθμο και δεν θεωρήθηκαν ως αντιστοιχίσεις.
- «False Positive». Πρόκειται για τις αντιστοιχίσεις σημείων που επιστράφηκαν από τον αλγόριθμο, αλλά δεν είναι σωστές.
- «False Negative». Πρόκειται για τις αντιστοιχίσεις σημείων οι οποίες δεν επιστράφηκαν από τον αλγόριθμο, μολονότι τα σημεία αυτά είναι ομόλογα και ορατά και στις δύο εικόνες.

Όπως αναφέρθηκε παραπάνω, η πιο απλή στρατηγική συνταύτισης θεωρεί ως ομόλογα τα σημεία εφόσον η απόσταση των διανυσμάτων περιγραφικών χαρακτηριστικών τους είναι η ελάχιστη και κάτω από ένα κατώφλι, το οποίο αντιστοιχεί σε μία μέγιστη απόσταση. Ωστόσο, ο ορισμός ενός υψηλού κατωφλίου κρύβει τον κίνδυνο μη σωστών αντιστοιχίσεων μεταξύ των σημείων ενδιαφέροντος, δηλαδή αυξάνεται η πιθανότητα επιστροφής πολλών «False Positive» αντιστοιχίσεων. Αντίθετα, ορίζοντας μία μικρή τιμή απόστασης ως κατώφλι, υφίσταται ο κίνδυνος να μην ταυτίζονται πολλά ομόλογα εικονοσημεία, μολονότι υπάρχουν και στις δύο εικόνες κάθε φορά, δηλαδή υπάρχει μεγάλη πιθανότητα αυξημένου αριθμού «False Negative». Για το λόγο αυτό, ο ορισμός ενός κατωφλίου πολλές φορές δεν αρκεί και πρέπει να γίνουν και κάποιες άλλες τεχνικές, επιπλέον του ορισμού του κατωφλίου ή χωρίς αυτόν, για την απόρριψη των άκυρων ομολογιών, γνωστών διεθνώς ως outliers. Πρόκειται για τις αντιστοιχίσεις τύπου «False Positive», καθώς για τις «False Negative» δεν είναι

δυνατή η παρέμβαση και διόρθωση με κάποια τεχνική, παρά μόνο αυξάνοντας το κατώφλι, καθώς είναι αυτές που απέρριψε ο αλγόριθμος συνταύτισης.

Στόχος λοιπόν είναι η ελαχιστοποίηση – ή και η εξάλειψη, εφόσον καθίσταται δυνατή – των άκυρων ομολογιών, η οποία επιτυγχάνεται με διάφορες μεθόδους. Στις επόμενες παραγράφους αναφέρονται ενδεικτικά τρεις μέθοδοι, οι δύο πρώτες εκ των οποίων βασίζονται σε πιο απλή λογική σε σχέση με την τελευταία, η οποία αφορά στο γεωμετρικό έλεγχο των ομολογιών.

2.3.3.1 Λόγος των αποστάσεων των εγγύτερων γειτόνων (ratio test)

Μία μέθοδος ελαχιστοποίησης των μη έγκυρων ομολογιών είναι το επονομαζόμενο ratio test. Στηρίζεται στο λόγο των δύο μικρότερων αποστάσεων των διανυσμάτων περιγραφικών χαρακτηριστικών, που είναι γνωστός και ως λόγος αποστάσεων εγγύτερων γειτόνων (Nearest Neighbour Distance Ratio - NNDR), και ορίζεται ως εξής:

$$\mathrm{NNDR} = \frac{\left\| \mathbf{v}_1 - \mathbf{v}_2 \right\|}{\left\| \mathbf{v}_1 - \mathbf{v}_3 \right\|} \tag{2.18}$$

όπου v_1 είναι το διάνυσμα περιγραφικών χαρακτηριστικών του σημείου για το οποίο γίνεται προσπάθεια εύρεσης του ομολόγου του, v_2 είναι το κοντινότερο διάνυσμα περιγραφικών χαρακτηριστικών σε αυτό και v_3 ο δεύτερος εγγύτερος γείτονας. Εφόσον ο λόγος αυτός είναι μικρότερος από μία προκαθορισμένη τιμή, εντός του διαστήματος (0,1), τότε μόνο θεωρείται ότι το σημείο ενδιαφέροντος με διάνυσμα περιγραφικών χαρακτηριστικών v_2 ταυτίζεται με το χαρακτηριστικό σημείο με περιγραφέα v_1. Με τον ορισμό λοιπόν μίας σχετικά μικρής τιμής του λόγου αυτού, επιστρέφονται από τον αλγόριθμο ομολογίες χαρακτηριστικών που θεωρούνται μοναδικά σε σχέση με το περιβάλλον τους και έχουν έντονη διακριτότητα. Με τον τρόπο αυτό, μπορούν να μειωθούν σημαντικά οι μη έγκυρες αντιστοιχίσεις.

2.3.3.2 Διασταυρωμένος έλεγχος (cross-check test)

Άλλος τρόπος μείωσης των μη έγκυρων ομολογιών είναι ο λεγόμενος διασταυρωμένος έλεγχος, κοινώς γνωστός ως cross-check test. Σύμφωνα με αυτόν, γίνεται σύγκριση κάθε διανύσματος περιγραφικών χαρακτηριστικών της μίας εικόνας με όλα τα διανύσματα περιγραφικών χαρακτηριστικών της δεύτερης εικόνας και επιστρέφονται κάποιες υποψήφιες ομολογίες. Αυτές ελέγχονται και μέσω της αντίστροφης σύγκρισης, δηλαδή της σύγκρισης του διανύσματος περιγραφικών χαρακτηριστικών της δεύτερης εικόνας με τα διανύσματα περιγραφικών χαρακτηριστικών της πρώτης. Μόνο εφόσον προκύψουν τα ίδια αποτελέσματα από την αμφίδρομη διαδικασία της σύγκρισης των διανυσμάτων, θεωρείται το ζεύγος σημείων ως ομολογία [69], [75].

2.3.3.3 Αλγόριθμος RANSAC

Μία διαφορετική τεχνική απόρριψης των άκυρων ομολογιών, σε σχέση με τις προηγούμενες, είναι ο έλεγχος αν αυτές επαληθεύουν τη γεωμετρική σχέση που συνδέει τις δύο εικόνες. Για το σκοπό αυτό υφίστανται διάφοροι αλγόριθμοι, εκ των οποίων ο συνηθέστερα χρησιμοποιούμενος είναι ο RANSAC (RANdom SAmple Consensus). Πρόκειται για αλγόριθμο που υπολογίζει τις παραμέτρους ενός μαθηματικού μοντέλου από ένα σύνολο δεδομένων, πολλά εκ των οποίων μπορεί να είναι λανθασμένα, και παρουσιάστηκε από τους Martin Fischler και Robert Bolles το 1981 [103]. Ο RANSAC μπορεί να εφαρμοστεί για τον εντοπισμό των χονδροειδών σφαλμάτων στις ομολογίες μέσω του υπολογισμού της γεωμετρικής σχέσης μεταξύ των εικόνων. Στηρίζεται στη λογική της χρήσης των ελάχιστων δυνατών δεδομένων και βασίζεται στο γεγονός ότι η πιθανότητα εύρεσης ενός καλού δείγματος δεδομένων από το σύνολο των δεδομένων, με την τυχαία επιλογή μίας ομάδας παρατηρήσεων, είναι μεγάλη. Συνεπώς, στοχεύει στη σημαντική μείωση του αριθμού των αναγκαίων δοκιμών στην περίπτωση μεγάλου αριθμού παρατηρήσεων, χωρίς ωστόσο να εγγυάται πάντα μία λύση χωρίς χονδροειδή σφάλματα [104]. Στην περίπτωση της χρήσης του στην απαλοιφή των μη έγκυρων ομολογιών, το μοντέλο είναι ο θεωρούμενος γεωμετρικός μετασχηματισμός μεταξύ των δύο εικόνων και τα δεδομένα είναι τα ομόλογα σημεία τα οποία εντοπίστηκαν από τον αλγόριθμο ψηφιακής συνταύτισης.

Τα βήματα του αλγορίθμου RANSAC, προσαρμοσμένα στον υπολογισμό του γεωμετρικού μετασχηματισμού μεταξύ των δύο εικόνων με την ταυτόχρονη εύρεση των λανθασμένων ομολογιών, μπορούν να περιγραφούν ως εξής:

- Από το σύνολο των ομολογιών επιλέγεται ένα τυχαίο δείγμα, το οποίο αποτελείται από τον ελάχιστο αριθμό των ομόλογων σημείων που είναι απαραίτητα για τον ορισμό του γεωμετρικού μετασχηματισμού μεταξύ των δύο εικόνων. Η επιλογή αυτή μπορεί να ελέγχεται, προκειμένου να εντοπίζονται διαφορετικές ομάδες ομόλογων σημείων σε κάθε προσπάθεια, ή ακόμη για να αποφεύγεται η δημιουργία μη «σταθερών» τυχαίων δειγμάτων. Αναφορικά με την τελευταία, μπορεί να καθοριστεί η μη επιλογή γειτονικών σημείων ως μελών του ίδιου τυχαίου δείγματος, καθώς η λύση που θα υπολογιζόταν μέσω ενός τυχαίου δείγματος, του οποίου όλα τα σημεία είναι συγκεντρωμένα σε μία περιοχή, δεν θα ήταν «σταθερή», με αποτέλεσμα η εκτίμηση της γεωμετρικής σχέσης των δύο εικόνων να ήταν ενδεχομένως λανθασμένη. Μία άλλη προσέγγιση είναι η ομαδοποίηση των ομολογιών ανά περιοχή και η επιλογή των ομόλογων σημείων, που θα αποτελέσουν τα μέλη του τυχαίου δείγματος, από διαφορετικές ομάδες.
- Με βάση το τυχαίο αυτό δείγμα γίνεται εκτίμηση των παραμέτρων που ορίζουν τη γεωμετρική σχέση μεταξύ των δύο εικόνων.

- Υπολογίζεται ο αριθμός των έγκυρων ομολογιών για τη συγκεκριμένη λύση, που δείχνει την ποιότητα της υπολογισμένης γεωμετρικής σχέσης. Συγκεκριμένα, για κάθε σημείο της μίας εικόνας, εκτός από εκείνα που ανήκουν στο δείγμα που χρησιμοποιήθηκε, υπολογίζεται η απόσταση (συνήθως η ευκλείδεια) μεταξύ της υπολογισμένης θέσης του στην άλλη εικόνα, με βάση το γεωμετρικό μετασχηματισμό, και της παρατηρούμενης θέσης του ομολόγου του, όπως προέκυψε από τη συνταύτιση. Αν η απόσταση αυτή είναι μικρότερη από ένα κατώφλι που έχει οριστεί, η ομολογία θεωρείται έγκυρη (inlier) για το υπολογισμένο μοντέλο από το συγκεκριμένο δείγμα, διαφορετικά θεωρείται άκυρη (outlier). Το κατώφλι αυτό είναι συνήθως πολλαπλάσιο της διάστασης του εικονοστοιχείου.
- Αν ο αριθμός των inliers είναι μεγαλύτερος από ένα κατώφλι, το οποίο ορίζεται ανάλογα με την εκάστοτε εφαρμογή, το μοντέλο που προέκυψε γίνεται δεκτό και ο αλγόριθμος τερματίζεται με επιτυχία, έχοντας υπολογίσει τις έγκυρες ομολογίες. Διαφορετικά, αν δεν έχει εντοπιστεί ο αναγκαίος αριθμός των inliers και αν τα παραπάνω βήματα έχουν επαναληφθεί N φορές, όπου N ο μέγιστος αριθμός των επαναλήψεων που συνήθως δίνεται από τη σχέση (2.19), ο αλγόριθμος τερματίζεται με αποτυχία. Στην περίπτωση αυτή, ενώ δεν έχει βρεθεί η λύση με την επιθυμητή «ποιότητα», έχει υπολογιστεί η λύση που οδηγεί στο μεγαλύτερο αριθμό των inliers, ο οποίος είναι μικρότερος του κατωφλίου. Διαφορετικά, αν ο αριθμός των επαναλήψεων των παραπάνω βημάτων είναι μικρότερος από N και δεν έχει εντοπιστεί ο αναγκαίος αριθμός των inliers, τα τέσσερα αυτά βήματα επαναλαμβάνονται εκ νέου.

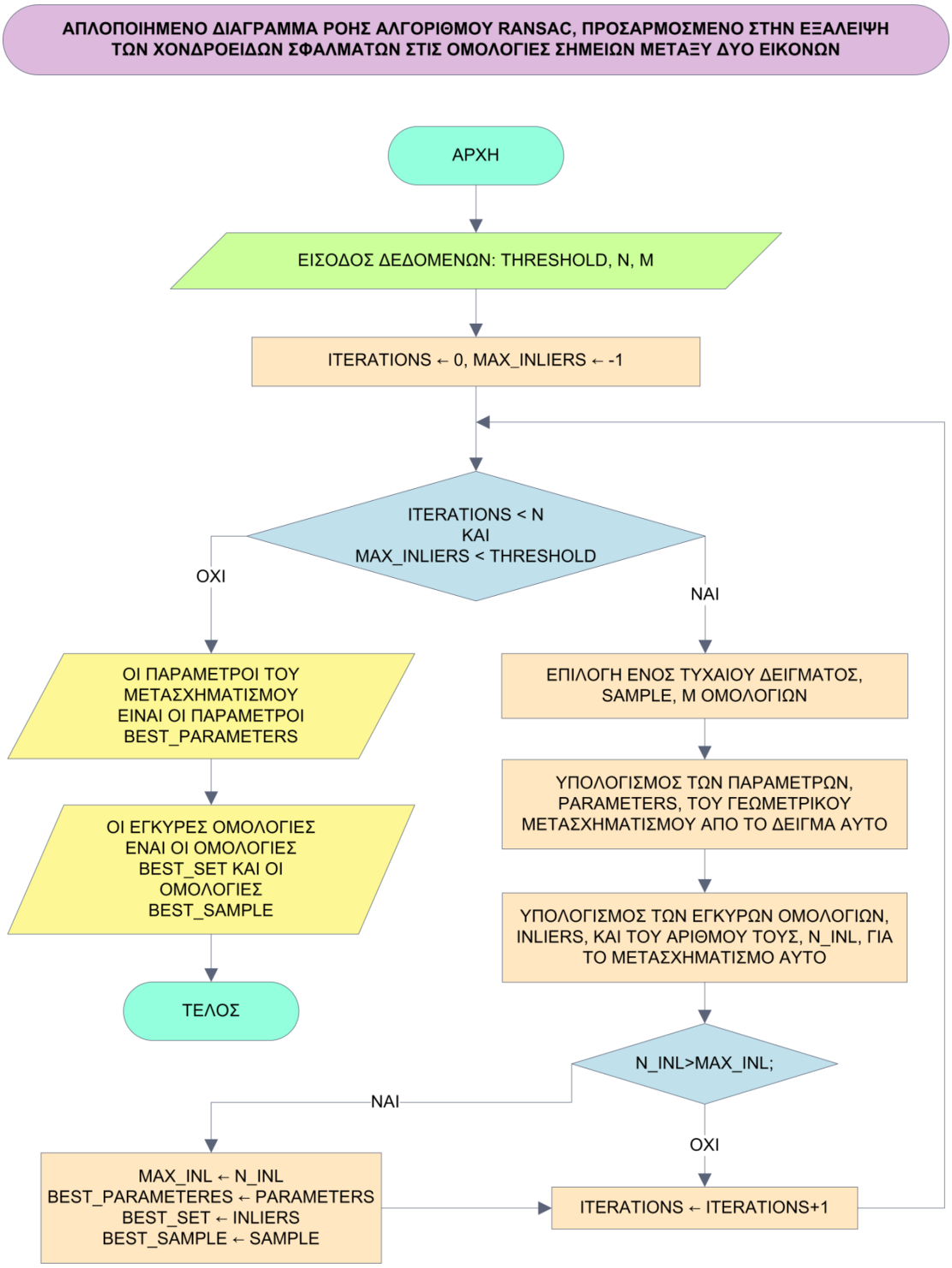
Με βάση τη διαδικασία αυτή, επιλέγεται η λύση που οδηγεί στο μεγαλύτερο αριθμό των έγκυρων αντιστοιχίσεων μεταξύ των σημείων ενδιαφέροντος. Η λύση αυτή ελαχιστοποιεί τα χονδροειδή σφάλματα στις ομολογίες, με αποτέλεσμα να βρίσκεται σε καλύτερη «συμφωνία» (consensus) με τα δεδομένα εισόδου, όπως άλλωστε δηλώνει και το όνομα του αλγορίθμου.

Ο αριθμός N των επαναλήψεων που απαιτούνται ώστε κάποιο από τα δείγματα να περιλαμβάνει μόνο έγκυρες ομολογίες (δηλαδή να οδηγεί σε σωστό εντοπισμό των χονδροειδών σφαλμάτων) με πιθανότητα p, δίνεται από την ακόλουθη σχέση:

$$N = \frac{\log\left(1-p\right)}{\log\left(1-\left(1-v\right)^{m}\right)} \tag{2.19}$$

όπου v η εκτίμηση του ποσοστού των μη έγκυρων ομολογιών και m ο αριθμός των αναγκαίων παραμέτρων που ορίζουν το μοντέλο [83], [94], [103], [104], [105], [106].

Ένα απλοποιημένο διάγραμμα ροής του αλγορίθμου RANSAC, στην περίπτωση της εύρεσης της γεωμετρικής σχέσης μεταξύ των δύο εικόνων, με ταυτόχρονη απαλοιφή των μη έγκυρων ομολογιών, παρατίθεται στη Διαγραμματική παρουσίαση 2-1[8].



Διαγραμματική παρουσίαση 2-1. Απλοποιημένο διάγραμμα ροής του αλγορίθμου RANSAC, προσαρμοσμένο στην εύρεση των χονδροειδών σφαλμάτων στις ομολογίες σημείων.

[8] Για τα χρησιμοποιούμενα σύμβολα των διαγραμμάτων ροής, βλ. Παράρτημα Α.

ΚΕΦΑΛΑΙΟ 3

ΚΕΝΤΡΙΚΗ ΠΡΟΒΟΛΗ ΚΑΙ ΕΠΙΛΥΣΗ ΒΑΣΙΚΩΝ ΦΩΤΟΓΡΑΜΜΕΤΡΙΚΩΝ ΠΡΟΒΛΗΜΑΤΩΝ

Το κεφάλαιο αυτό περιγράφει τα μαθηματικά μοντέλα που χρησιμοποιούνται στη φωτογραμμετρία και στην όραση υπολογιστών και εκφράζουν την κεντρική προβολή, η οποία μοντελοποιεί τη διαδικασία της φωτογράφησης. Για την πληρέστερη κατανόηση του προβολικού μετασχηματισμού, ο οποίος υιοθετείται από την όραση υπολογιστών αντί της συνθήκης συγγραμμικότητας, κρίνεται αναγκαία η αναφορά σε έννοιες στις οποίες αυτός στηρίζεται, όπως είναι η προβολική γεωμετρία και οι ομογενείς συντεταγμένες. Επίσης, γίνεται μία παρουσίαση των βασικών δισδιάστατων γραμμικών μετασχηματισμών με έμφαση στον προβολικό μετασχηματισμό στο χώρο των δύο διαστάσεων, ο οποίος χρησιμοποιήθηκε για τις ανάγκες των εφαρμογών της παρούσας εργασίας. Τέλος, παρουσιάζονται δύο βασικές φωτογραμμετρικές διαδικασίες, οι οποίες χρησιμοποιήθηκαν για τις εφαρμογές. Πρόκειται για τον υπολογισμό των στοιχείων του εσωτερικού προσανατολισμού της φωτογραφικής μηχανής και των στοιχείων του εξωτερικού προσανατολισμού της εικόνας. Τα προβλήματα αυτά επιλύονται στο παρόν κεφάλαιο θεωρητικά, μέσω της προσέγγισης που προτείνεται από την κοινότητα της όρασης υπολογιστών.

3.1 Προβολική Γεωμετρία

Η προβολική γεωμετρία είναι μία μη μετρική μορφή γεωμετρίας, δηλαδή δεν βασίζεται στην έννοια της απόστασης, και ασχολείται κατά βάση με τις ιδιότητες γεωμετρικών οντοτήτων και σχηματισμών υπό προβολή. Μπορεί να περιγράψει την απεικόνιση του χώρου και μέσω της κεντρικής προβολής, η οποία είναι το μαθηματικό μοντέλο το οποίο περιγράφει ικανοποιητικά τη διαδικασία της φωτογράφησης, και έτσι υιοθετείται τόσο από τη φωτογραμμετρία, όσο – κυρίως – και από την όραση υπολογιστών. Συνιστά μία γενικότερη γεωμετρία που ενσωματώνει τις υπόλοιπες (αφινική, ευκλείδεια κ.λπ.) και μελετά τις αναλλοίωτες ιδιότητες των προβολικών μετασχηματισμών. Στο πλαίσιο της προβολικής

γεωμετρίας, οι μετασχηματισμοί εκφράζονται μέσω ομογενών συντεταγμένων (βλ. ενότητα 3.2) και τα συστήματα εξισώσεων καταλήγουν να είναι γραμμικά [94], [107], [108].

3.2 Ομογενείς Συντεταγμένες

Οι ομογενείς συντεταγμένες (homogeneous coordinates) είναι ένα είδος συντεταγμένων που χρησιμοποιείται στην προβολική γεωμετρία, αντίστοιχα με τις καρτεσιανές συντεταγμένες που χρησιμοποιούνται στην ευκλείδεια γεωμετρία. Προτάθηκαν από τον August Ferdinand Möbius το 1827 και τυγχάνουν ευρείας αποδοχής σε διάφορα επιστημονικά πεδία, όπως στη φωτογραμμετρία και στην όραση υπολογιστών. Στα πλεονεκτήματά τους συγκαταλέγονται η δυνατότητα αναπαράστασης όλων των σημείων, ακόμα και των σημείων του απείρου, με πεπερασμένες συντεταγμένες, καθώς και η δυνατότητα έκφρασης ενός προβολικού μετασχηματισμού με τη βοήθεια ενός μόνο πίνακα [82], [109].

3.2.1 Ομογενείς συντεταγμένες σημείων στο προβολικό επίπεδο - θεωρητική προσέγγιση

Μία ευθεία του επιπέδου ορίζεται από την εξίσωση:

$$a \cdot x + b \cdot y + c = 0 \tag{3.1}$$

και ένα σημείο $\begin{bmatrix} x & y \end{bmatrix}^T$ ανήκει στην ευθεία αυτή εφόσον επαληθεύει την παραπάνω σχέση. Η εξίσωση (3.1) μπορεί να γραφεί ως το εσωτερικό γινόμενο δύο διανυσμάτων, των $\begin{bmatrix} a & b & c \end{bmatrix}$ και $\begin{bmatrix} x & y & 1 \end{bmatrix}^T$. Έτσι η ακόλουθη εξίσωση:

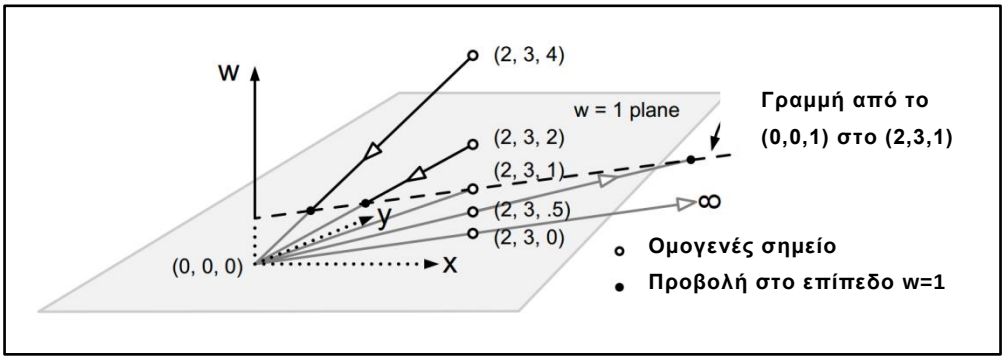
$$\begin{bmatrix} a & b & c \end{bmatrix} \cdot \begin{bmatrix} x & y & 1 \end{bmatrix}^T = 0 \tag{3.2}$$

αποτελεί ισοδύναμη αναπαράσταση ευθείας του επιπέδου, καθώς ισχύει: $\begin{bmatrix} a & b & c \end{bmatrix} \cdot \begin{bmatrix} x & y & 1 \end{bmatrix}^T = a \cdot x + b \cdot y + c$. Συνεπώς, το σημείο $\begin{bmatrix} x & y \end{bmatrix}^T$ του ευκλείδειου χώρου $\mathbf{R^2}$ αναπαρίσταται από ένα διάνυσμα τριών διαστάσεων, το $\begin{bmatrix} x & y & 1 \end{bmatrix}^T$. Ωστόσο, οι εξισώσεις $a \cdot x + b \cdot y + c = 0$ και $\lambda \cdot (a \cdot x) + \lambda \cdot (b \cdot y) + \lambda \cdot c = 0$, με $\lambda \neq 0$, αναπαριστούν την ίδια ευθεία. Επομένως, η εξίσωση (3.2) μπορεί να γραφεί και ως εξής:

$$\begin{bmatrix} a & b & c \end{bmatrix} \cdot \begin{bmatrix} \lambda \cdot x & \lambda \cdot y & \lambda \end{bmatrix}^T = 0 \quad \text{για } \lambda \neq 0 \tag{3.3}$$

καθώς $\begin{bmatrix} a & b & c \end{bmatrix} \cdot \begin{bmatrix} \lambda \cdot x & \lambda \cdot y & \lambda \end{bmatrix}^T = a \cdot (\lambda \cdot x) + b \cdot (\lambda \cdot y) + c \cdot \lambda = \lambda \cdot (a \cdot x) + \lambda \cdot (b \cdot y) + \lambda \cdot c$. Συνεπώς, τα διανύσματα της μορφής $\begin{bmatrix} \lambda \cdot x & \lambda \cdot y & \lambda \end{bmatrix}^T$ για οποιαδήποτε μη μηδενική τιμή του λ θεωρούνται ισοδύναμα και καλούνται ομογενή, αφού αποτελούν αναπαράσταση του ίδιου σημείου $\begin{bmatrix} x & y \end{bmatrix}^T$ του ευκλείδειου χώρου $\mathbf{R^2}$ [107].

Έστω ένα σημείο Ο του τρισδιάστατου χώρου, από το οποίο διέρχονται όλες οι πιθανές ευθείες γραμμές. Αν αυτές τμηθούν από ένα επίπεδο Π, το οποίο δεν περιέχει το Ο, τότε κάθε σημείο του Π σχετίζεται με μία από αυτές τις γραμμές και συνδέεται μέσω αυτής με το Ο. Ωστόσο, οι γραμμές που διέρχονται από το Ο και είναι παράλληλες στο επίπεδο Π δεν τέμνονται με αυτό. Συμβατικά, θεωρείται ότι οι παράλληλες στο επίπεδο γραμμές το τέμνουν στα ιδεατά σημεία. Το ίδιο συμβαίνει και με τα επίπεδα που διέρχονται από το σημείο Ο. Στην περίπτωση που ένα τέτοιο επίπεδο είναι παράλληλο με το επίπεδο Π, θεωρείται ότι το τέμνει στην ιδεατή γραμμή. Το προβολικό επίπεδο, συνεπώς, δεν μπορεί να αναπαρασταθεί μέσω του καρτεσιανού συστήματος συντεταγμένων, που χρησιμοποιείται στην ευκλείδεια γεωμετρία, και έτσι αναπαρίσταται μέσω των ομογενών συντεταγμένων. Συμβολίζεται ως $\mathbf{P^2}$, είναι αυτό που περιλαμβάνει τα ιδεατά σημεία και τις ιδεατές γραμμές και ορίζεται ως ένα κανονικό επίπεδο, επαυξημένο με τα σημεία του απείρου. Τα τελευταία είναι τα σημεία $\begin{bmatrix} x & y & 0 \end{bmatrix}^T$, των οποίων η τρίτη μηδενική συντεταγμένη υποδηλώνει ότι το προβαλλόμενο στο επίπεδο σημείο βρίσκεται σε άπειρη απόσταση από το Ο, κατά τη διεύθυνση $\begin{bmatrix} x & y \end{bmatrix}^T$ [110] (βλ. Εικόνα 3-1). Ο δισδιάστατος προβολικός χώρος $\mathbf{P^2}$, λοιπόν, σχηματίζεται από το σύνολο των ομογενών διανυσμάτων του τρισδιάστατου ευκλείδειου χώρου $\mathbf{R^3}$, εξαιρούμενου του διανύσματος $\begin{bmatrix} 0 & 0 & 0 \end{bmatrix}^T$ [107].

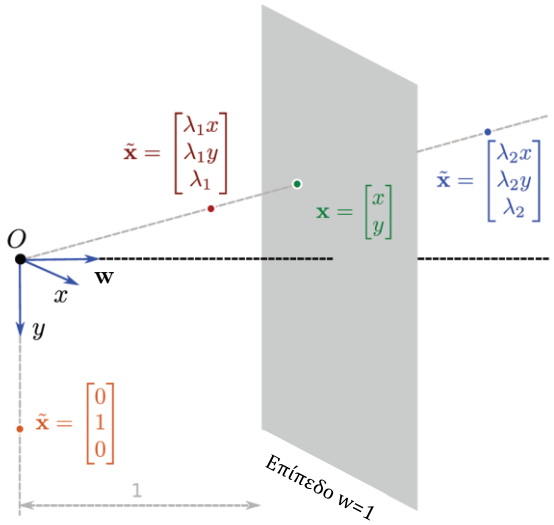


Εικόνα 3-1. Προβολή στο επίπεδο w=1 των ομογενών διανυσμάτων $[x\ y\ w_i]^T$ για διάφορες τιμές του i.

Πηγή: Bloomenthal & Rokne, 1994

Όπως υποδεικνύεται στην Εικόνα 3-2, το σύνολο των ομογενών διανυσμάτων $\begin{bmatrix} \lambda \cdot x & \lambda \cdot y & \lambda \cdot w \end{bmatrix}^T$, που ορίζονται για $\lambda \neq 0$, κείνται επί ευθείας, η οποία διέρχεται από το Ο. Το δισδιάστατο σημείο του επιπέδου w, που βρίσκεται στο σημείο τομής της ευθείας αυτής με το επίπεδο w, αντιστοιχεί στο διάνυσμα $\begin{bmatrix} x' & y' \end{bmatrix}^T = \begin{bmatrix} x/w & y/w \end{bmatrix}^T$ εκφρασμένο σε καρτεσιανές συντεταγμένες. Ειδικότερα, στην περίπτωση του επιπέδου $w = 1$ η εικονιζόμενη ευθεία αντιστοιχεί στα πολλαπλάσια του ομογενούς διανύσματος $\begin{bmatrix} x & y & 1 \end{bmatrix}^T$, δηλαδή στο σύνολο των διανυσμάτων $\begin{bmatrix} \lambda \cdot x & \lambda \cdot y & \lambda \end{bmatrix}^T$, και το σημείο τομής της με το επίπεδο αυτό είναι το σημείο με καρτεσιανές συντεταγμένες $\begin{bmatrix} x & y \end{bmatrix}^T$. Αντίστροφα, το σημείο $\begin{bmatrix} x & y \end{bmatrix}^T$ του επιπέδου αντιστοιχεί στα ομογενή διανύσματα $\begin{bmatrix} \lambda \cdot x & \lambda \cdot y & \lambda \end{bmatrix}^T$. Στην ίδια εικόνα, το

ομογενές διάνυσμα $\begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^T$ ορίζει μία ευθεία παράλληλη στο επίπεδο w=1 και αναπαριστά το λεγόμενο ιδεατό σημείο, ή σημείο του απείρου στη διεύθυνση $\begin{bmatrix} 0 & 1 \end{bmatrix}^T$ [111].



Εικόνα 3-2. Γεωμετρική ερμηνεία των ομογενών συντεταγμένων - προβολή στο επίπεδο w=1.

Πηγή: Prince, 2012

3.2.2 Περιγραφή και πλεονεκτήματα των ομογενών συντεταγμένων

Σύμφωνα με όσα αναφέρθηκαν στην υποενότητα 3.2.1, ο αριθμός των ομογενών συντεταγμένων που απαιτούνται για να εκφραστεί ένα σημείο στον προβολικό χώρο n διαστάσεων $\mathbf{P^n}$ είναι n+1. Επίσης, οι ομογενείς συντεταγμένες ενός σημείου όταν πολλαπλασιάζονται με ένα μη μηδενικό συντελεστή αναπαριστούν το ίδιο σημείο. Από την ιδιότητά τους αυτή προκύπτει το γεγονός ότι ενώ η μετατροπή από τις ομογενείς συντεταγμένες στις μη ομογενείς είναι μοναδική, δεν συμβαίνει το ίδιο και για την αντίστροφη μετατροπή. Έτσι, ένα σημείο του ευκλείδειου χώρου $\mathbf{R^n}$ με συντεταγμένες $\begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}^T$ αντιστοιχεί στο σημείο του προβολικού χώρου $\mathbf{P^n}$ με συντεταγμένες $\begin{bmatrix} x_1 & x_2 & \dots & x_n & 1 \end{bmatrix}^T$ και γενικότερα στο σημείο $\lambda \cdot \begin{bmatrix} x_1 & x_2 & \dots & x_n & 1 \end{bmatrix}^T$, όπου λ είναι ένας μη μηδενικός συντελεστής. Αντίστροφα, ένα σημείο που ορίζεται στον προβολικό χώρο $\mathbf{P^n}$ από το διάνυσμα $\begin{bmatrix} x_1 & x_2 & \dots & x_n & x_{n+1} \end{bmatrix}^T$ αναπαρίσταται στον ευκλείδειο χώρο $\mathbf{R^n}$ από το μη ομογενές διάνυσμα $\begin{bmatrix} x_1 / x_{n+1} & x_2 / x_{n+1} & \dots & x_n / x_{n+1} \end{bmatrix}^T$. Από τα παραπάνω γίνεται κατανοητό ότι ένα σημείο του προβολικού χώρου $\mathbf{P^n}$ με $x_{n+1} = 0$, δηλαδή το σημείο $\begin{bmatrix} x_1 & x_2 & \dots & x_n & 0 \end{bmatrix}^T$, που ονομάζεται, όπως ήδη αναφέρθηκε, ιδεατό σημείο ή σημείο του απείρου, δεν μπορεί να αποδοθεί ως σημείο του ευκλείδειου χώρου $\mathbf{R^n}$, αν και στο πλαίσιο της προβολικής γεωμετρίας αντιμετωπίζεται όπως όλα τα άλλα σημεία. [94], [107], [109], [112].

Για την πληρέστερη κατανόηση των παραπάνω, θεωρείται ένα σημείο της εικόνας με καρτεσιανές συντεταγμένες $\begin{bmatrix} x & y \end{bmatrix}^T$. Το αντίστοιχο σημείο, σε ομογενείς συντεταγμένες,

μπορεί να αναπαρασταθεί από το διάνυσμα $\lambda \cdot \begin{bmatrix} x & y & 1 \end{bmatrix}^T$, όπου $\lambda \neq 0$. Για παράδειγμα, τα ομογενή διανύσματα $\begin{bmatrix} 2 & 4 & 1 \end{bmatrix}^T$, $\begin{bmatrix} 4 & 8 & 2 \end{bmatrix}^T$ και $\begin{bmatrix} 6 & 12 & 3 \end{bmatrix}^T$ (για λ = 1, 2 και 3 αντίστοιχα) αναπαριστούν το ίδιο σημείο της εικόνας, που εκφράζεται σε καρτεσιανές συντεταγμένες από το διάνυσμα $\begin{bmatrix} 2 & 4 \end{bmatrix}^T$. Η μετατροπή της καρτεσιανής αναπαράστασης του σημείου της εικόνας $\begin{bmatrix} x & y \end{bmatrix}^T$ σε ομογενείς συντεταγμένες γίνεται επιλέγοντας λ=1 και προσθέτοντας μία τρίτη συντεταγμένη ίση με τη μονάδα στις προηγούμενες δύο. Δηλαδή η ομογενής αναπαράστασή του θεωρείται η $\begin{bmatrix} x & y & 1 \end{bmatrix}^T$. Αντίστοιχα, η μετατροπή από ομογενείς σε καρτεσιανές συντεταγμένες γίνεται διαιρώντας τα δύο πρώτα στοιχεία του διανύσματος με το τρίτο. Δηλαδή, από το ομογενές διάνυσμα $\begin{bmatrix} x_1 & y_1 & w_1 \end{bmatrix}^T$ με $w_1 \neq 0$ προκύπτει το καρτεσιανό διάνυσμα του δισδιάστατου σημείου $\begin{bmatrix} x & y \end{bmatrix}^T$ ως εξής: $x = x_1/w_1$ και $y = y_1/w_1$. Ακριβώς αντίστοιχη είναι και η λογική για σημεία του τρισδιάστατου χώρου, τα οποία ενώ σε καρτεσιανές συντεταγμένες αναπαρίστανται ως διανύσματα τριών διαστάσεων, στον προβολικό χώρο ορίζονται ως διανύσματα τεσσάρων διαστάσεων.

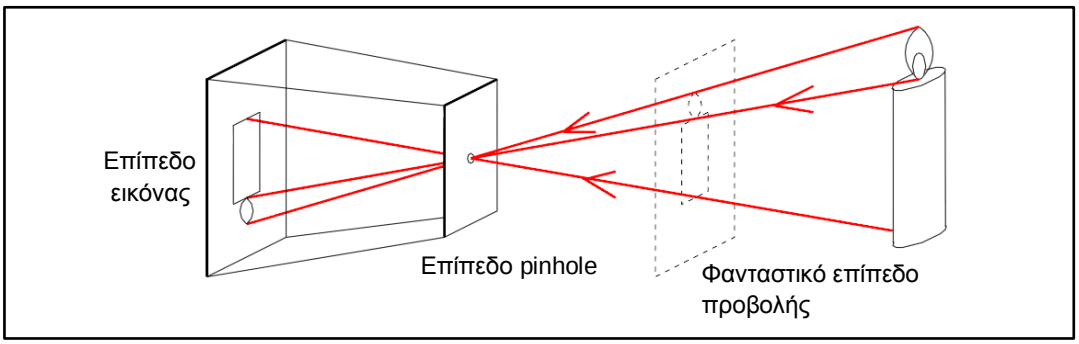
Οι ομογενείς συντεταγμένες χρησιμοποιούνται ευρέως σε προβλήματα της όρασης υπολογιστών και αποτελούν βολικές αναπαραστάσεις των σημείων της εικόνας και του τρισδιάστατου χώρου κατά την εφαρμογή διαφόρων μετασχηματισμών. Το πλεονέκτημά τους έγκειται στη δυνατότητα που δίνεται μέσω αυτών να εκφραστούν οι διάφοροι μετασχηματισμοί μέσω ενός μόνο ενιαίου πίνακα, γεγονός το οποίο οδηγεί σε μεγαλύτερη ευκολία κατά τη σύνθεση διαδοχικών μετασχηματισμών. Έτσι, αυτοί που περιλαμβάνουν μετάθεση ταυτόχρονα με κάποιο άλλο μετασχηματισμό ή με σύνθεση διαφόρων άλλων μετασχηματισμών (όπως στροφή, αλλαγή κλίμακας και στροφή κ.ο.κ.) εκφράζονται με τη βοήθεια ενός πίνακα. Αντίθετα, στην περίπτωση χρήσης καρτεσιανών συντεταγμένων, οι ίδιοι μετασχηματισμοί, εφόσον περιλαμβάνουν μετάθεση μαζί με κάποιο άλλο μετασχηματισμό, εκφράζονται μέσω δύο πινάκων και δημιουργούν την ανάγκη για επιπλέον πραγματοποίηση πρόσθεσης, εκτός από πολλαπλασιασμό. Το γεγονός αυτό θα γίνει περισσότερο κατανοητό στις επόμενες ενότητες. Επίσης, πολλά προβλήματα, που είναι μη γραμμικά στην περίπτωση χρήσης καρτεσιανών συντεταγμένων, καταλήγουν σε γραμμικά με χρήση των ομογενών συντεταγμένων, γεγονός το οποίο αποτελεί σημαντικό πλεονέκτημά τους, διότι δεν απαιτούνται προσεγγιστικές τιμές και δεν γίνονται επαναληπτικές διαδικασίες. Τέλος, η έκφραση των σημείων του απείρου μέσω αυτών αποτελεί ένα ακόμη πλεονέκτημά τους, σε σχέση με τις καρτεσιανές συντεταγμένες.

3.3 Κεντρική Προβολή

Το σύνηθες μαθηματικό μοντέλο το οποίο περιγράφει τη σχέση μεταξύ των συντεταγμένων χώρου ενός αντικειμένου και της προβολής τους στην εικόνα είναι το μοντέλο κεντρικής

προβολής. Περιγράφει την ιδανική φωτογραφική απεικόνιση και με ορισμένες παραδοχές μπορεί να περιγράψει με ικανοποιητική ακρίβεια το εξεταζόμενο φαινόμενο. Στην πραγματικότητα, η κεντρική προβολή περιγράφει με ακρίβεια μόνο τη μηχανή σημειακής οπής (pinhole camera), με αποτέλεσμα να αναφέρεται συχνά και ως pinhole camera model. Η εικόνα, λοιπόν, η οποία δεν αποτελεί μία αυστηρά κεντρική προβολή, λόγω ποικίλων σφαλμάτων που οφείλονται στους φακούς, στην ατμόσφαιρα, κ.λπ., προσεγγίζεται γεωμετρικά από το μοντέλο αυτό, τόσο στο πλαίσιο της φωτογραμμετρίας, όσο και της όρασης υπολογιστών [113], [114].

Η μηχανή σημειακής οπής περιγράφεται από ένα κουτί, στο μπροστινό μέρος του οποίου υπάρχει ένα άνοιγμα αμελητέων διαστάσεων, και οι ακτίνες που περνούν μέσω αυτού του ανοίγματος σχηματίζουν ένα είδωλο στο πίσω μέρος του κουτιού, όπως φαίνεται στην Εικόνα 3-3. Επειδή μέσω της κεντρικής προβολής δημιουργούνται ανεστραμμένες εικόνες, πολλές φορές καθίσταται βολική η θεώρηση του επιπέδου προβολής μπροστά από την μικροσκοπική οπή, στην ίδια απόσταση που βρισκόταν και προηγουμένως, με αποτέλεσμα η φανταστική αυτή εικόνα να μην είναι ανεστραμμένη [115], [116].



Εικόνα 3-3. Το μοντέλο κεντρικής προβολής (pinhole camera model).

Πηγή: Forsyth & Ponce, 2012

3.3.1 Μαθηματικό μοντέλο που χρησιμοποιείται στη φωτογραμμετρία

Στο πλαίσιο της φωτογραμμετρίας, το μαθηματικό μοντέλο το οποίο χρησιμοποιείται ευρέως και εκφράζει την κεντρική προβολή του χώρου είναι η συνθήκη συγγραμμικότητας, μέσω της οποίας συνδέονται τα σημεία του πραγματικού αντικειμένου με τα αντίστοιχα σημεία της εικόνας, μέσω των στοιχείων του εσωτερικού προσανατολισμού της φωτομηχανής και των παραμέτρων του εξωτερικού προσανατολισμού της λήψης. Η συνθήκη συγγραμμικότητας αποτελεί μία εξίσωση ευθείας στο χώρο, καθώς το κέντρο λήψης, το σημείο του χώρου και η προβολή του στην εικόνα κείνται επί ευθείας. Η συνθήκη συγγραμμικότητας εκφράζεται από τη σχέση (3.4):

$$\begin{bmatrix} x \\ y \\ -c \end{bmatrix} = \lambda \cdot \mathbf{R}_{\omega\varphi\kappa} \cdot \begin{bmatrix} X - X_0 \\ Y - Y_0 \\ Z - Z_0 \end{bmatrix} \qquad (3.4)$$

όπου X, Y, Z οι συντεταγμένες του σημείου στο χώρο, x, y οι διορθωμένες – και όχι οι μετρημένες – εικονοσυντεταγμένες του (στις οποίες εμπεριέχονται τα στοιχεία του εσωτερικού προσανατολισμού), X_0, Y_0, Z_0 οι συντεταγμένες του κέντρου λήψης στο χώρο, c η σταθερά της μηχανής ή πρωτεύουσα απόσταση, που αποτελεί τη διορθωμένη εστιακή απόσταση, λ η κλίμακα εικόνας στο συγκεκριμένο σημείο, και $\mathbf{R_{\omega\varphi\kappa}}$ ορθοκανονικός πίνακας στροφής στο χώρο, που αποτελεί συνάρτηση των γωνιών ω, φ, κ κατά τους άξονες X, Y και Z αντίστοιχα, κατά τις οποίες πρέπει να στραφεί το επίγειο σύστημα συντεταγμένων για να γίνει παράλληλο με το φωτογραμμετρικό. Ο πίνακας στροφής δίνεται από την παρακάτω σχέση:

$$\mathbf{R_{\omega\varphi\kappa}} = \begin{bmatrix} \cos\varphi\cdot\cos\kappa & \sin\omega\cdot\sin\varphi\cdot\cos\kappa+\cos\omega\cdot\sin\kappa & -\cos\omega\cdot\sin\varphi\cdot\cos\kappa+\sin\omega\cdot\sin\kappa \\ -\cos\varphi\cdot\sin\kappa & -\sin\omega\cdot\sin\varphi\cdot\sin\kappa+\cos\omega\cdot\cos\kappa & \cos\omega\cdot\sin\varphi\cdot\sin\kappa+\sin\omega\cdot\cos\kappa \\ \sin\varphi & -\sin\omega\cdot\cos\varphi & \cos\omega\cdot\cos\varphi \end{bmatrix} \quad (3.5)$$

Οι διορθωμένες εικονοσυντεταγμένες του σημείου δίνονται από τις σχέσεις:

$$x = x' - x_0 + \Delta x_r + \Delta x_d \quad (3.6)$$

$$y = y' - y_0 + \Delta y_r + \Delta y_d \quad (3.7)$$

όπου x', y' οι παρατηρούμενες εικονοσυντεταγμένες του σημείου, x_0, y_0 οι εικονοσυντεταγμένες του πρωτεύοντος σημείου, δηλαδή της προβολής του σημείου λήψης στο επίπεδο της εικόνας, Δx_r, Δy_r οι συνιστώσες της συμμετρικής ακτινικής διαστροφής στο συγκεκριμένο εικονοσημείο κατά x και y αντίστοιχα και Δx_d, Δy_d οι διορθώσεις κατά x και y που πρέπει να επιβληθούν στο σημείο λόγω της ασύμμετρης διαστροφής εκκεντρότητας.

Η διαστροφή οφείλεται σε σφάλματα που εισάγουν οι φακοί και προσδιορίζεται μέσω του υπολογισμού των συντελεστών των πολυωνύμων με βάση τα οποία προσεγγίζεται. Συγκεκριμένα, καθώς η φωτεινή ακτίνα διέρχεται μέσα από το σύστημα των φακών δεν ακολουθεί ευθεία γραμμή αλλά αποκλίνει από αυτή, με αποτέλεσμα την παραμόρφωσή της και την επακόλουθη μετατόπιση της θέσης της στο επίπεδο της εικόνας. Η διαστροφή των φακών χωρίζεται σε δύο είδη: τη συμμετρική ακτινική διαστροφή και την ασύμμετρη διαστροφή εκκεντρότητας [115], [117].

Η ακτινική διαστροφή αποτελεί το σημαντικότερο σφάλμα των φακών, εξαιτίας του οποίου οι προβολικές ακτίνες δεν εξέρχονται από το φακό υπό τη γωνία με την οποία προσπίπτουν. Αυτή εισάγει μία ενιαία μεγέθυνση ή σμίκρυνση για κάθε ακτίνα της εικόνας, λόγω της οποίας οι εικόνες των σημείων μετατοπίζονται ακτινικά και συμμετρικά ως προς το πρωτεύον σημείο, οι ευθείες του χώρου δεν απεικονίζονται ως ευθείες, αλλά καμπυλώνονται, και παρατηρείται τοπική διαφοροποίηση της κλίμακας της απεικόνισης. Πρόκειται για συστηματικό σφάλμα και επομένως διορθώνεται με την εφαρμογή του κατάλληλου τύπου [115], [117]. Ευρέως γνωστό πολυώνυμο, το οποίο προσεγγίζει ικανοποιητικά την ακτινική διαστροφή, είναι το πολυώνυμο των περιττών δυνάμεων της

ακτινικής απόστασης από το πρωτεύον σημείο της εικόνας, που δίνεται από την ακόλουθη σχέση:

$$\Delta r = k_0 \cdot r + k_1 \cdot r^3 + k_2 \cdot r^5 + k_3 \cdot r^7 + \ldots \tag{3.8}$$

Εφόσον η σταθερά της μηχανής θεωρείται άγνωστη και υπολογίζεται ταυτόχρονα με τα υπόλοιπα στοιχεία του εσωτερικού προσανατολισμού, ο γραμμικός όρος του παραπάνω πολυωνύμου απορροφάται από τη σταθερά της μηχανής, λόγω της μεγάλης συσχέτισης του συντελεστή k_0 με αυτή. Έτσι, στις περιπτώσεις αυτές λαμβάνεται $k_0=0$. Αντίθετα, αν η σταθερά θεωρείται εκ των προτέρων γνωστή, ο συντελεστής αυτός δεν θεωρείται μηδενικός.

Οι κατά x και y συνιστώσες της διαστροφής σε κάθε σημείο δίνονται από τις σχέσεις:

$$\Delta x_r = \frac{x' - x_0}{r} \cdot \Delta r \tag{3.9}$$

$$\Delta y_r = \frac{y' - y_0}{r} \cdot \Delta r \tag{3.10}$$

Συνεπώς, με βάση τις σχέσεις (3.8), (3.9) και (3.10), θεωρώντας ότι $k_0=0$ και λαμβάνοντας υπ' όψιν τους συντελεστές k_1, k_2 και k_3, που θεωρούνται επαρκείς για την περιγραφή της ακτινικής διαστροφής, οι διορθώσεις κατά x και y λόγω της τελευταίας εκφράζονται ως εξής:

$$\Delta x_r = \frac{x' - x_0}{r} \cdot \left(k_1 \cdot r^3 + k_2 \cdot r^5 + k_3 \cdot r^7\right) \rightarrow \Delta x_r = \left(x' - x_0\right) \cdot \left(k_1 \cdot r^2 + k_2 \cdot r^4 + k_3 \cdot r^6\right) \tag{3.11}$$

$$\Delta y_r = \frac{y' - y_0}{r} \cdot \left(k_1 \cdot r^3 + k_2 \cdot r^5 + k_3 \cdot r^7\right) \rightarrow \Delta y_r = \left(y' - y_0\right) \cdot \left(k_1 \cdot r^2 + k_2 \cdot r^4 + k_3 \cdot r^6\right) \tag{3.12}$$

Στις παραπάνω σχέσεις, r είναι η ακτινική απόσταση του μετρημένου σημείου από το πρωτεύον σημείο της εικόνας, η οποία δίνεται από την ακόλουθη σχέση:

$$r = \sqrt{\left(x' - x_0\right)^2 + \left(y' - y_0\right)^2} \tag{3.13}$$

Η ασύμμετρη διαστροφή εκκεντρότητας, η οποία συχνά αναφέρεται και ως εφαπτομενική διαστροφή, προέρχεται από την αδυναμία ακριβούς κέντρωσης των φακών μέσα στο σύστημα των φακών. Διακρίνεται σε δύο συνιστώσες, την ακτινική και την εγκάρσια, η οποία είναι μικρότερου μεγέθους. Η επίδραση του σφάλματος της διαστροφής εκκεντρότητας είναι σημαντικά μικρότερη από την αντίστοιχη της ακτινικής διαστροφής, με αποτέλεσμα πολλές φορές να θεωρείται αμελητέα και να παραλείπεται.

Οι διορθώσεις τις οποίες πρέπει να υποστούν οι εικονοσυντεταγμένες λόγω της ασύμμετρης διαστροφής εκκεντρότητας μπορούν να εκφραστούν σε πολυωνυμική μορφή από τις παρακάτω σχέσεις [118], [119]:

$$\Delta x_d = \left[2 \cdot p_1 \cdot \left(x' - x_0\right) \cdot \left(y' - y_0\right) + p_2 \cdot \left(r^2 + 2 \cdot \left(x' - x_0\right)^2\right)\right] \cdot \left[1 + p_3 \cdot r^2 + p_4 \cdot r^4 + \ldots\right] \tag{3.14}$$

$$\Delta y_d = \left[p_1 \cdot \left(r^2 + 2 \cdot \left(y' - y_0 \right)^2 \right) + 2 \cdot p_2 \cdot \left(x' - x_0 \right) \cdot \left(y' - y_0 \right) \right] \cdot \left[1 + p_3 \cdot r^2 + p_4 \cdot r^4 + \ldots \right] \quad (3.15)$$

όπου p_1, p_2, p_3 και p_4 είναι οι συντελεστές των πολυωνύμων. Ωστόσο, συνήθως οι συντελεστές p_3 και p_4 αγνοούνται, με αποτέλεσμα οι διορθώσεις να λαμβάνουν την παρακάτω μορφή [120]:

$$\Delta x_d = 2 \cdot p_1 \cdot \left(x' - x_0 \right) \cdot \left(y' - y_0 \right) + p_2 \cdot \left(r^2 + 2 \cdot \left(x' - x_0 \right)^2 \right) \quad (3.16)$$

$$\Delta y_d = p_1 \cdot \left(r^2 + 2 \cdot \left(y' - y_0 \right)^2 \right) + 2 \cdot p_2 \cdot \left(x' - x_0 \right) \cdot \left(y' - y_0 \right) \quad (3.17)$$

Η αναλυτική μορφή της συνθήκης συγγραμμικότητας, λαμβάνοντας υπ'όψιν και τη διαστροφή των φακών, είναι η εξής:

$$x' = x_0 - \Delta x_r - \Delta x_d - c \cdot \frac{(X - X_0) \cdot r_{11} + (Y - Y_0) \cdot r_{12} + (Z - Z_0) \cdot r_{13}}{(X - X_0) \cdot r_{31} + (Y - Y_0) \cdot r_{32} + (Z - Z_0) \cdot r_{33}} \quad (3.18)$$

$$y' = y_0 - \Delta y_r - \Delta y_d - c \cdot \frac{(X - X_0) \cdot r_{21} + (Y - Y_0) \cdot r_{22} + (Z - Z_0) \cdot r_{23}}{(X - X_0) \cdot r_{31} + (Y - Y_0) \cdot r_{32} + (Z - Z_0) \cdot r_{33}} \quad (3.19)$$

όπου r_{ij} τα στοιχεία του πίνακα στροφής. Τα στοιχεία του εξωτερικού προσανατολισμού είναι οι συντεταγμένες του προβολικού κέντρου X_0, Y_0, Z_0 και οι γωνίες στροφής ω, φ, κ. Τα στοιχεία c, x_0 και y_0 θεωρούνται τα κύρια στοιχεία του εσωτερικού προσανατολισμού της φωτογραφικής μηχανής, τα οποία μαζί με τις 6 παραμέτρους του εξωτερικού προσανατολισμού ορίζουν πλήρως την κεντρική προβολή. Στις παραμέτρους αυτές προστίθενται συνήθως και οι συντελεστές του πολυωνύμου ακτινικής διαστροφής και, σπανιότερα, οι συντελεστές του πολυωνύμου της ασύμμετρης διαστροφής εκκεντρότητας.

Οι παραπάνω εξισώσεις, όπως γίνεται κατανοητό, δεν είναι γραμμικές ως προς τα στοιχεία του εσωτερικού και εξωτερικού προσανατολισμού και ως προς τις συντεταγμένες χώρου του σημείου, με αποτέλεσμα να απαιτούνται προσεγγιστικές τιμές για αυτές προκειμένου να επιλυθούν σημαντικά φωτογραμμετρικά προβλήματα, όπως η βαθμονόμηση, η οπισθοτομία και η εμπροσθοτομία.

3.3.2 Μαθηματικό μοντέλο που χρησιμοποιείται στην όραση υπολογιστών

Στο πλαίσιο της επιστήμης της όρασης υπολογιστών, η κεντρική προβολή του τρισδιάστατου προβολικού χώρου στο επίπεδο της εικόνας εκφράζεται μέσω ομογενών συντεταγμένων από την παρακάτω εξίσωση:

$$\mathbf{x} = \mathbf{P} \cdot \mathbf{X} \quad \rightarrow \quad \begin{bmatrix} x \\ y \\ w \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \\ W \end{bmatrix} \tag{3.20}$$

η οποία συνήθως λαμβάνει τη μορφή:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{3.21}$$

όπου $\mathbf{X} = \begin{bmatrix} X & Y & Z & W \end{bmatrix}^T$ είναι το διάνυσμα ομογενών συντεταγμένων του σημείου στο χώρο, $\mathbf{x} = \begin{bmatrix} x & y & w \end{bmatrix}^T$ είναι το διάνυσμα ομογενών συντεταγμένων της προβολής του στο επίπεδο της εικόνας και **P** είναι ένας 3x4 ομογενής πίνακας, ο οποίος ονομάζεται προβολικός πίνακας της μηχανής (projective camera matrix).

Αν η εξίσωση (3.21) εκφραστεί μέσω μη ομογενών συντεταγμένων, τότε καταλήγει στην εξίσωση του άμεσου γραμμικού μετασχηματισμού (DLT - Direct Linear Transformation) [94], [107], ο οποίος ορίζεται ως εξής:

$$x = \frac{p_{11} \cdot X + p_{12} \cdot Y + p_{13} \cdot Z + p_{14}}{p_{31} \cdot X + p_{32} \cdot Y + p_{33} \cdot Z + p_{34}} \tag{3.22}$$

$$y = \frac{p_{21} \cdot X + p_{22} \cdot Y + p_{23} \cdot Z + p_{24}}{p_{31} \cdot X + p_{32} \cdot Y + p_{33} \cdot Z + p_{34}} \tag{3.23}$$

και προκύπτει με διαίρεση της πρώτης και δεύτερης με την τρίτη συντεταγμένη του διανύσματος ομογενών συντεταγμένων του δισδιάστατου προβολικού χώρου. Στην εξίσωση αυτή συνήθως τίθεται $p_{34} = 1$.

Ο πίνακας **P** μπορεί να εκφραστεί ως πολλαπλασιασμός τριών πινάκων, όπως υποδεικνύεται στη συνέχεια.

Με την παραδοχή ότι το επίγειο σύστημα συντεταγμένων συμπίπτει με το φωτογραμμετρικό[9] και τη θεώρηση ότι η αρχή του συστήματος συντεταγμένων της εικόνας ταυτίζεται με το πρωτεύον σημείο, όπως φαίνεται στην Εικόνα 3-4, είναι δυνατή η εύρεση της προβολής ενός σημείου του χώρου με συντεταγμένες X, Y, Z στο επίπεδο της εικόνας. Με τη βοήθεια όμοιων τριγώνων αποδεικνύεται ότι:

$$\frac{c}{Z} = \frac{x_c}{X} \quad \rightarrow \quad x_c = \frac{c}{Z} \cdot X \tag{3.24}$$

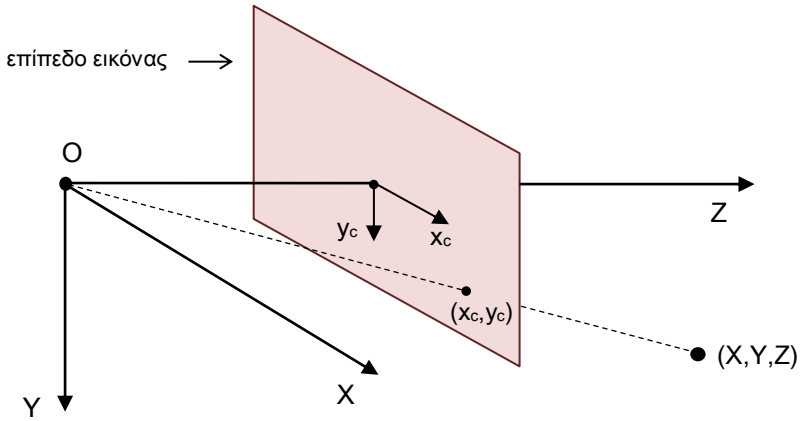
[9] Ο ορισμός του φωτογραμμετρικού συστήματος στις Εικόνες 3-4 και 3-6 γίνεται με βάση το αριστερόστροφο φωτογραμμετρικό σύστημα, όπως ορίζεται στη βιβλιοθήκη OpenCV που χρησιμοποιήθηκε στις εφαρμογές.

$$\frac{c}{Z}=\frac{y_c}{Y} \quad \rightarrow \quad y_c=\frac{c}{Z}\cdot Y \tag{3.25}$$

όπου c η πρωτεύουσα απόσταση και x_c, y_c οι εικονοσυντεταγμένες του σημείου του χώρου με επίγειες συντεταγμένες X, Y, Z [121], [122], [123]. Οι παραπάνω σχέσεις μπορούν να εκφραστούν μέσω ομογενών συντεταγμένων, και έτσι προκύπτει η ακόλουθη σχέση, που συνδέει τις συντεταγμένες του σημείου στο χώρο με τις συντεταγμένες του στην πραγματικότητα, υπό την παραπάνω παραδοχή [123]:

$$\begin{bmatrix} x_c \\ y_c \\ c \end{bmatrix}=\frac{c}{Z}\cdot\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}\cdot\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad \rightarrow \quad \begin{bmatrix} x_c \\ y_c \\ c \end{bmatrix}\sim\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}\cdot\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{3.26}$$

Να σημειωθεί ότι το σύμβολο $\sim$ δηλώνει ισότητα που περιλαμβάνει πολλαπλασιασμό με μη μηδενικό συντελεστή κλίμακας.



Εικόνα 3-4. Απλοποιημένη γεωμετρία κεντρικής προβολής. Ως O συμβολίζεται το προβολικό κέντρο, ενώ X, Y, Z είναι οι άξονες του επίγειου συστήματος αναφοράς, που θεωρείται ότι ταυτίζεται με το φωτογραμμετρικό.

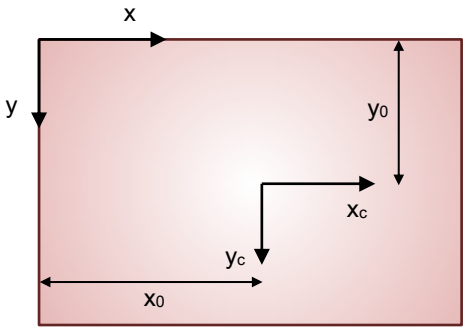
Ωστόσο, η παραδοχή που θεωρήθηκε παραπάνω, ότι η αρχή του συστήματος συντεταγμένων της εικόνας συμπίπτει με το πρωτεύον σημείο, γενικά δεν ισχύει. Θεωρώντας ότι η αρχή του συστήματος συντεταγμένων της εικόνας βρίσκεται στο επάνω αριστερά pixel – κάτι το οποίο υιοθετείται και από τη βιβλιοθήκη OpenCV που χρησιμοποιήθηκε για την ανάπτυξη των εφαρμογών – στις σχέσεις (3.24) και (3.25) πρέπει να προστεθούν και οι συντεταγμένες του πρωτεύοντος σημείου x_0 και y_0 (βλ. Εικόνα 3-5). Προς το παρόν, όλα τα μεγέθη είναι εκφρασμένα σε μονάδες μήκους και όχι σε pixels. Για τη μετατροπή του αποτελέσματος σε pixels πρέπει να γίνει διαίρεση με το μέγεθος του pixel. Επομένως, οι σχέσεις (3.24) και (3.25), με βάση τα παραπάνω αναφερόμενα, και θεωρώντας τη γενική περίπτωση μη τετραγωνικών εικονοστοιχείων, γίνονται:

$$x=\frac{c}{Z\cdot \text{width}_{pix}}\cdot X+x_0 \tag{3.27}$$

$$y = \frac{c}{Z \cdot height_{pix}} \cdot Y + y_0 \tag{3.28}$$

Στις παραπάνω σχέσεις, $c / width_{pix}$ είναι η σταθερά της μηχανής εκφρασμένη σε pixels μεγέθους ίσου με το μήκος της οριζόντιας πλευράς τους, η οποία συμβολίζεται συχνά ως c_x, και $c / height_{pix}$ είναι η σταθερά της μηχανής εκφρασμένη σε pixels μεγέθους ίσου με το μήκος της κατακόρυφης πλευράς τους και συμβολίζεται ως c_y [124]. Επίσης, στις παραπάνω σχέσεις οι συντεταγμένες του πρωτεύοντος σημείου x_0 και y_0 εκφράζονται ομοίως σε pixels, ώστε και το τελικό αποτέλεσμα να προκύψει σε pixels. Έτσι, για να καλυφθεί η περίπτωση των μη τετραγωνικών εικονοστοιχείων, εισάγεται ο συντελεστής a (aspect ratio), ο οποίος αποτελεί το λόγο των πλευρών του pixel:

$$a = \frac{c_y}{c_x} = \frac{width_{pix}}{height_{pix}} \tag{3.29}$$



Εικόνα 3-5. Σύστημα εικονοσυντεταγμένων και θεωρούμενο σύστημα αναφοράς των μετρήσεων της εικόνας.

Επίσης, πολλές φορές θεωρείται ένας ακόμη προβολικός συντελεστής, ο συντελεστής s (skewness), για να καλυφθεί η γενικότερη περίπτωση της απόκλισης από την ορθογωνικότητα των αξόνων του pixel. Με βάση τα παραπάνω, η σχέση (3.26) λαμβάνει την ακόλουθη μορφή:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} c_x & s & x_0 \\ 0 & a \cdot c_x & y_0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{3.30}$$

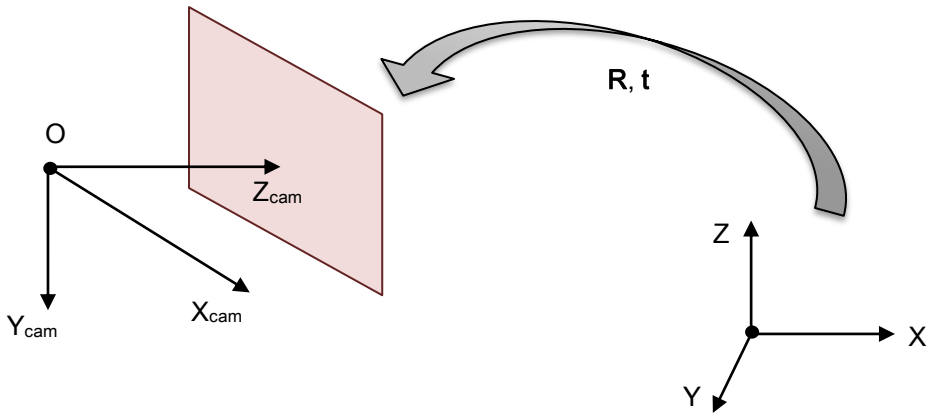
Ο 3x3 άνω τριγωνικός πίνακας ονομάζεται και πίνακας εσωτερικού προσανατολισμού της φωτομηχανής (camera calibration matrix) ή και πίνακας της μηχανής (camera matrix). Συνήθως, ο πίνακας αυτός, έστω **K**, συναντάται με την εξής ισοδύναμη μορφή:

$$\mathbf{K} = \begin{bmatrix} c_x & s & x_0 \\ 0 & c_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3.31}$$

όπου $c_y = a \cdot c_x$ [94], [121], [122], [123].

Σε πολλές εφαρμογές βέβαια, υιοθετείται το μοντέλο της απλής κεντρικής προβολής, με συνέπεια να θεωρείται $s = 0$ και $c_x = c_y = c$. Κατά τη χρήση των ψηφιακών μηχανών με αισθητήρα CCD, είναι πιθανό τα pixels να μην μπορούν να θεωρηθούν τετράγωνα, με αποτέλεσμα να μην αγνοείται το aspect ratio και επομένως να θεωρούνται διαφορετικές τιμές για τη σταθερά c. Η περίπτωση της μη καθετότητας των πλευρών των εικονοστοιχείων ($s \neq 0$) είναι σπανιότερη [107].

Έως το σημείο αυτό, θεωρήθηκε ότι το φωτογραμμετρικό σύστημα συντεταγμένων και το επίγειο συμπίπτουν. Στην πραγματικότητα ωστόσο, αυτά δεν ταυτίζονται, με αποτέλεσμα το προβολικό κέντρο να μη βρίσκεται στη θέση (0,0,0) και επιπλέον ο άξονας Z να μην είναι κάθετος στην εικόνα. Συνεπώς, το επίγειο σύστημα συντεταγμένων πρέπει να στραφεί και να μετατεθεί, ώστε να συμπέσει με το φωτογραμμετρικό, όπως φαίνεται και στην Εικόνα 3-6 [121], [122], [123].



Εικόνα 3.6. Μετάθεση και στροφή του επίγειου συστήματος συντεταγμένων για να ταυτιστεί με το φωτογραμμετρικό.

Έτσι, αν συμβολιστούν ως X_{cam}, Y_{cam}, Z_{cam} οι συντεταγμένες του προβολικού κέντρου, τότε ισχύει η ακόλουθη εξίσωση [122], [123]:

$$\begin{bmatrix} X_{cam} \\ Y_{cam} \\ Z_{cam} \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ 0^T & 1 \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{3.32}$$

όπου **R** ο 3x3 πίνακας στροφής και **t** το τριών διαστάσεων διάνυσμα μετατόπισης.

Σύμφωνα με τα παραπάνω, λοιπόν, οι εικονοσυντεταγμένες ενός σημείου του χώρου προκύπτουν από την ακόλουθη σχέση:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \mathbf{K} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ 0^T & 1 \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{3.33}$$

η οποία γράφεται ως εξής:

$$\mathbf{x} = \mathbf{K} \cdot \mathbf{I} \cdot \left[\mathbf{R} \,|\, \mathbf{t}\right] \cdot \mathbf{X} \quad \rightarrow \quad \mathbf{x} = \mathbf{K} \cdot \left[\mathbf{R} \,|\, \mathbf{t}\right] \cdot \mathbf{X} \tag{3.34}$$

Επομένως, ο προβολικός πίνακας της μηχανής **P** δίνεται από την ακόλουθη σχέση [122]:

$$\mathbf{P} = \mathbf{K} \cdot \left[\mathbf{R} \,|\, \mathbf{t}\right] = \mathbf{K} \cdot \mathbf{R} \cdot \left[\mathbf{I} \,|\, \mathbf{R}^{\mathrm{T}} \cdot \mathbf{t}\right] \tag{3.35}$$

Επίσης, να σημειωθεί ότι το 3x1 διάνυσμα $\mathbf{R}^{\mathrm{T}} \cdot \mathbf{t}$ είναι το διάνυσμα με την αντίθετη θέση του κέντρου λήψης, δηλαδή:

$$\mathbf{R}^{\mathrm{T}} \cdot \mathbf{t} = -\mathbf{C} = -\begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \end{bmatrix} \tag{3.36}$$

Συνεπώς, ο πίνακας **P** γράφεται ως εξής:

$$\mathbf{P} = \mathbf{K} \cdot \mathbf{R} \cdot \left[\mathbf{I} \,|\, -\mathbf{C}\right] \tag{3.37}$$

Τελικά, με βάση τις σχέσεις (3.21) και (3.35) το σύστημα εξισώσεων υπό τη μορφή πινάκων είναι το ακόλουθο:

$$\lambda \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} c_x & s & x_0 \\ 0 & c_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{3.38}$$

όπου r_{ij} τα στοιχεία του πίνακα στροφής, t_i τα στοιχεία του διανύσματος μετάθεσης και λ ένας συντελεστής κλίμακας. Ο 3x4 πίνακας μετάθεσης - στροφής $\left[\mathbf{R} \,|\, \mathbf{t}\right]$ ονομάζεται και πίνακας εξωτερικού προσανατολισμού της λήψης. Συνήθως, στην εξίσωση (3.38) τίθεται $s = 0$, καθώς στην πλειονότητα των περιπτώσεων δεν παρατηρείται μη καθετότητα των πλευρών των εικονοστοιχείων. Στην εξίσωση αυτή, οι εικονοσυντεταγμένες x, y του σημείου με συντεταγμένες χώρου X, Y, Z, οι δύο τιμές για τη σταθερά της μηχανής c_x, c_y καθώς και οι συντεταγμένες του πρωτεύοντος σημείου x_0, y_0 εκφράζονται σε pixel. Ο παραπάνω μετασχηματισμός είναι ισοδύναμος με τον επόμενο, για $w \neq 0$, στην περίπτωση κατά την οποία s=0 (για απόδειξη βλ. Παράρτημα Β):

$$\begin{bmatrix} x' \\ y' \\ w' \end{bmatrix} = \mathbf{R} \cdot \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \mathbf{t} \quad \text{όπου} \quad \left\{ \begin{array}{l} x = c_x \cdot x'' + x_0 \\ y = c_y \cdot y'' + y_0 \\ \text{για } x'' = x'/w' \\ \text{και } y'' = y'/w' \end{array} \right\} \rightarrow \left\{ \begin{array}{l} x' = \dfrac{(x - x_0) \cdot w'}{c_x} \\ y' = \dfrac{(y - y_0) \cdot w'}{c_y} \end{array} \right\} \tag{3.39}$$

Το μοντέλο αυτό μπορεί να επεκταθεί για να συμπεριλάβει και τα σφάλματα των φακών, και συγκεκριμένα την ακτινική διαστροφή και την ασύμμετρη διαστροφή εκκεντρότητας. Συνεπώς, το παραπάνω μοντέλο, μπορεί να γραφεί ως εξής [125]:

$$\begin{bmatrix} x' \\ y' \\ w' \end{bmatrix} = \mathbf{R} \cdot \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \mathbf{t} \quad \text{όπου}$$

$$\left\{ \begin{array}{c} x = c_x \cdot x''' + x_0 \\ y = c_y \cdot y''' + y_0 \\ \text{για } x''' = x'' \cdot \left(1 + k_1 \cdot r^2 + k_2 \cdot r^4 + k_3 \cdot r^6\right) + 2 \cdot p_1 \cdot x'' \cdot y'' + p_2 \cdot \left(r^2 + 2 \cdot x''^2\right) \\ \text{και } y''' = y'' \cdot \left(1 + k_1 \cdot r^2 + k_2 \cdot r^4 + k_3 \cdot r^6\right) + p_1 \cdot \left(r^2 + 2 \cdot y''^2\right) + 2 \cdot p_2 \cdot x'' \cdot y'' \\ \text{όπου } x'' = x'/w, \quad y'' = y'/w \quad \text{και} \quad r = \sqrt{x''^2 + y''^2} \end{array} \right\} \tag{3.40}$$

3.3.3 Σύγκριση συνθήκης συγγραμμικότητας και προβολικού μετασχηματισμού

Σύμφωνα με όσα αναφέρθηκαν στην υποενότητα 3.3.2, με βάση την προσέγγιση που ακολουθείται από την επιστήμη της όρασης υπολογιστών, η προβολική σχέση εικόνας - τρισδιάστατου χώρου εκφράζεται μέσω 11 παραμέτρων (3 γωνίες στροφής κατά τους άξονες X, Y, Z, 3 μεταθέσεις κατά τους άξονες αυτούς και 5 στοιχεία της εσωτερικής γεωμετρίας της μηχανής), αν δεν ληφθεί υπ' όψιν η διαστροφή των φακών. Ωστόσο, το γεωμετρικό πρόβλημα της κεντρικής προβολής έχει 9 παραμέτρους (3 βασικές παραμέτρους του εσωτερικού προσανατολισμού και 6 του εξωτερικού προσανατολισμού), με αποτέλεσμα ο προβολικός αυτός μετασχηματισμός να μην περιγράφει αυστηρά μία κεντρική προβολή, αλλά να είναι γενικότερος αυτής. Η διαφορά των δύο βαθμών ελευθερίας αίρεται εφόσον μηδενιστούν η απόκλιση από την ορθογωνικότητα (s=0) και η διαφορά κλίμακας μεταξύ των πλευρών των εικονοστοιχείων (a=1 ή αντίστοιχα c_x=c_y=c).

Γενικά, η κύρια διαφορά της συνθήκης συγγραμμικότητας και του προβολικού μετασχηματισμού, που χρησιμοποιείται στην επιστήμη της όρασης υπολογιστών, έγκειται στη διάκριση της προοπτικότητας, που είναι προϊόν της κεντρικής προβολής, από την προβολικότητα, η οποία είναι πιο γενική. Έτσι, στο πλαίσιο της εφαρμογής του προβολικού μετασχηματισμού, οι συντεταγμένες των εικονοσημείων δεν απαιτείται να μετρηθούν στο σύστημα εικονοσυντεταγμένων ή να αναχθούν σε αυτό, μέσω ενός αφινικού μετασχηματισμού (βλ. παράγραφο 3.4.1.4). Ο τελευταίος ενσωματώνεται στους 11 συντελεστές και έτσι σε αντίθεση με τη συνθήκη συγγραμμικότητας, στην οποία οι συντεταγμένες των σημείων της εικόνας πρέπει να αναφέρονται στο σύστημα εικονοσυντεταγμένων με αρχή το πρωτεύον σημείο, στο πλαίσιο του προβολικού μετασχηματισμού μπορούν να χρησιμοποιηθούν απευθείας οι συντεταγμένες pixel [126].

Σημαντικό πλεονέκτημα της εξίσωσης του προβολικού μετασχηματισμού, σε σχέση με την εξίσωση συγγραμμικότητας, είναι το γεγονός ότι ο υπολογισμός των 11 παραμέτρων είναι γραμμικός, με αποτέλεσμα να μην υπάρχει ανάγκη για προσεγγιστικές τιμές. Συνεπώς, η

επίλυση καθίσταται πολύ ευκολότερη αλλά και γρηγορότερη, καθώς δεν απαιτείται να πραγματοποιηθούν επαναλήψεις για την εύρεση της τελικής λύσης.

Επίσης, μία άλλη διαφορά σε σχέση με τη συνθήκη συγγραμμικότητας προκύπτει από το γεγονός ότι ο 3x4 προβολικός πίνακας της μηχανής **P** αποτελείται από παραμέτρους οι οποίες δεν έχουν άμεσα κάποιο φυσικό νόημα, όπως ο πίνακας στροφής, οι συντεταγμένες του σημείου λήψης και τα στοιχεία του εσωτερικού προσανατολισμού στο πλαίσιο της συνθήκης συγγραμμικότητας. [127].

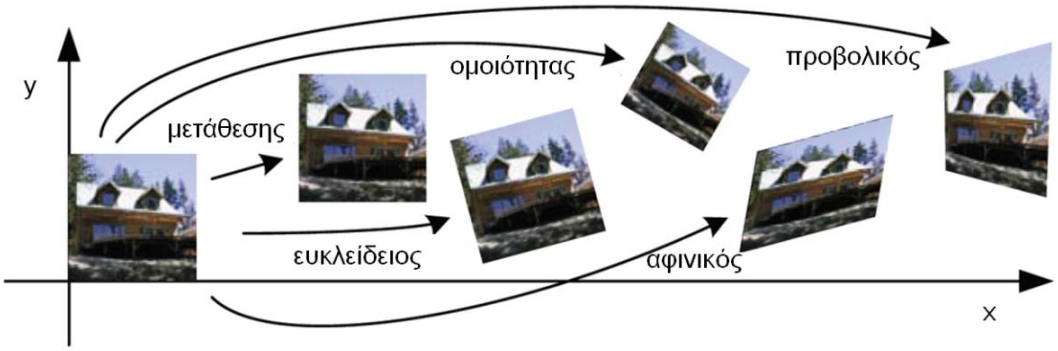
Ως μειονεκτήματα του προβολικού μετασχηματισμού μπορεί να θεωρηθούν το γεγονός ότι δεν μοντελοποιεί σωστά τη φυσική διαδικασία σχηματισμού της εικόνας, με αποτέλεσμα να απαιτούνται περισσότερες ομολογίες σημείων, καθώς επίσης και η μεγαλύτερη ευαισθησία του σε σφάλματα των μετρήσεων σε σχέση με τη συνθήκη συγγραμμικότητας, που στηρίζεται στην αληθή προοπτική προβολή. Η τελευταία παράγει περισσότερο ακριβή αποτελέσματα [128].

3.4 Δισδιάστατοι Γραμμικοί Μετασχηματισμοί

Στην ενότητα αυτή γίνεται μία παρουσίαση των βασικών δισδιάστατων γραμμικών μετασχηματισμών, με έμφαση στον προβολικό μετασχηματισμό μεταξύ δύο δισδιάστατων συστημάτων συντεταγμένων. Αυτός συναντάται συχνά με τον όρο ομογραφία και αποτελεί το γενικότερο δισδιάστατο γραμμικό μετασχηματισμό, υπό την έννοια ότι όλοι οι υπόλοιποι μπορούν να θεωρηθούν υποπεριπτώσεις του. Ο μετασχηματισμός αυτός χρησιμοποιήθηκε τόσο κατά την ανάπτυξη των εφαρμογών, όσο και στη βαθμονόμηση, η οποία έγινε για τις ανάγκες της παρούσας εργασίας.

3.4.1 Ιεραρχία των δισδιάστατων γραμμικών μετασχηματισμών

Στην παράγραφο 2.3.3.3 έγινε λόγος για απαλοιφή των μη έγκυρων ομολογιών, που προκύπτουν από τη διαδικασία της αυτόματης ψηφιακής συνταύτισης, με την εφαρμογή ενός γεωμετρικού μετασχηματισμού μεταξύ των εικόνων που συνταυτίστηκαν. Στην παρούσα υποενότητα περιγράφονται κάποιοι βασικοί δισδιάστατοι γραμμικοί μετασχηματισμοί, ένας εκ των οποίων πρέπει να επιλεγεί για να εκφράσει τη γεωμετρική σχέση μεταξύ των εικόνων ενός δισδιάστατου αντικειμένου. Προτού όμως γίνει σύντομη αναφορά σε κάθε έναν από αυτούς με έμφαση στο γενικότερο όλων, την ομογραφία, κρίνεται σκόπιμη η παρουσίαση της ιεραρχίας των μετασχηματισμών, όπως φαίνεται στην Εικόνα 3-7.



Εικόνα 3-7. Αλληλουχία των γραμμικών δισδιάστατων μετασχηματισμών.

Πηγή: Szeliski, 2011

Η Εικόνα 3-7 παρουσιάζει τους γραμμικούς γεωμετρικούς δισδιάστατους μετασχηματισμούς που μπορούν, υπό κάποιες προϋποθέσεις, να εκφράσουν τη σχέση μεταξύ επίπεδων αντικειμένων που απεικονίζονται στις εικόνες, με τον πιο γενικό μετασχηματισμό να φαίνεται στα δεξιά της εικόνας. Ανάλογα με την εφαρμογή και τις υποθέσεις που γίνονται, μπορεί να γίνει η επιλογή του κατάλληλου, κάθε φορά, μετασχηματισμού. Πιο απλός είναι ο μετασχηματισμός μετάθεσης και αν σε αυτόν προστεθεί και μία στροφή προκύπτει ο ευκλείδειος μετασχηματισμός. Ο μετασχηματισμός ομοιότητας προκύπτει από τον προηγούμενο με μία επιπλέον ενιαία μεταβολή κλίμακας, ενώ ο αφινικός μετασχηματισμός εισάγει τη μη ομοιόμορφη μεταβολή κλίμακας και τη μη καθετότητα των αξόνων του ενός δισδιάστατου συστήματος συντεταγμένων. Τέλος, ο προβολικός μετασχηματισμός ή ομογραφία είναι ο γενικότερος όλων και μπορεί να εφαρμοστεί σε όλες τις περιπτώσεις μετασχηματισμού μεταξύ εικόνων του ίδιου δισδιάστατου αντικειμένου.

Μετασχηματισμός	Βαθμοί ελευθερίας	Μεγέθη που διατηρούνται				
		Προσανατολισμός	Μήκη	Γωνίες	Παράλληλες γραμμές	Ευθείες Γραμμές
Μετάθεσης	**2**	✓	✓	✓	✓	✓
Ευκλείδειος	**3**		✓	✓	✓	✓
Ομοιότητας	**4**			✓	✓	✓
Αφινικός	**6**				✓	✓
Προβολικός	**8**					✓

Πίνακας 3-1. Ιεραρχία των γραμμικών δισδιάστατων μετασχηματισμών.

Στον Πίνακα 3-1 φαίνονται τα μεγέθη που διατηρούνται σε κάθε μετασχηματισμό, καθώς επίσης και οι βαθμοί ελευθερίας του καθενός. Όπως προκύπτει και από τον πίνακα αυτό, οι

μετασχηματισμοί αυτοί ονομάζονται γραμμικοί επειδή οι ευθείες γραμμές μετασχηματίζονται ομοίως σε ευθείες [126]. Ωστόσο, πρέπει να σημειωθεί ότι αυτοί δεν είναι οι μοναδικοί μετασχηματισμοί και υπάρχει πληθώρα άλλων, όπως μετασχηματισμός στροφής, κλίμακας, ορθογωνικός κ.ά., οι οποίοι δεν εξετάζονται στο πλαίσιο της παρούσας εργασίας, καθώς μπορεί να προκύψουν από τη θεώρηση ως μηδενικών κάποιων εκ των παραμέτρων ενός γενικότερου μετασχηματισμού. Για παράδειγμα, ο ορθογωνικός προκύπτει από τον αφινικό θεωρώντας μηδενική την απόκλιση από την καθετότητα, ο μετασχηματισμός στροφής προκύπτει από τον ευκλείδειο μετασχηματισμό θεωρώντας μηδενικό το διάνυσμα μετατόπισης, κ.ο.κ.

3.4.1.1 Μετασχηματισμός μετάθεσης

Ο μετασχηματισμός μετάθεσης εκφράζεται μέσω δύο παραμέτρων, των μεταθέσεων t_x και t_y κατά τις διευθύνσεις x και y αντίστοιχα μεταξύ των δύο δισδιάστατων συστημάτων συντεταγμένων. Αν συμβολιστούν ως $\begin{bmatrix} x_1 & y_1 \end{bmatrix}^T$ και $\begin{bmatrix} x_2 & y_2 \end{bmatrix}^T$ τα διανύσματα που εκφράζουν τις καρτεσιανές συντεταγμένες του ίδιου σημείου στα δύο επίπεδα συστήματα συντεταγμένων που διαφέρουν μόνο κατά μία μετάθεση, η σχέση που τα συνδέει είναι η εξής:

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \tag{3.41}$$

όπου t_x και t_y οι μεταθέσεις κατά τις διευθύνσεις x και y αντίστοιχα. Ο ίδιος μετασχηματισμός μπορεί να εκφραστεί μέσω των ομογενών διανυσμάτων $\begin{bmatrix} x_1 & y_1 & w_1 \end{bmatrix}^T$ και $\begin{bmatrix} x_2 & y_2 & w_2 \end{bmatrix}^T$ ως εξής:

$$\begin{bmatrix} x_2 \\ y_2 \\ w_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ y_1 \\ w_1 \end{bmatrix} \tag{3.42}$$

Βέβαια, στην περίπτωση που διατίθενται οι καρτεσιανές συντεταγμένες και επιδιώκεται η έκφραση του ανωτέρου μετασχηματισμού με ομογενείς συντεταγμένες, τίθεται $w_1 = w_2 = 1$. Αυτό ισχύει και για τους μετασχηματισμούς που θα περιγραφούν στη συνέχεια και αποτελεί συνήθη τεχνική στη φωτογραμμετρία και στην όραση υπολογιστών.

3.4.1.2 Ευκλείδειος μετασχηματισμός

Ο μετασχηματισμός στροφής και μετάθεσης είναι ευρέως γνωστός και ως ευκλείδειος μετασχηματισμός, καθώς διατηρεί τις ευκλείδειες αποστάσεις, δηλαδή τα μήκη μεταξύ των σημείων του επιπέδου. Ο μετασχηματισμός, στην περίπτωση χρήσης καρτεσιανών συντεταγμένων, εκφράζεται μέσω ενός πολλαπλασιασμού και μίας πρόσθεσης, ως εξής:

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \tag{3.43}$$

Αν ως θ εκφραστεί η αριστερόστροφη γωνία κατά την οποία πρέπει να στραφεί το σύστημα συντεταγμένων του δεύτερου επιπέδου για να γίνει παράλληλο με το πρώτο, ο πίνακας στροφής του δεύτερου συστήματος είναι ο ακόλουθος [126]:

$$\mathbf{R} = \begin{bmatrix} a & b \\ -b & a \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \tag{3.44}$$

Επομένως η σχέση (3.43) μπορεί να γραφεί ως εξής:

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \tag{3.45}$$

ενώ υφίσταται και η δέσμευση των παραμέτρων:

$$a^2 + b^2 = \cos^2\theta + \sin^2\theta = 1 \tag{3.46}$$

Στην περίπτωση που χρησιμοποιούνται ομογενείς συντεταγμένες, ο ίδιος μετασχηματισμός εκφράζεται μέσω ενός μόνο πολλαπλασιασμού πινάκων, ως εξής:

$$\begin{bmatrix} x_2 \\ y_2 \\ w_2 \end{bmatrix} = \begin{bmatrix} a & b & t_x \\ -b & a & t_y \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ y_1 \\ w_1 \end{bmatrix} \quad \text{υπό τη δέσμευση: } a^2 + b^2 = \lambda^2 \tag{3.47}$$

Δεδομένου ότι $a = \cos\theta$ και $b = \sin\theta$, ο μετασχηματισμός μπορεί να γραφεί και ως εξής:

$$\begin{bmatrix} x_2 \\ y_2 \\ w_2 \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & t_x \\ -\sin\theta & \cos\theta & t_y \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ y_1 \\ w_1 \end{bmatrix} \tag{3.48}$$

Ο ευκλείδειος μετασχηματισμός έχει τρεις βαθμούς ελευθερίας, όσα και τα γεωμετρικά στοιχεία που εκφράζουν τη σχέση μεταξύ των δισδιάστατων συστημάτων συντεταγμένων.

3.4.1.3 Μετασχηματισμός ομοιότητας

Ο μετασχηματισμός ομοιότητας εφαρμόζεται όταν τα δύο δισδιάστατα συστήματα συντεταγμένων διαφέρουν κατά μία στροφή, κατά μία μετάθεση και κατά μία ενιαία κλίμακα. Στην περίπτωση των καρτεσιανών συντεταγμένων, ο μετασχηματισμός αυτός εκφράζεται ως εξής:

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \lambda \cdot \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} = \begin{bmatrix} \lambda \cdot \cos\theta & \lambda \cdot \sin\theta \\ -\lambda \cdot \sin\theta & \lambda \cdot \cos\theta \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \tag{3.49}$$

ή ισοδύναμα [126]:

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} a' & b' \\ -b' & a' \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \quad \text{υπό τη δέσμευση:} \quad (a')^2 + (b')^2 = \lambda^2 \tag{3.50}$$

όπου $a' = \lambda \cdot \cos\theta$ και $b' = \lambda \cdot \sin\theta$.

Σε ομογενείς συντεταγμένες, ο ίδιος μετασχηματισμός γράφεται ως εξής [94]:

$$\begin{bmatrix} x_2 \\ y_2 \\ w_2 \end{bmatrix} = \begin{bmatrix} \lambda \cdot \cos\theta & \lambda \cdot \sin\theta & t_x \\ -\lambda \cdot \sin\theta & \lambda \cdot \cos\theta & t_y \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ y_1 \\ w_1 \end{bmatrix} \tag{3.51}$$

$$\text{ή} \quad \begin{bmatrix} x_2 \\ y_2 \\ w_2 \end{bmatrix} = \begin{bmatrix} a' & b' & t_x \\ -b' & a' & t_y \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ y_1 \\ w_1 \end{bmatrix} \quad \text{υπό τη δέσμευση:} \quad (a')^2 + (b')^2 = \lambda^2 \tag{3.52}$$

Ο μετασχηματισμός ομοιότητας συχνά ονομάζεται και σύμμορφος μετασχηματισμός, αφού διατηρεί τα σχήματα. Για τον προσδιορισμό του απαιτείται η εύρεση τεσσάρων παραμέτρων, όσες και οι γεωμετρικές παράμετροι που περιγράφουν τη σχέση των δύο δισδιάστατων συστημάτων συντεταγμένων (δύο μεταθέσεις, στροφή και κλίμακα) [126]. Αυτές μπορούν να υπολογιστούν από τη λύση ενός συστήματος τεσσάρων εξισώσεων. Επομένως, απαιτούνται τουλάχιστον δύο σημεία γνωστά στα δύο συστήματα αφού το κάθε ένα από αυτά δίνει δύο εξισώσεις.

3.4.1.4 Αφινικός μετασχηματισμός

Ο αφινικός μετασχηματισμός εφαρμόζεται στην περίπτωση που το σύστημα συντεταγμένων του δεύτερου επιπέδου έχει μετατεθεί κατά x και y σε σχέση με το πρώτο, έχει υποστεί αλλαγή κλίμακας, διαφορετική κατά τους δύο άξονες, είναι στραμμένο σε σχέση με το πρώτο και παρατηρείται απόκλιση από την καθετότητα των δύο αξόνων του. Συνεπώς, περιλαμβάνει δύο επιπλέον παραμέτρους, σε σχέση με το μετασχηματισμό ομοιότητας [126]. Συγκεντρωτικά, οι έξι παράμετροι του αφινικού μετασχηματισμού είναι οι εξής:

- δύο μεταθέσεις κατά τις διευθύνσεις x και y
- στροφή
- δύο μεταβολές κλίμακας κατά τους άξονες x και y
- απόκλιση των αξόνων από την καθετότητα

Στην περίπτωση που χρησιμοποιούνται καρτεσιανές συντεταγμένες, ο αφινικός μετασχηματισμός εκφράζεται με την παρακάτω σχέση:

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ \varepsilon & 1 \end{bmatrix} \cdot \begin{bmatrix} \lambda_x & 0 \\ 0 & \lambda_y \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \tag{3.53}$$

όπου λ_x και λ_y οι διαφορετικοί συντελεστές κλίμακας κατά τις διευθύνσεις x και y αντίστοιχα και ε η απόκλιση των αξόνων από την καθετότητα. Κάνοντας τον πολλαπλασιασμό των τριών πρώτων πινάκων, προκύπτει η ακόλουθη σχέση:

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} \lambda_x \cdot (\cos\theta + \varepsilon \cdot \sin\theta) & \lambda_y \cdot \sin\theta \\ \lambda_x \cdot (-\sin\theta + \varepsilon \cdot \cos\theta) & \lambda_y \cdot \cos\theta \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \tag{3.54}$$

Επομένως, συμβολίζοντας ως $a_1 = \lambda_x \cdot (\cos\theta + \varepsilon \cdot \sin\theta)$, $b_1 = \lambda_y \cdot \sin\theta$, $a_2 = \lambda_x \cdot (-\sin\theta + \varepsilon \cdot \cos\theta)$ και $b_2 = \lambda_y \cdot \cos\theta$, η σχέση (3.54) γράφεται ως εξής:

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \tag{3.55}$$

Ο αφινικός μετασχηματισμός, μέσω των ομογενών διανυσμάτων, μπορεί να εκφραστεί με την παρακάτω απλή σχέση:

$$\begin{bmatrix} x_2 \\ y_2 \\ w_2 \end{bmatrix} = \begin{bmatrix} a_1 & b_1 & t_x \\ a_2 & b_2 & t_y \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ y_1 \\ w_1 \end{bmatrix} \tag{3.56}$$

ή ισοδύναμα:

$$\begin{bmatrix} x_2 \\ y_2 \\ w_2 \end{bmatrix} = \begin{bmatrix} \lambda_x \cdot (\cos\theta + \varepsilon \cdot \sin\theta) & \lambda_y \cdot \sin\theta & t_x \\ \lambda_x \cdot (-\sin\theta + \varepsilon \cdot \cos\theta) & \lambda_y \cdot \cos\theta & t_y \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ y_1 \\ w_1 \end{bmatrix} \tag{3.57}$$

Ο μετασχηματισμός αυτός, ο οποίος ονομάζεται και ομοπαράλληλος, έχει 6 βαθμούς ελευθερίας και για να προσδιοριστεί απαιτούνται ομολογίες τουλάχιστον τριών σημείων.

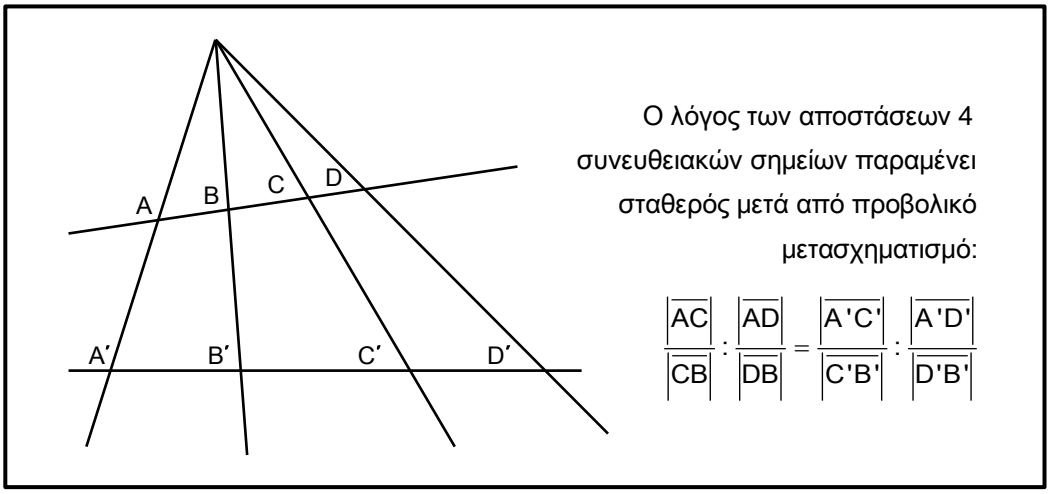
3.4.1.5 Προβολικός μετασχηματισμός (ομογραφία)

Ο προβολικός μετασχηματισμός, ή ομογραφία, είναι ένας αντιστρέψιμος γραμμικός μετασχηματισμός του προβολικού χώρου μεταξύ ομώνυμων γεωμετρικών στοιχείων (σημείων, ευθειών, κ.λπ.). Εφαρμόζεται σε ομογενείς συντεταγμένες και αναφέρεται συχνά και ως συγγραμμισμός (collineation) ή και προβολικότητα (projectivity). Πρόκειται για μετασχηματισμό ο οποίος ενώ δεν διατηρεί μήκη ή γωνίες, διατηρεί τις ευθείες γραμμές, ενώ άλλα αναλλοίωτα από το μετασχηματισμό μεγέθη αποτελούν ο διπλός λόγος (βλ. Εικόνα 3-8) και τα σημεία τομής των ευθειών, τα οποία εξακολουθούν να αποτελούν σημεία τομής και μετά το μετασχηματισμό [94], [129].

Ο δισδιάστατος προβολικός μετασχηματισμός εκφράζεται από έναν αντιστρέψιμο πίνακα διαστάσεων 3x3 μέσω ομογενών συντεταγμένων, σύμφωνα με την εξίσωση [94], [107]:

$$\begin{bmatrix} x_2 \\ y_2 \\ w_2 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ y_1 \\ w_1 \end{bmatrix} \quad \leftrightarrow \quad \mathbf{x_2} = \mathbf{H} \cdot \mathbf{x_1} \tag{3.58}$$

η οποία μετασχηματίζει ένα σημείο $\mathbf{x_1} = \begin{bmatrix} x_1 & y_1 & w_1 \end{bmatrix}^T$ στο $\mathbf{x_2} = \begin{bmatrix} x_2 & y_2 & w_2 \end{bmatrix}^T$. Ο μετασχηματισμός αυτός δεν μεταβάλλεται με τον πολλαπλασιασμό του πίνακα **H** με μη μηδενικό συντελεστή, με αποτέλεσμα ο **H** να αποτελεί ομογενή πίνακα. Συνεπώς, σημασία έχει μόνο ο λόγος των στοιχείων του πίνακα και καθώς μεταξύ των εννέα στοιχείων του είναι δυνατό να προκύψουν οκτώ διαφορετικοί λόγοι, η ομογραφία στο δισδιάστατο προβολικό χώρο έχει οκτώ βαθμούς ελευθερίας [107], [130], [131]. Επομένως, για τον προσδιορισμό της απαιτούνται ομολογίες τουλάχιστον τεσσάρων σημείων. Επίσης, ο συνδυασμός δύο ή και περισσότερων ομογραφιών συνιστά επίσης ομογραφία [94].



Εικόνα 3-8. Διπλός λόγος.

Η σχέση (3.58) μπορεί να αναπτυχθεί ώστε να εκφραστεί και σε μη ομογενείς (καρτεσιανές) συντεταγμένες. Από τη σχέση αυτή προκύπτει:

$$x_2 = h_{11} \cdot x_1 + h_{12} \cdot y_1 + h_{13} \cdot w_1 \tag{3.59}$$

$$y_2 = h_{21} \cdot x_1 + h_{22} \cdot y_1 + h_{23} \cdot w_1 \tag{3.60}$$

$$w_2 = h_{31} \cdot x_1 + h_{32} \cdot y_1 + h_{33} \cdot w_1 \tag{3.61}$$

Οι καρτεσιανές συντεταγμένες προκύπτουν ως: $x_2' = x_2 / w_2$ και $y_2' = y_2 / w_2$. Επομένως, σύμφωνα με τις παραπάνω σχέσεις και θεωρώντας $w_1 = 1$ [132], η ομογραφία σε καρτεσιανές συντεταγμένες εκφράζεται ως εξής:

$$x_2' = \frac{h_{11} \cdot x_1 + h_{12} \cdot y_1 + h_{13}}{h_{31} \cdot x_1 + h_{32} \cdot y_1 + h_{33}} \tag{3.62}$$

$$y_2' = \frac{h_{21} \cdot x_1 + h_{22} \cdot y_1 + h_{23}}{h_{31} \cdot x_1 + h_{32} \cdot y_1 + h_{33}} \tag{3.63}$$

Λαμβάνοντας υπ'όψιν ότι προκύπτει ο ίδιος προβολικός μετασχηματισμός με τον πολλαπλασιασμό του πίνακα **H** με μη μηδενική σταθερά, πολλαπλασιάζοντάς τον με $1/h_{33}$ η ομογραφία, βάσει της σχέσης (3.58), λαμβάνει την ακόλουθη μορφή:

$$\begin{bmatrix} x_2 \\ y_2 \\ w_2 \end{bmatrix} = \frac{1}{h_{33}} \cdot \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ y_1 \\ w_1 \end{bmatrix} \quad \leftrightarrow \tag{3.64}$$

$$\begin{bmatrix} x_2 \\ y_2 \\ w_2 \end{bmatrix} = \begin{bmatrix} h_{11}/h_{33} & h_{12}/h_{33} & h_{13}/h_{33} \\ h_{21}/h_{33} & h_{22}/h_{33} & h_{23}/h_{33} \\ h_{31}/h_{33} & h_{32}/h_{33} & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ y_1 \\ w_1 \end{bmatrix} \tag{3.65}$$

Και επομένως, μπορεί να συναντηθεί και ως εξής:

$$\begin{bmatrix} x_2 \\ y_2 \\ w_2 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ y_1 \\ w_1 \end{bmatrix} \tag{3.66}$$

Τότε, ο προβολικός μετασχηματισμός σε καρτεσιανές συντεταγμένες λαμβάνει την εξής μορφή:

$$x_2' = \frac{h_{11} \cdot x_1 + h_{12} \cdot y_1 + h_{13}}{h_{31} \cdot x_1 + h_{32} \cdot y_1 + 1} \tag{3.67}$$

$$y_2' = \frac{h_{21} \cdot x_1 + h_{22} \cdot y_1 + h_{23}}{h_{31} \cdot x_1 + h_{32} \cdot y_1 + 1} \tag{3.68}$$

Οι σχέσεις αυτές αποτελούν τις γνωστές σχέσεις της φωτογραμμετρικής αναγωγής, δηλαδή του προβολικού μετασχηματισμού της εικόνας σε ορθή προβολή [117]. Εκτός από τη χρήση του προβολικού μετασχηματισμού δύο διαστάσεων στην έκφραση της σχέσης μεταξύ της εικόνας και του επιπέδου του χώρου, η ομογραφία αποτελεί τη σχέση μεταξύ εικονοσημείων διαφορετικών εικόνων του ίδιου επίπεδου αντικειμένου οι οποίες μπορεί να λήφθηκαν από διαφορετικό σημείο λήψης και υπό διαφορετική γωνία, καθώς επίσης και μεταξύ εικονοσημείων ενός μη επίπεδου αντικειμένου σε εικόνες που λήφθηκαν από το ίδιο σημείο λήψης και διαφέρουν μόνο κατά μία στροφή [94], [111].

3.4.2 Μετασχηματισμοί μεταξύ εικονοσημείων διαφόρων εικόνων ενός δισδιάστατου αντικειμένου

Η ομογραφία αποτελεί τη γενικότερη σχέση η οποία συνδέει τα εικονοσημεία ενός δισδιάστατου αντικειμένου σε διαφορετικές εικόνες, οι οποίες μπορεί να διαφέρουν μεταξύ τους τόσο κατά θέση, όσο και κατά προσανατολισμό. Για να αποδειχτεί αυτό, αρκεί να αποδειχθεί ότι ο προβολικός αυτός μετασχηματισμός συνδέει την εικόνα ενός επίπεδου αντικειμένου με το πραγματικό αυτό δισδιάστατο αντικείμενο. Η ιδιότητα αυτή, εξάλλου, θα

συναντηθεί στην υποενότητα 3.5.2, όπου περιγράφεται μία προσέγγιση βαθμονόμησης από επίπεδο αντικείμενο, αλλά και στην ενότητα 3.6, όπου περιγράφεται μία προσέγγιση εύρεσης των στοιχείων του εξωτερικού προσανατολισμού μίας λήψης με χρήση σημείων γνωστών συντεταγμένων ενός επιπέδου. Έτσι, η εξίσωση (3.38) για εφαρμογή σε επίπεδο αντικείμενο, δηλαδή θέτοντας Z=0, λαμβάνει την ακόλουθη μορφή:

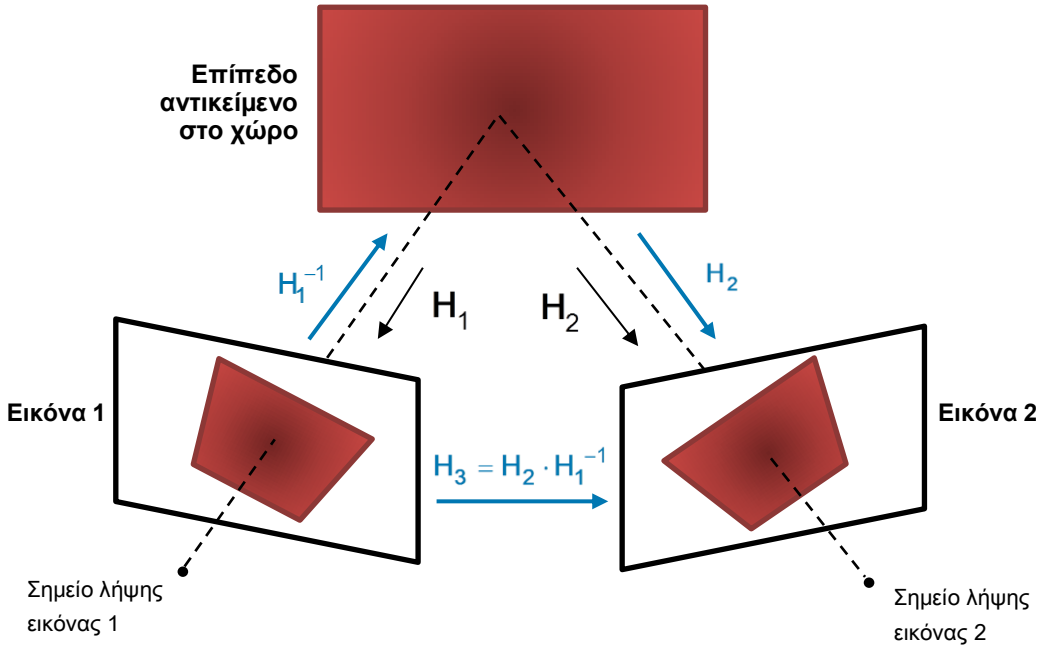
$$\lambda \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & s & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_X \\ r_{21} & r_{22} & r_{23} & t_Y \\ r_{31} & r_{32} & r_{33} & t_Z \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix} \quad \rightarrow$$

$$\lambda \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & s & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & t_X \\ r_{21} & r_{22} & t_Y \\ r_{31} & r_{32} & t_Z \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \quad \rightarrow \quad \lambda \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \quad \rightarrow$$

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \tag{3.69}$$

η οποία αποτελεί την εξίσωση του δισδιάστατου προβολικού μετασχηματισμού [111].

Αφού αποδείχθηκε ότι η ομογραφία συνδέει την εικόνα ενός δισδιάστατου αντικειμένου, με το πραγματικό αυτό αντικείμενο, μπορεί να αποδειχθεί επιπλέον ότι συνδέει και τις εικόνες του ίδιου επίπεδου αντικειμένου. Όπως φαίνεται και στην Εικόνα 3-9, έστω ότι συμβολίζεται ως $\mathbf{H_1}$ ο 3x3 πίνακας προβολικού μετασχηματισμού που μετατρέπει τις συντεταγμένες χώρου του επίπεδου αντικειμένου σε εικονοσυντεταγμένες της εικόνας 1 και ως $\mathbf{H_2}$ ο αντίστοιχος πίνακας που μετατρέπει τις συντεταγμένες χώρου του ίδιου επίπεδου αντικειμένου σε εικονοσυντεταγμένες της εικόνας 2. Για τον υπολογισμό των εικονοσυντεταγμένων του επίπεδου αντικειμένου της εικόνας 2, με βάση τις αντίστοιχες της εικόνας 1, εφαρμόζεται πρώτα ο προβολικός μετασχηματισμός από την εικόνα 1 στο επίπεδο του χώρου (με βάση τον αντίστροφο πίνακα μετασχηματισμού, $\mathbf{H_1}^{-1}$) και εν συνεχεία ο μετασχηματισμός από το επίπεδο του χώρου στην εικόνα 2 (με βάση τον πίνακα μετασχηματισμού $\mathbf{H_2}$). Έτσι, ο υπολογισμός των εικονοσυντεταγμένων του δισδιάστατου αντικειμένου στην εικόνα 2, με βάση τις εικονοσυντεταγμένες του στην εικόνα 1, προκύπτει μέσω της αλληλουχίας των δύο παραπάνω μετασχηματισμών. Συνεπώς, αν συμβολιστεί ως $\mathbf{H_3}$ ο πίνακας προβολικού μετασχηματισμού των εικονοσυντεταγμένων του επίπεδου αντικειμένου της εικόνας 1 σε αντίστοιχες της εικόνας 2, τότε ισχύει: $\mathbf{H_3} = \mathbf{H_2} \cdot \mathbf{H_1}^{-1}$. Καθώς ο συνδυασμός δύο ή και περισσότερων ομογραφιών συνιστά επίσης ομογραφία, η τελευταία είναι η σχέση που συνδέει τις εικόνες του ίδιου επίπεδου αντικειμένου [111].



Εικόνα 3-9. Προβολικός μετασχηματισμός μεταξύ επίπεδου αντικειμένου του χώρου και της εικόνας του και μεταξύ εικόνων του ίδιου δισδιάστατου αντικειμένου.

Εκτός από την ομογραφία, η οποία μπορεί να εφαρμοστεί σε όλες τις περιπτώσεις μετασχηματισμού εικονοσυντεταγμένων μεταξύ εικόνων του ίδιου δισδιάστατου αντικειμένου, με κάποιες παραδοχές και υπό προϋποθέσεις μπορεί να εφαρμοστεί και ο αφινικός μετασχηματισμός. Θεωρείται, έτσι, το παράδειγμα του αφινικού μετασχηματισμού μεταξύ μίας εικόνας ενός δισδιάστατου αντικειμένου, που έχει ληφθεί με μηδενικές γωνίες στροφής, και μίας άλλης τυχαίας εικόνας του ίδιου αντικειμένου. Στην περίπτωση που η δεύτερη εικόνα έχει μικρές προοπτικές παραμορφώσεις (βλ. Εικόνα 3-10a), που μπορεί να είναι αποτέλεσμα μικρών γωνιών στροφής, μεγάλης απόστασης από το αντικείμενο, και χρήσης οξυγώνιου φακού (μεγάλης εστιακής απόστασης), ο αφινικός μετασχηματισμός μεταξύ αυτής και της πρώτης μπορεί να δώσει ικανοποιητικά αποτελέσματα. Στην Εικόνα 3-10a, με κόκκινο περίγραμμα ενώνονται τα άκρα της πρώτης εικόνας, χωρίς στροφές, του δισδιάστατου αντικειμένου που μετασχηματίστηκαν μέσω του αφινικού μετασχηματισμού σε άκρα του ίδιου αντικειμένου στη δεύτερη εικόνα. Ωστόσο, στην περίπτωση που η δεύτερη εικόνα παρουσιάζει έντονες προοπτικές παραμορφώσεις, όπως στην Εικόνα 3-10b, λόγω ευρυγώνιου φακού, μικρής απόστασης από το αντικείμενο και λήψης με γωνίες στροφής που δεν μπορούν να θεωρηθούν μηδενικές, ο αφινικός μετασχηματισμός δεν μπορεί να περιγράψει τη σχέση μεταξύ των εικόνων. Ο βασικός περιορισμός της χρήσης του είναι η διατήρηση της παραλληλίας των γραμμών. Στην περίπτωση αυτή, ο προβολικός μετασχηματισμός είναι αυτός ο οποίος μπορεί να περιγράψει τη σχέση μεταξύ των εικόνων (βλ. Εικόνα 3-10c), ο οποίος διατηρεί μόνο τις ευθείες γραμμές και όχι την παραλληλία μεταξύ αυτών [111].

Υπό πολύ συγκεκριμένες προϋποθέσεις και υιοθετώντας απλές παραδοχές, θα μπορούσαν να εφαρμοστούν και οι υπόλοιποι γραμμικοί δισδιάστατοι μετασχηματισμοί στην έκφραση των σχέσεων μεταξύ δύο εικόνων. Στην πράξη, ωστόσο, προτιμάται η χρήση του μετασχηματισμού ομογραφίας για μεγαλύτερη ασφάλεια και ακριβή αποτελέσματα.

Εικόνα 3-10. Εφαρμογή αφινικού μετασχηματισμού (a,b) και ομογραφίας (c).

Πηγή: Prince, 2012

3.5 Βαθμονόμηση

Με την αποκατάσταση του εσωτερικού προσανατολισμού πραγματοποιείται σωστή ανάπλαση της δέσμης των ακτίνων που δημιούργησε την εικόνα, δηλαδή αποκαθίσταται η πορεία των φωτεινών ακτίνων στο εσωτερικό της μηχανής, όπως αυτή υπήρχε κατά τη στιγμή της λήψης. Προκειμένου να αποκατασταθεί ο εσωτερικός προσανατολισμός, απαιτείται ο προσδιορισμός των στοιχείων του, ο οποίος πραγματοποιείται μέσω της βαθμονόμησης της φωτογραφικής μηχανής. Η τελευταία, αναφέρεται και ως καλιμπράρισμα (από τον αγγλικό όρο calibration) και αποτελεί το σύνολο των μετρητικών και υπολογιστικών διαδικασιών που προσδιορίζουν τα στοιχεία του εσωτερικού προσανατολισμού της φωτομηχανής [117]. Αυτά είναι η σταθερά της μηχανής, δηλαδή η απόσταση του προβολικού κέντρου από το επίπεδο της εικόνας, και η θέση του πρωτεύοντος σημείου στο επίπεδο της εικόνας και σε αυτά συνήθως προστίθενται οι συντελεστές του πολυωνύμου της συμμετρικής ακτινικής διαστροφής και – σπανιότερα – της ασύμμετρης διαστροφής εκκεντρότητας.

3.5.1 Ανασκόπηση διαφόρων μεθόδων βαθμονόμησης

Υφίστανται διάφορες μέθοδοι βαθμονόμησης, οι οποίες μπορούν να ταξινομηθούν με βάση διάφορα κριτήρια, ένα εκ των οποίων είναι η απαίτηση για λήψη εικόνων ή η πραγματοποίηση της βαθμονόμησης χωρίς τη φωτογράφηση. Με βάση αυτό το κριτήριο, οι μέθοδοι βαθμονόμησης διακρίνονται σε οπτικές και φωτογραφικές.

Οι οπτικές μέθοδοι βαθμονόμησης ονομάζονται συχνά και εργαστηριακές, αφού διενεργούνται σε ειδικά εξοπλισμένα εργαστήρια, και περιλαμβάνουν τη μέθοδο των ευθυγραμμιστών και τη μέθοδο του γωνιομέτρου. Αυτές βασίζονται στη λογική

προσδιορισμού της σταθεράς της μηχανής μέσω παρατηρήσεων γωνιών και ακτινικών αποστάσεων με χρήση της συνάρτησης της εικόνας, η οποία εκφράζει την κεντρική προβολή συνδέοντας γωνίες του χώρου με στοιχεία της εικόνας μέσω παραμέτρων του εσωτερικού προσανατολισμού [115], [117], [133].

Μία κοινή στη φωτογραμμετρία φωτογραφική μέθοδος βαθμονόμησης είναι η χρήση τρισδιάστατου πεδίου ελέγχου με σημεία γνωστών συντεταγμένων για την εύρεση του εσωτερικού προσανατολισμού ή την ταυτόχρονη εύρεση του εσωτερικού και εξωτερικού προσανατολισμού από μία ή περισσότερες εικόνες, μέσω της συνθήκης συγγραμμικότητας. Στην περίπτωση αυτή, η ύπαρξη μεγάλου αριθμού σημείων κατανεμημένων σε διαφορετικά επίπεδα έχει μεγάλη σημασία για την εύρεση των στοιχείων του εσωτερικού προσανατολισμού. Επίσης, η βαθμονόμηση με χρήση επίπεδων πεδίων ελέγχου με σημεία άγνωστων συντεταγμένων ή και με χρήση χαρακτηριστικών επίπεδων αντικειμένων, που περιλαμβάνουν εύκολα αναγνωρίσιμα σημεία, αποτελεί συχνά προτιμώμενη μέθοδο βαθμονόμησης. Συνήθως, στο πλαίσιο μίας τέτοιας προσέγγισης δεν είναι γνωστός ούτε ο εξωτερικός προσανατολισμός, με αποτέλεσμα να πραγματοποιείται μία ταυτόχρονη διαδικασία εύρεσης εσωτερικού και εξωτερικού προσανατολισμού μέσω πολλαπλών φωτογραφικών λήψεων του επίπεδου αντικειμένου από διάφορες γωνίες και αποστάσεις από αυτό. Μία τέτοια προσέγγιση υιοθετείται και στην παρούσα εργασία για τις ανάγκες των εφαρμογών. Άλλες προσεγγίσεις φωτογραφικών μεθόδων βαθμονόμησης είναι η βαθμονόμηση με χρήση πεδίου ελέγχου με χαρακτηριστικά από το χώρο του αντικειμένου, όπως για παράδειγμα γραμμικά στοιχεία. Ευρύτερη έννοια της βαθμονόμησης αποτελεί η αυτοβαθμονόμηση, κατά την οποία τα στοιχεία του εσωτερικού προσανατολισμού προσδιορίζονται την ίδια στιγμή σε μία ενιαία λύση με τον προσδιορισμό των τρισδιάστατων συντεταγμένων των στόχων [133], [134].

Η κατηγοριοποίηση των μεθόδων βαθμονόμησης μπορεί να γίνει και με βάση τη χρησιμοποιούμενη μαθηματική εξίσωση, η οποία μπορεί να είναι είτε η συνθήκη συγγραμμικότητας είτε ο προβολικός μετασχηματισμός. Ακόμη, η διάκριση γίνεται και με βάση τη γραμμικότητα ή μη του συστήματος εξισώσεων που δημιουργείται για την εύρεση των στοιχείων του εσωτερικού προσανατολισμού. Μάλιστα, μπορεί να επιτευχθεί και ένας συνδυασμός γραμμικών και μη γραμμικών εξισώσεων, όπου υπολογίζονται αρχικές τιμές των παραμέτρων από μία γραμμική επίλυση και στη συνέχεια αυτές εισάγονται ως προσεγγιστικές τιμές στο μη γραμμικό σύστημα εξισώσεων, για τη βελτίωση της προηγούμενης λύσης [135].

3.5.2 Βαθμονόμηση με λήψη εικόνων επίπεδης σκακιέρας

Γενικά, η βαθμονόμηση από πολλές λήψεις ενός επίπεδου αντικειμένου με χαρακτηριστικά σημεία, όπως είναι η σκακιέρα, επιτυγχάνεται με δεδομένα τις συντεταγμένες χώρου και

εικόνας των αντίστοιχων σημείων, οι πρώτες εκ των οποίων προκύπτουν αυτόματα, ενώ οι τελευταίες αναγνωρίζονται στην εικόνα ομοίως με αυτόματο τρόπο. Ωστόσο, για την εξασφάλιση σωστών αποτελεσμάτων από τη διαδικασία της βαθμονόμησης, οι εικόνες του αντικειμένου πρέπει να ληφθούν από διαφορετικά σημεία και υπό διαφορετικές γωνίες, δημιουργώντας με αυτόν τον τρόπο «ανάγλυφο» στο αντικείμενο. Επίσης, η διαδικασία αυτή απαιτεί τον επιπρόσθετο υπολογισμό του εξωτερικού προσανατολισμού κάθε λήψης, με αποτέλεσμα να αποτελεί στην ουσία μία διαδικασία ταυτόχρονης εύρεσης του εσωτερικού προσανατολισμού της μηχανής και του εξωτερικού προσανατολισμού κάθε εικόνας.

Στο πλαίσιο της παρούσας εργασίας, για τις ανάγκες των εφαρμογών έγινε βαθμονόμηση με λήψη εικόνων επίπεδης σκακιέρας, με χρήση ελεύθερου κώδικα της βιβλιοθήκης OpenCV. Στην παρούσα υποενότητα παρουσιάζονται τα κύρια σημεία της διαδικασίας της βαθμονόμησης, τα οποία προσεγγίζουν εκείνη που υιοθετείται από τη βιβλιοθήκη OpenCV στην περίπτωση διεξαγωγής της βαθμονόμησης με χρήση επίπεδης σκακιέρας. Ωστόσο, δεν δίνεται έμφαση στη λεπτομερή παρουσίαση κάθε βήματος, καθώς ξεφεύγει από τους σκοπούς της παρούσας διπλωματικής εργασίας. Η διαδικασία αυτή είναι αυτόματη, ενώ τα μόνα δεδομένα εισόδου είναι οι εικόνες της σκακιέρας και το σχέδιό της, δηλαδή ο αριθμός των εσωτερικών γωνιών της στις δύο κάθετες διευθύνσεις.

3.5.2.1 Αρχική επεξεργασία κάθε εικόνας

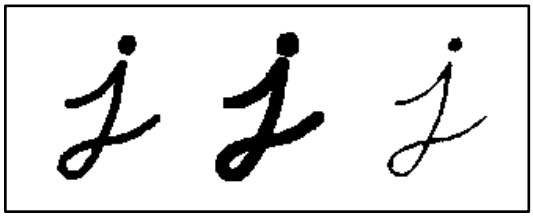
Αρχικό βήμα της βαθμονόμησης είναι η ανίχνευση των εσωτερικών γωνιών της επίπεδης ασπρόμαυρης σκακιέρας, δηλαδή των κοινών κορυφών των μαύρων τετραγώνων σε κάθε εικόνα εισόδου. Προκειμένου αυτές να εντοπιστούν, διεξάγεται μία σειρά βημάτων, τα οποία περιγράφονται στις επόμενες παραγράφους. Πρώτο βήμα αυτής της διαδικασίας είναι η μετατροπή της κάθε εικόνας εισόδου από έγχρωμη σε ασπρόμαυρη – βήμα το οποίο παραλείπεται στην περίπτωση εισόδου ασπρόμαυρων εικόνων – και η ισοστάθμιση του ιστογράμματός της, ώστε να αυξηθεί η αντίθεσή της.

3.5.2.2 Έλεγχος εντοπισμού της σκακιέρας σε κάθε εικόνα

Στη συνέχεια γίνεται έλεγχος αν μπορεί να εντοπιστεί η σκακιέρα στη συγκεκριμένη ασπρόμαυρη και ισοσταθμισμένη εικόνα, ώστε να μην ακολουθήσει η διαδικασία ανίχνευσης των εσωτερικών γωνιών της σε περίπτωση που η εικόνα δεν περιέχει τη σκακιέρα ή αν αυτή κρύβεται μερικώς. Ο έλεγχος αυτός περιλαμβάνει αρχικά συστολή (erosion) της εικόνας και διαστολή (dilation) της ίδιας εικόνας.

Η συστολή και η διαστολή είναι δύο μορφολογικές επεξεργασίες, οι οποίες μπορούν να πραγματοποιηθούν με φιλτράρισμα της εικόνας με φίλτρο ελάχιστης και μέγιστης τιμής αντίστοιχα (βλ. Εικόνα 3-11). Συγκεκριμένα, στην περίπτωση της συστολής, το εικονοστοιχείο που βρίσκεται κάθε φορά στο κέντρο του φίλτρου αντικαθίσταται από την

ελάχιστη τιμή των (n x n – 1) γειτονικών του εικονοστοιχείων, όπου n x n είναι το μέγεθος του φίλτρου, εφόσον αυτό είναι τετράγωνο. Συνεπώς, το μέγεθος των άσπρων περιοχών της εικόνας, και γενικά των περιοχών πιο ανοικτών τόνων, ελαττώνεται. Αντίστοιχα, κατά την εφαρμογή της διαστολής, το εικονοστοιχείο που βρίσκεται κάθε φορά στο κέντρο του φίλτρου αντικαθίσταται από τη μέγιστη τιμή των γειτονικών του εικονοστοιχείων. Αποτέλεσμα της διαστολής είναι η αύξηση του μεγέθους των άσπρων και γενικότερα των περιοχών ανοικτών τόνων της εικόνας.

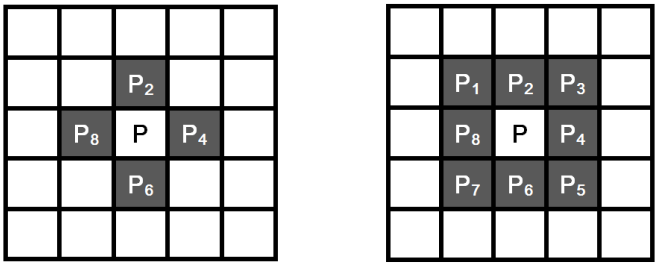


Εικόνα 3-11. Αρχική εικόνα [αριστερά], συστολή της [κέντρο] και διαστολή της [δεξιά].

Πηγή: http://docs.opencv.org/doc/tutorials/imgproc/erosion_dilatation/erosion_dilatation.html

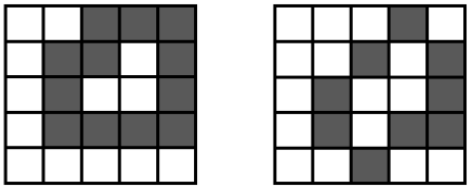
Επόμενο βήμα είναι η κατωφλίωση της εικόνας που έχει υποστεί συστολή, έτσι ώστε να μετατραπεί σε δυαδική εικόνα, δηλαδή εικόνα που έχει δύο στάθμες φωτεινότητας, το άσπρο και το μαύρο που αναπαρίστανται από τους αριθμούς 0 και 255 [104], [136]. Ακολούθως, γίνεται ανίχνευση των ορίων - περιγραμμάτων (contours) της ασπρόμαυρης δυαδικής εικόνας. Αυτά αποτελούν μία λίστα σημείων που απεικονίζουν μία συνοριακή γραμμή / καμπύλη και υποδηλώνουν τα όρια ενός αντικειμένου της εικόνας. Στην προκειμένη περίπτωση, ανιχνεύονται όλα τα περιγράμματα και οργανώνονται σε ιεραρχία δύο επιπέδων. Το πρώτο επίπεδο περιλαμβάνει τα εξωτερικά όρια (outer borders), ενώ στο δεύτερο επίπεδο περιλαμβάνονται τα όρια των τρυπών (hole borders). Αν υπάρχει ένα ακόμη περίγραμμα εντός μίας τρύπας, αυτό τοποθετείται στο πρώτο επίπεδο, κ.ο.κ. Ο αλγόριθμος που χρησιμοποιείται από την OpenCV για την ανίχνευση των περιγραμμάτων στηρίζεται στην εργασία των Satoshi Suzuki και Keiichi Abe (1985) [137], στην οποία προτείνεται μία τεχνική κατά την οποία ακολουθούνται τα όρια - περιγράμματα που εντοπίζονται στην εικόνα για την εξαγωγή της τοπολογικής δομής της, με αποτέλεσμα την εξαγωγή ακολουθιών συντεταγμένων που ορίζουν τα περιγράμματα αλλά και πληροφοριών για τη σχέση μεταξύ των ορίων.

Πριν την περιγραφή της λογικής με βάση την οποία ο αλγόριθμος εντοπίζει τα όρια αυτά, κρίνεται σκόπιμη η παράθεση κάποιων βασικών ορισμών. Οι τέσσερις άμεσοι γείτονες ενός εικονοστοιχείου P είναι τα εικονοστοιχεία P_2, P_4, P_6 και P_8 που φαίνονται στην Εικόνα 3-12, αριστερά. Κάθε ένα από τα εικονοστοιχεία αυτά έχει κοινή πλευρά με το P, ονομάζεται 4-γείτονας (4-neighbour) του P και θεωρείται ότι ανήκει στο σύνολο $N_4(P)$. Οι οκτώ γείτονες του P είναι τα εικονοστοιχεία P_1, …, P_8 που φαίνονται στην Εικόνα 3-12, δεξιά, καθένα από τα οποία έχει κοινή πλευρά ή κορυφή με το εικονοστοιχείο P. Τα εικονοστοιχεία P_1, …, P_8 ονομάζονται 8-γείτονες (8-neighbours) του P και θεωρείται ότι ανήκουν στο σύνολο $N_8(P)$.



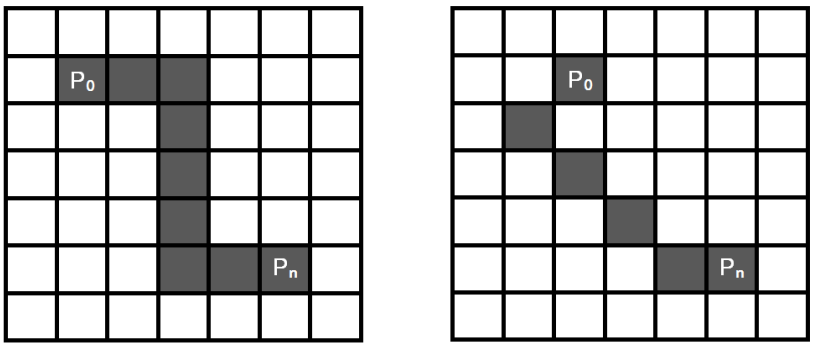
Εικόνα 3-12. 4-γείτονες (4-neighbours) [αριστερά] και 8-γείτονες (8-neighbours) [δεξιά] ενός εικονοστοιχείου P.

Συνδεδεμένο συστατικό (connected component) ονομάζεται ένα σύνολο εικονοστοιχείων που έχουν την ίδια τιμή και είναι συνδεδεμένα μεταξύ τους, δηλαδή είναι γείτονες. Ειδικότερα, σε ένα 4-συνδεδεμένο συστατικό (4-connected component) ισχύει η ακόλουθη δέσμευση: για κάθε δύο γειτονικά εικονοστοιχεία του, P και Q, το Q ανήκει στην ομάδα $N_4(P)$ (βλ. Εικόνα 3-13, αριστερά). Αντίστοιχα, σε ένα 8-συνδεδεμένο συστατικό (8-connected component) ισχύει η δέσμευση ότι για κάθε δύο γειτονικά εικονοστοιχεία του, P και Q, το Q ανήκει στην ομάδα $N_8(P)$ (βλ. Εικόνα 3-13, δεξιά).



Εικόνα 3-13. Παράδειγμα 4-συνδεδεμένου συστατικού (4-connected component) [αριστερά] και 8-συνδεδεμένου συστατικού (8-connected component) [δεξιά].

Μία 4-διαδρομή (4-path) από το εικονοστοιχείο P_0 στο εικονοστοιχείο P_n είναι μία αλληλουχία εικονοστοιχείων $P_0, P_1, P_2, \ldots, P_n$ τέτοια ώστε το εικονοστοιχείο P_k να ανήκει στο $N_4(P_{k+1})$, για k=0,1,…,n-1 (βλ. Εικόνα 3-14, αριστερά). Αντίστοιχα, μία 8-διαδρομή (8-path) από το P_0 στο P_n είναι μία αλληλουχία εικονοστοιχείων $P_0, P_1, P_2, \ldots, P_n$ τέτοια ώστε το εικονοστοιχείο P_k να ανήκει στο $N_8(P_{k+1})$ για k=0,1,…,n-1 (βλ. Εικόνα 3-14, δεξιά) [138], [139].



Εικόνα 3-14. Παράδειγμα 4-διαδρομής (4-path) [αριστερά] και 8-διαδρομής (8-path) [δεξιά].

Ο αλγόριθμος εφαρμόζεται σε δυαδικές εικόνες που αναπαριστούν το μαύρο τόνο με την τιμή 0 και τον άσπρο τόνο με την τιμή 1, ενώ το περίγραμμα της εικόνας λαμβάνει την τιμή 0.

Επίσης, θεωρεί ως φόντο της εικόνας (background) ένα συνδεδεμένο συστατικό μαύρων εικονοστοιχείων που περιλαμβάνει το περίγραμμα της εικόνας και ως τρύπα (hole) ένα συνδεδεμένο συστατικό μαύρων εικονοστοιχείων που δεν περιλαμβάνει το περίγραμμα της εικόνας [137].

Στο πλαίσιο του αλγορίθμου, ο όρος: «περίπτωση 4-σύνδεσης» αναφέρεται στο 4-συνδεδεμένο συστατικό άσπρων εικονοστοιχείων, που αντιστοιχεί στο 8-συνδεδεμένο συστατικό μαύρων εικονοστοιχείων. Αντίστοιχα, ο όρος «περίπτωση 8-σύνδεσης» αναφέρεται στο 8-συνδεδεμένο συστατικό άσπρων εικονοστοιχείων, που αντιστοιχεί στο 4-συνδεδεμένο συστατικό μαύρων εικονοστοιχείων [137].

Αρχικά βήματα της διαδικασίας είναι η θεώρηση της τιμής 0 για τα εικονοστοιχεία του περιγράμματος της εικόνας και των τιμών 0 και 1 για τα μαύρα και άσπρα εικονοστοιχεία αντίστοιχα της δυαδικής εικόνας. Ακολούθως, καθορίζονται τα όρια (contours) και εξάγεται η τοπολογική δομή της δυαδικής εικόνας, με βάση τη λογική και τις παραδοχές που αναφέρονται παρακάτω [137]:

- Στην περίπτωση 4- (ή 8-) σύνδεσης, έστω ένα άσπρο pixel p και ένα μαύρο pixel q. Αν το q ανήκει στο σύνολο $N_8(p)$ (ή $N_4(p)$ αντίστοιχα), το p ονομάζεται σημείο ορίου (border point).
- Για δύο συνδεδεμένα συστατικά S_1 και S_2, αν υπάρχει ένα εικονοστοιχείο που ανήκει στο S_2, για οποιαδήποτε 4-διαδρομή από ένα εικονοστοιχείο στο S_1, σε ένα εικονοστοιχείο στο περίγραμμα της εικόνας, τότε θεωρείται ότι το S_2 περιβάλλει το S_1. Αν επιπρόσθετα υπάρχει ένα σημείο ορίου μεταξύ των S_1 και S_2, το S_2 περιβάλλει άμεσα το S_1.
- Ως εξωτερικό όριο (outer border) ορίζεται το σύνολο των σημείων ορίου μεταξύ ενός συνδεδεμένου συστατικού άσπρων εικονοστοιχείων και ενός συνδεδεμένου συστατικού μαύρων εικονοστοιχείων που το περιβάλλει άμεσα. Αντίστοιχα, όριο τρύπας (hole border) είναι το σύνολο των σημείων ορίου μεταξύ μίας τρύπας (hole) και του συνδεδεμένου συστατικού άσπρων εικονοστοιχείων που την περιβάλλει άμεσα. Τόσο το όριο τρύπας όσο και το εξωτερικό όριο ορίζονται ως η ομάδα των άσπρων εικονοστοιχείων.
- Το όριο - γονέας (parent border) ενός εξωτερικού ορίου μεταξύ ενός συνδεδεμένου συστατικού άσπρων εικονοστοιχείων S_1 και ενός συνδεδεμένου συστατικού μαύρων εικονοστοιχείων S_2, που περιβάλλει το S_1, ορίζεται ως το όριο τρύπας μεταξύ του S_2 και του συνδεδεμένου συστατικού άσπρων εικονοστοιχείων που περιβάλλει άμεσα το S_2, αν το S_2 είναι τρύπα, ή το περίγραμμα της εικόνας, αν το S_2 είναι το φόντο (background) της εικόνας. Αντίστοιχα, το όριο - γονέας ενός ορίου τρύπας μεταξύ μίας τρύπας S_3 και του συνδεδεμένου συστατικού άσπρων εικονοστοιχείων S_4 που

περιβάλλει άμεσα την S_3 ορίζεται ως το εξωτερικό όριο μεταξύ του S_4 και του συνδεδεμένου συστατικού μαύρων εικονοστοιχείων που περιβάλλει άμεσα το S_4.

- Για δύο όρια της εικόνας, B_0 και B_n, θεωρείται ότι το B_n περιβάλλει το B_0 αν υπάρχει μία ακολουθία ορίων B_0, B_1, …, B_n, τέτοια ώστε το όριο B_k να είναι το όριο - γονέας του ορίου B_{k-1} για όλα τα $k \in [1,n]$.
- Για ένα τυχαίο συνδεδεμένο συστατικό άσπρων εικονοστοιχείων, το εξωτερικό του όριο είναι μοναδικό. Ομοίως, το όριο τρύπας είναι μοναδικό.
- Ένα συνδεδεμένο συστατικό άσπρων εικονοστοιχείων περιβάλλεται από το εξωτερικό του όριο. Μία «τρύπα» περιβάλλεται από το όριο «τρύπας» μεταξύ της τρύπας και του συνδεδεμένου συστατικού άσπρων εικονοστοιχείων που την περιβάλλει άμεσα. Το φόντο περιβάλλεται από το περίγραμμα της εικόνας.

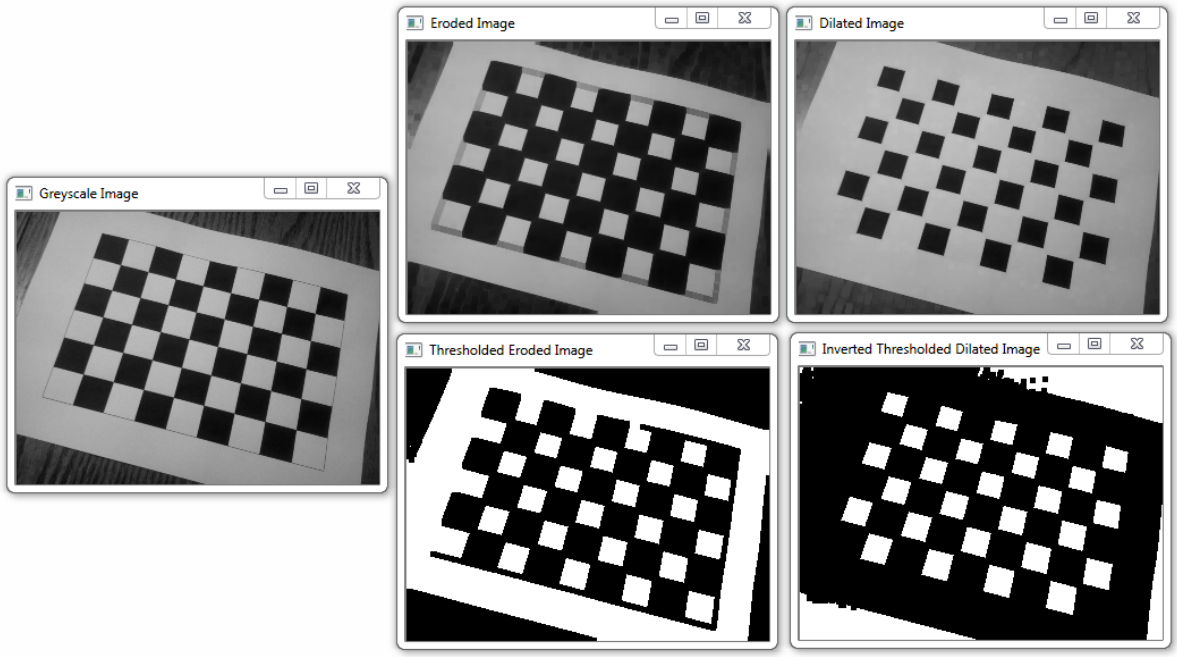
Με βάση τους παραπάνω ορισμούς και ιδιότητες, ο αλγόριθμος ακολουθεί τα όρια της εικόνας, αναλύοντας την τοπολογική δομή της. Έτσι, ξεκινά σκανάροντας την εικόνα από το επάνω αριστερά εικονοστοιχείο, έως ότου εντοπίσει ένα εικονοστοιχείο που είναι είτε σημείο εξωτερικού ορίου, είτε σημείο ορίου τρύπας. Από τη στιγμή που το εντοπίζει, το ακολουθεί, «σημειώνοντας» τα εικονοστοιχεία του ορίου, δηλαδή δίνοντας σε αυτά ένα μοναδικό κωδικό αριθμό, και προσδιορίζει το όριο - γονέα του. Αφού ακολουθηθεί όλο το όριο, το σκανάρισμα συνεχίζει από το σημείο στο οποίο σταμάτησε και ο αλγόριθμος τελειώνει όταν το σκανάρισμα της εικόνας φτάσει στο κάτω δεξιά εικονοστοιχείο της.

Από την παραπάνω διαδικασία εύρεσης των περιγραμμάτων, μπορούν να εντοπιστούν οι γραμμές που οριοθετούν τα τετράγωνα της σκακιέρας. Ωστόσο, μπορεί να εντοπιστούν περιγράμματα και για άλλα αντικείμενα της εικόνας. Για το λόγο αυτό, ακολουθεί η απόρριψη των περιγραμμάτων που δεν οριοθετούν τετράγωνα της σκακιέρας. Έτσι, τα περιγράμματα απλοποιούνται ώστε να έχουν μόνο τέσσερις κορυφές. Προκειμένου να πραγματοποιηθεί η διαδικασία αυτή, εντοπίζεται το ορθογώνιο που περιβάλλει την κάθε γραμμή - περίγραμμα (contour) και έχει το μικρότερο εμβαδό. Αν η μεγαλύτερη διάστασή του είναι μικρότερη από ένα προκαθορισμένο μέγεθος, θεωρείται ότι το ορθογώνιο δεν αποτελεί τετράγωνο της σκακιέρας και συνεπώς δεν λαμβάνεται υπ'όψιν. Επίσης, το ορθογώνιο δεν θεωρείται ότι είναι εικόνα του τετραγώνου της σκακιέρας εφόσον ο λόγος των πλευρών του είναι μεγαλύτερος ή μικρότερος κάποιων κατωφλίων. Με βάση τη διαδικασία αυτή, εντοπίζονται τα πιθανά περιγράμματα των άσπρων τετραγώνων της σκακιέρας.

Επόμενο βήμα είναι η κατωφλίωση της εικόνας που έχει υποστεί διαστολή έτσι ώστε να μετατραπεί σε δυαδική εικόνα. Η κατωφλίωση αυτή γίνεται με μικρότερο κατώφλι σε σχέση με την κατωφλίωση της εικόνας που έχει υποστεί συστολή. Επιπλέον, επιχειρείται αντιστροφή των τιμών, δηλαδή τα σκούρα pixels που έχουν τιμή κάτω του κατωφλίου λαμβάνουν την τιμή 255, ενώ αντίθετα τα ανοικτά pixels λαμβάνουν την τιμή 0 και γίνονται μαύρα. Στη συνέχεια, ομοίως με τον αλγόριθμο των Suzuki και Abe [137], γίνεται εξαγωγή

των περιγραμμάτων της εικόνας και κατόπιν εντοπίζονται τα πιθανά μαύρα τετράγωνα, σύμφωνα με τη διαδικασία που περιγράφηκε παραπάνω.

Τα περιγράμματα των πιθανών τετραγώνων της σκακιέρας, που εντοπίστηκαν με βάση την εξαγωγή πληροφοριών από την εικόνα που έχει υποστεί τη μορφολογική επεξεργασία της συστολής και από εκείνη που έχει υποστεί διαστολή, ελέγχονται ως προς το μέγεθός τους και τα ορθογώνια που διαφέρουν σε μέγεθος απορρίπτονται. Στη συνέχεια ελέγχεται ο αριθμός των μαύρων και άσπρων τετραγώνων. Αν τα άσπρα ή τα μαύρα τετράγωνα είναι λιγότερα από ένα ποσοστό του αριθμού τους στην πραγματική σκακιέρα, τότε θεωρείται ότι αυτή δεν βρέθηκε και η διαδικασία που περιγράφηκε μετά από τη συστολή και τη διαστολή της ίδιας εικόνας επαναλαμβάνεται για μεγαλύτερη τιμή κατωφλίων, έως ότου εντοπιστεί η σκακιέρα ή μέχρι το μικρότερο κατώφλι (στην περίπτωση της διαστολής) φτάσει μία προκαθορισμένη τιμή [140], [141], [142].



Εικόνα 3-15. Παράδειγμα συστολής, διαστολής και κατωφλίωσης σε ασπρόμαυρη εικόνα σκακιέρας.

Όσον αφορά στο διαχωρισμό των μαύρων και άσπρων τετραγώνων της σκακιέρας, αυτός επιτυγχάνεται χάριν της συστολής και της διαστολής, καθώς μέσω της πρώτης εντοπίζονται τα άσπρα τετράγωνα, ενώ μέσω της δεύτερης τα μαύρα. Η διαδικασία αυτή μπορεί να γίνει εύκολα κατανοητή από την Εικόνα 3-15, στην οποία αριστερά απεικονίζεται μία ασπρόμαυρη εικόνα σκακιέρας. Η εικόνα της συστολής (βλ. επάνω - κέντρο στην Εικόνα 3-15) έχει ως αποτέλεσμα την αύξηση του μεγέθους των μαύρων τετραγώνων, με συνέπεια την εύκολη εξαγωγή των περιγραμμάτων των άσπρων τετραγώνων, ύστερα από την κατωφλίωση (βλ. κάτω - κέντρο στην Εικόνα 3-15). Αντίστοιχα, στην εικόνα της διαστολής (βλ επάνω - δεξιά στην Εικόνα 3-15), αυξάνεται το μέγεθος των άσπρων τετραγώνων, με αποτέλεσμα να μπορούν, μέσω της ανίχνευσης των περιγραμμάτων στην κατωφλιωμένη δυαδική εικόνα, να εντοπιστούν τα μαύρα τετράγωνα, τα οποία πλέον έχουν απομακρυνθεί

το ένα από το άλλο. Μάλιστα, στο πλαίσιο της βαθμονόμησης επιλέγεται να διεξαχθεί η αντίστροφη κατωφλίωση, ώστε τα μαύρα τετράγωνα να απεικονίζονται ως άσπρα (βλ. κάτω - δεξιά στην Εικόνα 3-15).

3.5.2.3 Ανίχνευση των εσωτερικών γωνιών της σκακιέρας σε κάθε εικόνα

Εφόσον, με βάση την προηγούμενη διαδικασία, η σκακιέρα εντοπίστηκε στη συγκεκριμένη εικόνα, η επεξεργασία της τελευταίας συνεχίζεται για την ανίχνευση των εσωτερικών γωνιών της σκακιέρας. Πρώτο βήμα μετά τον εντοπισμό της σκακιέρας είναι η μετατροπή της ασπρόμαυρης ισοσταθμισμένης εικόνας σε δυαδική εικόνα τιμών 0 για το μαύρο και 255 για το άσπρο και η εν συνεχεία διαστολή της, ώστε να διαχωριστούν οι εσωτερικές γωνίες της σκακιέρας και να εντοπιστούν τα μαύρα τετράγωνα. Κατόπιν, ακολουθεί η ανίχνευση των κορυφών των τετραγώνων, με δεξιόστροφη φορά. Προκειμένου να επιτευχθεί η τελευταία, εντοπίζονται τα περιγράμματα με βάση τον αλγόριθμο των Suzuki και Abe [137], όπως παρουσιάστηκε στην προηγούμενη παράγραφο. Ειδικότερα, για κάθε περίγραμμα που εντοπίζεται, ελέγχεται αν η περίμετρός του είναι μικρότερη από ένα όριο, οπότε και απορρίπτεται. Διαφορετικά, αν δεν έχει πολύ μικρή περίμετρο, γίνεται δεκτό ως πιθανό όριο τετραγώνου της σκακιέρας και ακολουθεί η απλοποίηση της καμπύλης με τον αλγόριθμο των Douglas και Peucker. Κατόπιν, το περίγραμμα απορρίπτεται αν δεν είναι τετράπλευρο, ενώ αν είναι τετράπλευρο γίνεται δεκτό μόνο αν το σχήμα του προσεγγίζει το τετράγωνο. Η διαδικασία αυτή πραγματοποιείται για κάθε περίγραμμα το οποίο εντοπίζεται από τον αλγόριθμο των Suzuki και Abe και τελικό αποτέλεσμά της είναι ο αριθμός των μαύρων τετραγώνων που εντοπίστηκαν και οι συντεταγμένες pixel (x,y) των τεσσάρων κορυφών του καθενός.

Επόμενο βήμα είναι η εύρεση των γειτονικών κορυφών των τετραγώνων που εντοπίστηκαν. Συγκεκριμένα, για κάθε τετράγωνο επιλέγονται οι κορυφές του, οι οποίες είναι κοντά στις κορυφές άλλων τετραγώνων, γεγονός το οποίο συμβαίνει λόγω της απομάκρυνσης των κορυφών εξαιτίας της διαστολής. Στο πλαίσιο αυτό, για κάθε κορυφή κάθε τετραγώνου εντοπίζεται η πλησιέστερη κορυφή ανάμεσα στις κορυφές όλων των υπολοίπων τετραγώνων. Αν η απόσταση της πλησιέστερης κορυφής από τη συγκεκριμένη κορυφή του τετραγώνου είναι μικρότερη από ένα κατώφλι και εφόσον οι αντίστοιχες πλευρές είναι συμβατές, δηλαδή το μήκος τους δεν διαφέρει περισσότερο από μία καθορισμένη τιμή, τότε θεωρείται ότι η πλησιέστερη κορυφή της συγκεκριμένης κορυφής του τετραγώνου είναι η ζητούμενη αντίστοιχη κορυφή του. Κατόπιν, γίνεται μία σειρά άλλων ελέγχων, όπως π.χ. αν κάποια άλλη κορυφή του συγκεκριμένου τετραγώνου είναι πιο κοντά στην κορυφή που βρέθηκε ως πλησιέστερη, αναγκαίος στην περίπτωση των μικρών τετραγώνων, αν όλες οι κορυφές του τετραγώνου είναι γείτονες διαφορετικών τετραγώνων κ.ά., ώστε τελικά από την παραπάνω διαδικασία να εντοπιστούν οι γειτονικές κορυφές των μαύρων τετραγώνων, οι οποίες στην αρχική εικόνα, πριν τη διαστολή, ταυτίζονται. Ακολούθως, υπολογίζονται οι

συντεταγμένες pixel (x,y) των κοινών κορυφών των μαύρων τετραγώνων στην αρχική εικόνα (δηλαδή οι συντεταγμένες pixel των εσωτερικών γωνιών της σκακιέρας) ως ο μέσος όρος των αντίστοιχων συντεταγμένων x και y των γειτονικών κορυφών.

Στη συνέχεια, εντοπίζονται οι ομάδες συνδεδεμένων τετραγώνων και υπολογίζεται ο τελικός αριθμός τους, ενώ ακολουθεί ο καθορισμός της σειράς των κορυφών τους, με την ταυτόχρονη διαγραφή ή πρόσθεση κάποιων κορυφών. Έτσι, οι κορυφές των τετραγώνων λαμβάνουν τους κωδικούς αριθμούς 0, 1, 2 και 3, ξεκινώντας από την επάνω αριστερή γωνία και συνεχίζοντας δεξιόστροφα. Ο όρος «επάνω αριστερή γωνία» είναι σχετικός και εξαρτάται από την αρχική διάταξη του πρώτου τετραγώνου που εντοπίστηκε, ενώ οι κορυφές των υπόλοιπων τετραγώνων αριθμούνται με τον ίδιο τρόπο. Αν με την παραπάνω διαδικασία εντοπίστηκαν περισσότερα συνδεδεμένα τετράγωνα από τα αναμενόμενα, με βάση τον αριθμό των εσωτερικών γωνιών της σκακιέρας που έχει δοθεί ως αρχικό δεδομένο για τη βαθμονόμηση, απορρίπτονται εκείνα που θεωρείται ότι δεν ανήκουν στο σχέδιο της σκακιέρας. Τελικά, οι συντεταγμένες pixel των εσωτερικών γωνιών της σκακιέρας αποθηκεύονται με την κατάλληλη σειρά, ώστε σε κάθε εικόνα οι συντεταγμένες που ανιχνεύτηκαν να αντιστοιχούν στις ίδιες εσωτερικές γωνίες της σκακιέρας.

Αν με βάση την παραπάνω διαδικασία θεωρηθεί ότι η σκακιέρα δεν εντοπίστηκε, η διαδικασία αυτή, από την κατωφλίωση και τον υπολογισμό της εικόνας της διαστολής, επαναλαμβάνεται με επιπρόσθετη διαστολή της εικόνας, ώστε να εντοπιστεί η σκακιέρα ή ο αριθμός των επαναλήψεων να φτάσει έναν προκαθορισμένο αριθμό. Αν η σκακιέρα εντοπιστεί, ελέγχεται αν κάθε θεωρούμενη οριζόντια και κάθετη γραμμή της σκακιέρας, που ορίζεται από τις συντεταγμένες των εσωτερικών γωνιών που ανιχνεύτηκαν, είναι μία «μονότονη καμπύλη». Για τη διεξαγωγή του ελέγχου αυτού, για κάθε γωνία της σκακιέρας, που βρίσκεται μεταξύ των τελικών γωνιών της ίδιας οριζόντιας ή κάθετης γραμμής, ελέγχεται αν η προβολή της στην ευθεία βρίσκεται ανάμεσα από τις προβολές των γειτονικών της γωνιών στην ίδια οριζόντια ή κάθετη γραμμή.

Τελικό βήμα είναι η παροχή υποψηφιδικής ακρίβειας για τη θέση κάθε εσωτερικής γωνίας της σκακιέρας που εντοπίστηκε. Ο εντοπισμός των γωνιών με υποψηφιδική ακρίβεια βασίζεται στην παρατήρηση ότι το διάνυσμα με αρχή την εσωτερική γωνία της σκακιέρας $\mathbf{q}$, δηλαδή την κορυφή ενός τετραγώνου της, προς ένα σημείο $\mathbf{p_i}$ στη γειτονιά του $\mathbf{q}$ είναι κάθετο στην κλίση της εικόνας στο σημείο $\mathbf{p_i}$ (βλ. Εικόνα 3-16). Συνεπώς, το $\mathbf{q}$ υπολογίζεται έτσι ώστε η σχέση (3.70) να ελαχιστοποιηθεί:

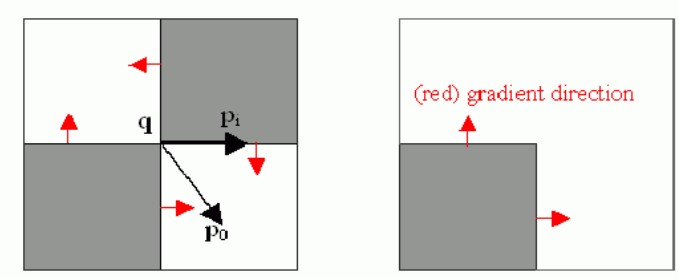
$$\varepsilon_i = \mathbf{DI}_{\mathbf{p_i}}^{T} \cdot \left(\mathbf{q} - \mathbf{p_i}\right) \tag{3.70}$$

Στην παραπάνω σχέση, ως $\mathbf{DI}_{\mathbf{p_i}}$ έχει συμβολιστεί η κλίση της εικόνας στο σημείο $\mathbf{p_i}$ που βρίσκεται στη γειτονιά του $\mathbf{q}$.

Με βάση τη λογική αυτή, θεωρείται το ακόλουθο σύστημα εξισώσεων:

$$\sum_i \left(DI_{p_i} \cdot DI_{p_i}^T \right) \cdot q = \sum_i \left(DI_{p_i} \cdot DI_{p_i}^T \cdot p_i \right) \quad (3.71)$$

από το οποίο προκύπτει το ζητούμενο σημείο **q** που αποτελεί την κορυφή ενός τετραγώνου της σκακιέρας με υποψηφιδική ακρίβεια [142], [143], [144].

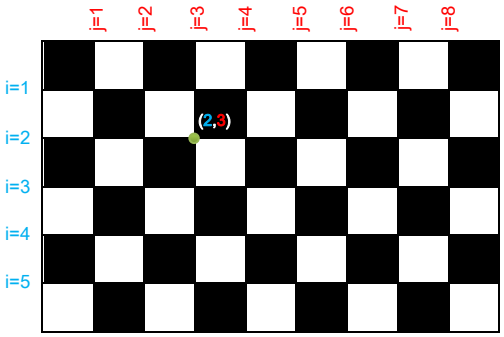


Εικόνα 3-16. Εντοπισμός των κορυφών των τετραγώνων της σκακιέρας με υποψηφιδική ακρίβεια, με βάση την κλίση της εικόνας.

Πηγή: http://docs.opencv.org/modules/imgproc/doc/feature_detection.html

3.5.2.4 Καθορισμός των συντεταγμένων χώρου των εσωτερικών γωνιών της σκακιέρας

Μετά την επεξεργασία όλων των διατιθέμενων εικόνων, με αποτέλεσμα την εξαγωγή ενός πίνακα με τις συντεταγμένες pixel των εσωτερικών γωνιών της σκακιέρας για κάθε εικόνα, ξεκινά η κυρίως διαδικασία της βαθμονόμησης με ταυτόχρονη εύρεση των στοιχείων του εξωτερικού προσανατολισμού κάθε λήψης. Στην παράγραφο αυτή και στις επόμενες που ακολουθούν περιγράφονται τα βήματα που απαιτούνται προκειμένου να ολοκληρωθεί η βαθμονόμηση.



Εικόνα 3-17. Αρίθμηση των εσωτερικών γωνιών της σκακιέρας.

Οι συντεταγμένες χώρου των εσωτερικών γωνιών της σκακιέρας μπορούν να καθοριστούν με βάση τον αριθμό της σειράς και της στήλης (i,j) που αντιστοιχεί σε κάθε γωνία της συγκεκριμένης σκακιέρας (βλ. Εικόνα 3-17) και μία τιμή για το μέγεθος των τετραγώνων της σκακιέρας. Το τελευταίο μπορεί να θεωρηθεί από το χρήστη, ή να επιλεγεί αυθαίρετα από το πρόγραμμα της βαθμονόμησης, καθώς δεν επηρεάζει τα στοιχεία του εσωτερικού προσανατολισμού, αλλά μόνο τις παραμέτρους του εξωτερικού προσανατολισμού.

Έτσι, θεωρώντας ως αρχή του επίγειου συστήματος συντεταγμένων την επάνω αριστερή γωνία της σκακιέρας, οι συντεταγμένες χώρου των εσωτερικών γωνιών της στο ορισθέν σύστημα αναφοράς υπολογίζονται με βάση τη σχέση (3.72), στην οποία ως d συμβολίζεται το θεωρούμενο μήκος της πλευράς των τετραγώνων της σκακιέρας.

$$\begin{aligned} X &= j \cdot d \\ Y &= i \cdot d \\ Z &= 0 \end{aligned} \tag{3.72}$$

3.5.2.5 Εκτίμηση προσεγγιστικών τιμών για τα στοιχεία του εσωτερικού προσανατολισμού

Επόμενο βήμα αποτελεί η θεώρηση αρχικών προσεγγιστικών τιμών για τα στοιχεία του εσωτερικού προσανατολισμού της φωτομηχανής. Η διαδικασία αυτή παραλείπεται στην περίπτωση που δίνονται προσεγγιστικές τιμές από το χρήστη, οπότε και ελέγχεται η ορθότητα των τιμών. Εφόσον δεν δοθούν τιμές για τα στοιχεία του εσωτερικού προσανατολισμού, αυτά αρχικοποιούνται με βάση τη διαδικασία και τις παραδοχές που αναφέρονται στη συνέχεια.

Το πρωτεύον σημείο θεωρείται ότι βρίσκεται στο κέντρο της εικόνας και οι συντελεστές των πολυωνύμων της ακτινικής διαστροφής και της ασύμμετρης διαστροφής εκκεντρότητας τίθενται μηδενικοί. Απομένει η θεώρηση αρχικής τιμής για τη σταθερά της μηχανής, η οποία προσδιορίζεται μέσω των σημείων φυγής[10]. Συγκεκριμένα, αν υπολογιστούν τα σημεία φυγής δύο ομάδων παράλληλων ευθειών του τρισδιάστατου χώρου και εφόσον είναι γνωστές οι συντεταγμένες του πρωτεύοντος σημείου και η γωνία μεταξύ των διευθύνσεων των δύο ομάδων παράλληλων ευθειών, η σταθερά της μηχανής μπορεί να προσδιοριστεί. Στην περίπτωση κατά την οποία οι διευθύνσεις των δύο ομάδων ευθειών είναι κάθετες μεταξύ τους, όπως στην περίπτωση της σκακιέρας, η σταθερά της μηχανής προκύπτει από την ακόλουθη σχέση [104]:

$$c^2 = -\left(x_{\infty 1} - x_0\right) \cdot \left(x_{\infty 2} - x_0\right) - \left(y_{\infty 1} - y_0\right) \cdot \left(y_{\infty 2} - y_0\right) \tag{3.73}$$

όπου c η σταθερά της μηχανής, (x_0, y_0) οι εικονοσυντεταγμένες του πρωτεύοντος σημείου και $(x_{\infty 1}, y_{\infty 1})$ και $(x_{\infty 2}, y_{\infty 2})$ οι εικονοσυντεταγμένες των δύο σημείων φυγής των δύο κάθετων διευθύνσεων των οριζόντιων και κάθετων ομάδων ευθειών της σκακιέρας. Συνεπώς, για τον προσδιορισμό αρχικής τιμής για τη σταθερά της μηχανής αρκεί η εύρεση των δύο σημείων φυγής. Μία τέτοια διαδικασία μπορεί να επιτευχθεί μέσω της προσαρμογής ευθειών στους

[10] Όλες οι ευθείες του χώρου που είναι παράλληλες σε κάποια διεύθυνση συγκλίνουν στην εικόνα στο σημείο φυγής της διεύθυνσης αυτής. Το σημείο αυτό ορίζεται ως η τομή της ευθείας της παράλληλης στη διεύθυνση διά του κέντρου προβολής με το επίπεδο της εικόνας και έτσι αποτελεί την προβολή του ιδεατού σημείου τομής των παράλληλων ευθειών στο επίπεδο της εικόνας [107], [114].

κόμβους που διατάσσονται σε δέσμες συγκλινουσών ευθειών. Έτσι, οι παρατηρήσεις σημείων σε συγκλίνουσες γραμμές της εικόνας μπορούν να οδηγήσουν στην εύρεση των γραμμών που προσαρμόζονται στα σημεία αυτά, υπό τον περιορισμό ότι οι τελευταίες συγκλίνουν στο σημείο φυγής, και στην κατ' επέκταση εύρεση του αντίστοιχου σημείου φυγής [145], [146], [147].

Ωστόσο, η θεώρηση προσεγγιστικών τιμών για τα στοιχεία του εσωτερικού προσανατολισμού με τη λογική που παρουσιάστηκε δεν είναι η μόνη λύση. Σύμφωνα με την αυτόματη διαδικασία βαθμονόμησης με χρήση επίπεδης σκακιέρας που προτάθηκε από το Zhengyou Zhang [148], στην οποία άλλωστε στηρίζεται σε μεγάλο βαθμό η διαδικασία βαθμονόμησης που ακολουθείται από την OpenCV, οι προσεγγιστικές τιμές για τα στοιχεία του εσωτερικού προσανατολισμού μπορούν να προκύψουν όπως αναφέρεται στη συνέχεια.

Στην υποενότητα 3.4.2 (βλ. σχέση (3.69)) αποδείχθηκε ότι ο προβολικός μετασχηματισμός μεταξύ των σημείων ενός δισδιάστατου αντικειμένου του χώρου και των αντίστοιχων σημείων στο επίπεδο της εικόνας εκφράζεται από μία σχέση ομογραφίας στο δισδιάστατο προβολικό χώρο:

$$\lambda \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} c_x & 0 & x_0 \\ 0 & c_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & t_X \\ r_{21} & r_{22} & t_Y \\ r_{31} & r_{32} & t_Z \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \quad (3.74)$$

Στη σχέση (3.74) ο συντελεστής skewness τέθηκε μηδενικός (s=0), θεωρώντας ότι οι πλευρές των εικονοστοιχείων είναι κάθετες μεταξύ τους. Η σχέση αυτή ισχύει στην περίπτωση που το δισδιάστατο αντικείμενο βρίσκεται στο επίπεδο Z=0, όπως θεωρείται και στην περίπτωση της επίπεδης σκακιέρας. Συνεπώς, ο πίνακας της ομογραφίας δίνεται από την ακόλουθη σχέση:

$$\begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} = \frac{1}{\lambda} \cdot \begin{bmatrix} c_x & 0 & x_0 \\ 0 & c_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} r_{11} & r_{12} & t_X \\ r_{21} & r_{22} & t_Y \\ r_{31} & r_{32} & t_Z \end{bmatrix} \quad (3.75)$$

Ο πίνακας της ομογραφίας για κάθε εικόνα υπολογίζεται ελαχιστοτετραγωνικά με δεδομένα τις συντεταγμένες χώρου και εικόνας των εσωτερικών γωνιών της σκακιέρας. Η λύση αυτή βελτιώνεται εν συνεχεία μέσω του αλγορίθμου βελτιστοποίησης Levenberg-Marquardt, με κριτήριο την ελαχιστοποίηση του σφάλματος επαναπροβολής (reprojection error), στην περίπτωση που οι αντιστοιχίες των σημείων σε κάθε εικόνα είναι περισσότερες από τέσσερις, που είναι οι ελάχιστες αναγκαίες. Το σφάλμα επαναπροβολής είναι το άθροισμα των τετραγώνων των αποστάσεων μεταξύ των προβολών των εσωτερικών γωνιών της σκακιέρας από το χώρο στο επίπεδο της εκάστοτε εικόνας – με βάση τον υπολογισμένο πίνακα της ομογραφίας – και των αντίστοιχων σημείων που εντοπίστηκαν στο επίπεδο της εικόνας [125], [149].

Ο αλγόριθμος Levenberg-Marquardt αποτελεί μία επαναληπτική διαδικασία που παρέχει αριθμητική λύση στο πρόβλημα ελαχιστοποίησης μίας συνάρτησης πολλών μεταβλητών, γενικά μη γραμμικής. Γενικά, το πρόβλημα που επιλύει ο αλγόριθμος μπορεί να ορισθεί ως εξής: με δεδομένο ένα σύνολο m ζευγών ανεξάρτητων και εξαρτημένων μεταβλητών (x_i, y_i), ο αλγόριθμος βελτιστοποιεί το διάνυσμα των παραμέτρων **β** του μοντέλου έτσι ώστε το άθροισμα των τετραγώνων των αποκλίσεων να ελαχιστοποιηθεί:

$$C(\beta) = \sum_{i=1}^{m} \left[y_i - f\left(x_i, \boldsymbol{\beta}\right) \right]^2 \tag{3.76}$$

Ο αλγόριθμος είναι επαναληπτικός, απαιτεί μία αρχική εκτίμηση του διανύσματος **β** με τις παραμέτρους του μοντέλου και σε κάθε βήμα επανάληψης το διάνυσμα **β** αντικαθίσταται από τη νέα εκτίμηση **β+δ**, ενώ το τελικό διάνυσμα είναι εκείνο το οποίο ελαχιστοποιεί τη συνάρτηση (3.76) [150].

Αν συμβολιστεί ως **K** ο πίνακας με τις εσωτερικές παραμέτρους της μηχανής, και αν τα τρία στοιχεία κάθε στήλης του πίνακα **H** της ομογραφίας και του πίνακα [**R**|**t**] συμβολιστούν μέσω ενός διανύσματος στήλης, η εξίσωση (3.75) λαμβάνει την ακόλουθη μορφή:

$$\begin{bmatrix} \mathbf{h_1} & \mathbf{h_2} & \mathbf{h_3} \end{bmatrix} = \frac{1}{\lambda} \cdot \mathbf{K} \cdot \begin{bmatrix} \mathbf{r_1} & \mathbf{r_2} & \mathbf{t} \end{bmatrix} \tag{3.77}$$

Από την εξίσωση αυτή είναι δυνατός ο προσδιορισμός της στροφής και της μετάθεσης του επίγειου συστήματος συντεταγμένων για να ταυτιστεί με το φωτογραμμετρικό:

$$\mathbf{h_1} = \frac{1}{\lambda} \cdot \mathbf{K} \cdot \mathbf{r_1} \quad \rightarrow \quad \mathbf{r_1} = \lambda \cdot \mathbf{K}^{-1} \cdot \mathbf{h_1} \tag{3.78}$$

$$\mathbf{h_2} = \frac{1}{\lambda} \cdot \mathbf{K} \cdot \mathbf{r_2} \quad \rightarrow \quad \mathbf{r_2} = \lambda \cdot \mathbf{K}^{-1} \cdot \mathbf{h_2} \tag{3.79}$$

$$\mathbf{h_3} = \frac{1}{\lambda} \cdot \mathbf{K} \cdot \mathbf{t} \quad \rightarrow \quad \mathbf{t} = \lambda \cdot \mathbf{K}^{-1} \cdot \mathbf{h_3} \tag{3.80}$$

Τα διανύσματα $\mathbf{r_1}$ και $\mathbf{r_2}$ είναι ορθοκανονικά. Συνεπώς, ισχύουν δύο περιορισμοί:

$$\mathbf{r_1}^{\mathrm{T}} \cdot \mathbf{r_2} = 0 \tag{3.81}$$

$$\left\| \mathbf{r_1} \right\| = \left\| \mathbf{r_2} \right\| \quad \rightarrow \quad \mathbf{r_1}^{\mathrm{T}} \cdot \mathbf{r_1} = \mathbf{r_2}^{\mathrm{T}} \cdot \mathbf{r_2} \tag{3.82}$$

Ο πρώτος περιορισμός, ο οποίος δίνεται από τη σχέση (3.81), με βάση τις σχέσεις (3.78) και (3.79) και την ιδιότητα $(\mathbf{ab})^{\mathrm{T}}=\mathbf{a}^{\mathrm{T}}\mathbf{b}^{\mathrm{T}}$ για δύο τυχαία διανύσματα **a** και **b**, λαμβάνει την ακόλουθη μορφή:

$$\mathbf{h_1}^{\mathrm{T}} \cdot \mathbf{K}^{-\mathrm{T}} \cdot \mathbf{K}^{-1} \cdot \mathbf{h_2} = 0 \tag{3.83}$$

στην οποία ως $\mathbf{K}^{-\mathrm{T}}$ συμβολίζεται ο πίνακας $\left(\mathbf{K}^{-1}\right)^{\mathrm{T}}$.

Αντίστοιχα, ο δεύτερος περιορισμός που ορίστηκε στη σχέση (3.82), με βάση τις σχέσεις (3.78) και (3.79) και την ιδιότητα $(\mathbf{ab})^T=\mathbf{a}^T\mathbf{b}^T$ για δύο τυχαία διανύσματα $\mathbf{a}$ και $\mathbf{b}$, λαμβάνει την ακόλουθη μορφή:

$$\mathbf{h}_1^T \cdot \mathbf{K}^{-T} \cdot \mathbf{K}^{-1} \cdot \mathbf{h}_1 = \mathbf{h}_2^T \cdot \mathbf{K}^{-T} \cdot \mathbf{K}^{-1} \cdot \mathbf{h}_2 \tag{3.84}$$

Οι σχέσεις (3.83) και (3.84) ορίζουν τους δύο βασικούς περιορισμούς στα στοιχεία του εσωτερικού προσανατολισμού με βάση μία ομογραφία. Επειδή η ομογραφία του δισδιάστατου προβολικού χώρου έχει 8 βαθμούς ελευθερίας και τα στοιχεία του εξωτερικού προσανατολισμού, που συνδέονται με τη συγκεκριμένη ομογραφία, είναι 6 (3 παράμετροι για τη στροφή και 3 για την κλίμακα), από την κάθε λήψη λαμβάνονται δύο περιορισμοί για τα στοιχεία του εσωτερικού προσανατολισμού. Συνεπώς, για τη διεξαγωγή της βαθμονόμησης απαιτούνται τουλάχιστον δύο εικόνες που να περιλαμβάνουν τη σκακιέρα, στην οποία προφανώς πρέπει να εντοπιστούν τουλάχιστον τέσσερις γωνίες για τον υπολογισμό του πίνακα της ομογραφίας. Αυτό αποδεικνύεται και ως εξής. Αν ληφθούν n εικόνες, οι άγνωστοι είναι 6n+4, όπου ο αριθμός 4 υποδηλώνει τα στοιχεία του εσωτερικού προσανατολισμού x_0, y_0, c_x και c_y (θεωρούνται μηδενικές οι συντελεστές των πολυωνύμων της διαστροφής). Οι n εικόνες δίνουν 8n στοιχεία (τους συντελεστές των πινάκων της ομογραφίας όλων των εικόνων). Συνεπώς πρέπει να ισχύει:

$$8n \geq 6n+4 \quad \rightarrow \quad 2n \geq 4 \quad \rightarrow \quad n \geq 2 \tag{3.85}$$

Στην περίπτωση που ως άγνωστος στα στοιχεία του εσωτερικού προσανατολισμού θεωρηθεί και ο συντελεστής skewness – αν οι πλευρές των εικονοστοιχείων αποκλίνουν από την καθετότητα – απαιτούνται τουλάχιστον τρεις εικόνες.

Για λόγους απλότητας, συμβολίζεται ως $\mathbf{B}$ ο 3x3 πίνακας $\mathbf{K}^{-T} \cdot \mathbf{K}^{-1}$:

$$\mathbf{B} = \mathbf{K}^{-T} \cdot \mathbf{K}^{-1} = \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \\ B_{31} & B_{32} & B_{33} \end{bmatrix} \quad \rightarrow \quad \mathbf{B} = \begin{bmatrix} \frac{1}{c_x^2} & 0 & -\frac{x_0}{c_x^2} \\ 0 & \frac{1}{c_y^2} & -\frac{y_0}{c_y^2} \\ -\frac{x_0}{c_x^2} & -\frac{y_0}{c_y^2} & \frac{x_0^2}{c_x^2}+\frac{y_0^2}{c_y^2}+1 \end{bmatrix} \tag{3.86}$$

Στο σημείο αυτό αξίζει να σημειωθεί ότι ο πίνακας $\mathbf{B}$ δίνεται από τη σχέση (3.86) στην περίπτωση που ο συντελεστής s (skewness) τίθεται μηδενικός στον πίνακα $\mathbf{K}$, παραδοχή η οποία θεωρήθηκε στην περίπτωση των εφαρμογών που αναπτύχθηκαν για τις ανάγκες της παρούσας διπλωματικής εργασίας. Στη γενικότερη περίπτωση, ο πίνακας $\mathbf{B}$ δίνεται από τη σχέση (3.87):

$$\mathbf{B}_{s\neq 0}=\begin{bmatrix} \frac{1}{c_x^2} & \frac{-s}{c_x^2\cdot c_y} & \frac{s\cdot y_0-x_0\cdot c_y}{c_x^2\cdot c_y} \\ \frac{-s}{c_x^2\cdot c_y} & \frac{s^2}{c_x^2\cdot c_y^2}+\frac{1}{c_y^2} & -\frac{s\cdot\left(s\cdot y_0-x_0\cdot c_y\right)}{c_x^2\cdot c_y^2}-\frac{y_0}{c_y^2} \\ \frac{s\cdot y_0-x_0\cdot c_y}{c_x^2\cdot c_y} & -\frac{s\cdot\left(s\cdot y_0-x_0\cdot c_y\right)}{c_x^2\cdot c_y^2}-\frac{y_0}{c_y^2} & \frac{\left(s\cdot y_0-x_0\cdot c_y\right)^2}{c_x^2\cdot c_y^2}+\frac{y_0^2}{c_y^2}+1 \end{bmatrix} \tag{3.87}$$

Χρησιμοποιώντας τον πίνακα **B**, οι περιορισμοί που ορίζονται από τις σχέσεις (3.83) και (3.84) λαμβάνουν την ακόλουθη μορφή:

$$\mathbf{h_1}^T\cdot\mathbf{B}\cdot\mathbf{h_2}=0 \tag{3.88}$$

$$\mathbf{h_1}^T\cdot\mathbf{B}\cdot\mathbf{h_1}=\mathbf{h_2}^T\cdot\mathbf{B}\cdot\mathbf{h_2} \tag{3.89}$$

Επειδή ο πίνακας **B** είναι συμμετρικός, μπορεί να οριστεί από ένα διάνυσμα έξι διαστάσεων, ως εξής:

$$\mathbf{b}=\begin{bmatrix} B_{11} & B_{12} & B_{22} & B_{13} & B_{23} & B_{33} \end{bmatrix}^T \tag{3.90}$$

για το οποίο ισχύει:

$$B_{12}=0 \tag{3.91}$$

στην περίπτωση που εξετάζεται (s=0).

Οι περιορισμοί (3.88) και (3.89) περιλαμβάνουν τη γενική έκφραση $\mathbf{h_i}^T\cdot\mathbf{B}\cdot\mathbf{h_j}$. Η έκφραση αυτή μπορεί να γραφεί συναρτήσει του διανύσματος **b** ως εξής:

$$\mathbf{h_i}^T\cdot\mathbf{B}\cdot\mathbf{h_j}=\mathbf{v_{ij}}^T\cdot\mathbf{b} \tag{3.92}$$

Το διάνυσμα $\mathbf{v_{ij}}$ της σχέσης (3.92) δίνεται από την ακόλουθη σχέση:

$$\mathbf{v_{ij}}=\begin{bmatrix} h_{i1}\cdot h_{j1} & h_{i1}\cdot h_{j2}+h_{i2}\cdot h_{j1} & h_{i2}\cdot h_{j2} & h_{i3}\cdot h_{j1}+h_{i1}\cdot h_{j3} & h_{i3}\cdot h_{j2}+h_{i2}\cdot h_{j3} & h_{i3}\cdot h_{j3} \end{bmatrix}^T \tag{3.93}$$

Στις παραπάνω σχέσεις ως $\mathbf{h_i}=\begin{bmatrix} \mathbf{h_{i1}} & \mathbf{h_{i2}} & \mathbf{h_{i3}} \end{bmatrix}^T$ έχει συμβολιστεί το διάνυσμα με τα τρία στοιχεία της στήλης i του πίνακα της ομογραφίας.

Συνεπώς, σύμφωνα με τη σχέση (3.92), οι δύο περιορισμοί των σχέσεων (3.88) και (3.89) γράφονται ως εξής:

$$\mathbf{v_{12}}^T\cdot\mathbf{b}=0 \tag{3.94}$$

$$\mathbf{v_{11}}^T\cdot\mathbf{b}=\mathbf{v_{22}}^T\cdot\mathbf{b}\ \rightarrow\ \left(\mathbf{v_{11}}-\mathbf{v_{22}}\right)^T\cdot\mathbf{b}=0 \tag{3.95}$$

Επιπρόσθετα, ο περιορισμός (3.91) για την περίπτωση κατά την οποία s=0 γράφεται ως εξής:

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \cdot \mathbf{b} = 0 \tag{3.96}$$

Με βάση τους περιορισμούς (3.94) και (3.95) λαμβάνεται η ακόλουθη εξίσωση υπό μορφή πινάκων:

$$\begin{bmatrix} \mathbf{v}_{12}^{T} \\ \left(\mathbf{v}_{11} - \mathbf{v}_{22}\right)^{T} \end{bmatrix} \cdot \mathbf{b} = 0 \tag{3.97}$$

Αν ληφθούν n εικόνες που να περιλαμβάνουν την επίπεδη σκακιέρα, μπορούν να γραφούν n τέτοιες εξισώσεις. Τελικά, το σύστημα των εξισώσεων, το οποίο είναι γραμμικό, μπορεί να γραφεί ως εξής:

$$\mathbf{V} \cdot \mathbf{b} = 0 \tag{3.98}$$

όπου $\mathbf{V}$ είναι ένας πίνακας διαστάσεων 2n x 6. Στην περίπτωση που s=0, ισχύει επιπρόσθετα η εξίσωση (3.96).

Από τη στιγμή που το διάνυσμα $\mathbf{b}$ έχει υπολογιστεί, τα στοιχεία του εσωτερικού προσανατολισμού μπορούν να προσδιοριστούν κατά μοναδικό τρόπο από τη σχέση (3.99), η οποία προκύπτει από τη σχέση (3.86) και τη θεώρηση του παράγοντα κλίμακας λ, ο οποίος απαλείφθηκε και στο σημείο αυτό πρέπει να αποκατασταθεί:

$$\begin{aligned} c_x &= \sqrt{\frac{\lambda}{B_{11}}} \\ c_y &= \sqrt{\frac{\lambda}{B_{22}}} \\ x_0 &= -\frac{B_{13} \cdot c_x^2}{\lambda} \\ y_0 &= -\frac{B_{23}}{B_{22}} \\ \text{όπου} \quad \lambda &= B_{33} - \frac{B_{13}^2 - c_y \cdot B_{11} \cdot B_{23}}{B_{11}} \end{aligned} \tag{3.99}$$

Στη γενική περίπτωση κατά την οποία $s \neq 0$, η σχέση που δίνει τις παραμέτρους του εσωτερικού προσανατολισμού από τα στοιχεία του διανύσματος $\mathbf{b}$ είναι οι ακόλουθη:

$$
\begin{aligned}
c_x &= \sqrt{\frac{\lambda}{B_{11}}} \\
c_y &= \sqrt{\frac{\lambda \cdot B_{11}}{B_{11} \cdot B_{22} - B_{12}^2}} \\
y_0 &= \frac{B_{12} \cdot B_{13} - B_{11} \cdot B_{23}}{B_{11} \cdot B_{22} - B_{12}^2} \\
x_0 &= \frac{s \cdot y_0}{c_x} - \frac{B_{13} \cdot c_x^2}{\lambda} \\
s &= -\frac{B_{12} \cdot c_x^2 \cdot c_y}{\lambda} \\
\text{όπου} \quad \lambda &= B_{33} - \frac{B_{13}^2 + c_y \cdot \left(B_{12} \cdot B_{13} - B_{11} \cdot B_{23} \right)}{B_{11}}
\end{aligned}
\tag{3.100}
$$

3.5.2.6 Υπολογισμός αρχικών τιμών για τα στοιχεία του εξωτερικού προσανατολισμού κάθε λήψης

Αφού υπολογιστούν τα προσεγγιστικά στοιχεία του εσωτερικού προσανατολισμού, είναι δυνατή η δημιουργία του πίνακα **K**. Συνεπώς, από τις σχέσεις (3.78), (3.79) και (3.80) μπορούν να υπολογιστούν τα στοιχεία του εξωτερικού προσανατολισμού $\mathbf{r_1}$, $\mathbf{r_2}$, **t**. Απομένει ο υπολογισμός του διανύσματος $\mathbf{r_3}$. Ο πίνακας στροφής **R** είναι ορθοκανονικός. Αυτό σημαίνει ότι οι τρεις στήλες του, δηλαδή τα διανύσματα $\mathbf{r_1}$, $\mathbf{r_2}$ και $\mathbf{r_3}$ είναι ορθοκανονικά διανύσματα, με αποτέλεσμα το διάνυσμα $\mathbf{r_3}$ να είναι κάθετο στο επίπεδο των $\mathbf{r_1}$ και $\mathbf{r_2}$. Επομένως, υπολογίζεται από το εξωτερικό γινόμενο των τελευταίων:

$$\mathbf{r_3} = \mathbf{r_1} \times \mathbf{r_2} \tag{3.101}$$

Η κλίμακα λ στις σχέσεις (3.78), (3.79) και (3.80) υπολογίζεται ως εξής:

$$\lambda = \frac{1}{\left\| \mathbf{K}^{-1} \cdot \mathbf{h_1} \right\|} = \frac{1}{\left\| \mathbf{K}^{-1} \cdot \mathbf{h_2} \right\|} \tag{3.102}$$

Στο σημείο αυτό πρέπει να σημειωθεί μία σημαντική λεπτομέρεια. Αν δεν δίνονται προσεγγιστικές τιμές για τα στοιχεία του εσωτερικού προσανατολισμού της φωτομηχανής, ο πίνακας της ομογραφίας υπολογίζεται όπως αναφέρθηκε παραπάνω με βάση τις εσωτερικές γωνίες της σκακιέρας στην εικόνα και τα αντίστοιχα σημεία του χώρου, καθώς έχουν θεωρηθεί μηδενικές οι τιμές της διαστροφής. Ωστόσο, στην περίπτωση που δίνονται προσεγγιστικές τιμές για τη διαστροφή των φακών, αρχικά υπολογίζονται οι συντεταγμένες εικόνας των ιδανικών θέσεων των εσωτερικών γωνιών της σκακιέρας (x_{ideal}, y_{ideal}), με βάση τις παρατηρούμενες συντεταγμένες τους ($x_{dist.}$, $y_{dist.}$) και τους συντελεστές των πολυωνύμων

της διαστροφής των φακών, σύμφωνα με τη σχέση (3.103), και στη συνέχεια υπολογίζεται ο πίνακας της ομογραφίας μεταξύ των ιδανικών θέσεων των σημείων αυτών στην εικόνα και των αντίστοιχων σημείων του χώρου. Η διαδικασία αυτή γίνεται επειδή η σχέση (3.74), από την οποία προκύπτουν τα στοιχεία του εξωτερικού προσανατολισμού με βάση τον πίνακα της ομογραφίας, δεν λαμβάνει υπ'όψιν τη διαστροφή των φακών.

$$\begin{bmatrix} x_{ideal} \\ y_{ideal} \end{bmatrix} = \left(1 + k_1 \cdot r^2 + k_2 \cdot r^4 + k_3 \cdot r^6\right) \cdot \begin{bmatrix} x_{dist.} \\ y_{dist.} \end{bmatrix} + \begin{bmatrix} 2 \cdot p_1 \cdot x_{dist.} \cdot y_{dist} + p_2 \cdot \left(r^2 + 2 \cdot x_{dist.}^2\right) \\ p_1 \cdot \left(r^2 + 2 \cdot y_{dist.}^2\right) + 2 \cdot p_2 \cdot x_{dist.} \cdot y_{dist.} \end{bmatrix} \quad (3.103)$$

Ωστόσο, λόγω «θορύβου» στα δεδομένα, ο πίνακας $\mathbf{R} = \begin{bmatrix} \mathbf{r_1} & \mathbf{r_2} & \mathbf{r_3} \end{bmatrix}$ που υπολογίζεται με βάση την παραπάνω διαδικασία δεν πληροί γενικά τις ιδιότητες ενός πίνακα στροφής και έτσι μπορεί να μην τηρείται η ορθογωνικότητα, σύμφωνα με την οποία: $\mathbf{R} \cdot \mathbf{R}^T = \mathbf{R}^T \cdot \mathbf{R} = \mathbf{I}$. Για το λόγο αυτό, απαιτείται ο υπολογισμός του καλύτερου δυνατού πίνακα στροφής, που μπορεί να προσεγγίσει τον υπολογισμένο πίνακα στροφής $\mathbf{R}$ [148]. Η λύση στο πρόβλημα αυτό προέρχεται από την ανάλυση ιδιαζουσών τιμών (Singular Value Decomposition - SVD) του υπολογισμένου πίνακα $\mathbf{R}$. Προτού παρουσιαστεί η εύρεση του καλύτερου δυνατού πίνακα στροφής, μέσω ανάλυσης ιδιαζουσών τιμών, κρίνεται σκόπιμη η σύντομη παρουσίαση του προβλήματος αυτού.

Κάθε πίνακας $\mathbf{A}$ διαστάσεων m x n, του οποίου ο αριθμός των γραμμών m είναι μεγαλύτερος ή ίσος από τον αριθμό των στηλών n, μπορεί να αναλυθεί σε γινόμενο τριών πινάκων, ως εξής:

$$\mathbf{A} = \mathbf{U} \cdot \mathbf{W} \cdot \mathbf{V}^T \quad (3.104)$$

Στη σχέση (3.104) ο πίνακας $\mathbf{U}$ είναι ένας ορθογώνιος πίνακας διαστάσεων m x n, για τον οποίο ισχύει:

$$\mathbf{U}^T \cdot \mathbf{U} = \mathbf{I_n} \quad (3.105)$$

Ο πίνακας $\mathbf{V}$ είναι ένας ορθογώνιος πίνακας διαστάσεων n x n, για τον οποίο ισχύει:

$$\mathbf{V}^T \cdot \mathbf{V} = \mathbf{V} \cdot \mathbf{V}^T = \mathbf{I_n} \quad (3.106)$$

Τέλος, ο πίνακας $\mathbf{W}$ είναι ένας διαγώνιος πίνακας διαστάσεων n x n, του οποίου τα διαγώνια στοιχεία είναι θετικά ή μηδενικά και αποτελούν τις ιδιάζουσες τιμές του πίνακα $\mathbf{A}$ [104].

Σύμφωνα με τα παραπάνω, ο πίνακας $\mathbf{R}$ που έχει υπολογιστεί, μπορεί να αναλυθεί στους τρεις αυτούς πίνακες ως εξής:

$$\mathbf{R} = \mathbf{U} \cdot \mathbf{W} \cdot \mathbf{V}^T \quad (3.107)$$

Ωστόσο, επειδή ο πίνακας $\mathbf{R}$ πρέπει να είναι ορθοκανονικός, ο πίνακας $\mathbf{W}$ πρέπει να είναι ο μοναδιαίος πίνακας 3 x 3. Συνεπώς, ο πίνακας $\mathbf{R}$ μπορεί να εξαναγκαστεί να γίνει πίνακας στροφής, με την ανάλυση ιδιαζουσών τιμών του με βάση τη σχέση (3.107) και

χρησιμοποιώντας τους πίνακες **U** και **V** για τον υπολογισμό του ορθοκανονικού πίνακα στροφής ως εξής [148]:

$$\mathbf{R} = \mathbf{U} \cdot \mathbf{I} \cdot \mathbf{V}^{\mathrm{T}} = \mathbf{U} \cdot \mathbf{V}^{\mathrm{T}} \tag{3.108}$$

Ωστόσο, ο πίνακας στροφής **R** αποτελείται από εννέα στοιχεία, ενώ η στροφή στο χώρο ορίζεται από τρία μόνο στοιχεία, που αποτελούν τις γωνίες στροφής γύρω από τους τρεις άξονες. Μία διαφορετική αναπαράσταση της στροφής είναι η έκφρασή της μέσω ενός διανύσματος $\mathbf{r} = \begin{bmatrix} r_1 & r_2 & r_3 \end{bmatrix}^{\mathrm{T}}$ που είναι παράλληλο στον άξονα στροφής και το μέτρο του είναι ίσο με την αριστερόστροφη γωνία στροφής [104], [127], [151], [152]. Η αναπαράσταση αυτή είναι όχι μόνο πιο εύκολα κατανοητή από τους ανθρώπους αλλά και βολική για διαδικασίες αριθμητικής βελτιστοποίησης, καθώς είναι πιο αποδοτική η αντιμετώπιση των τριών στοιχείων του διανύσματος αυτού, σε σχέση με τα εννέα στοιχεία του πίνακα στροφής. Στο πλαίσιο αυτό, ο πίνακας στροφής μετατρέπεται στο διάνυσμα στροφής **r**.

Αν συμβολιστεί ως θ η γωνία στροφής, τότε $\theta = \|\mathbf{r}\|$. Έστω $\mathbf{r}' = \begin{bmatrix} r_x & r_y & r_z \end{bmatrix}^{\mathrm{T}}$ το μοναδιαίο διάνυσμα που υποδηλώνει τον άξονα στροφής. Τότε ισχύει:

$$\mathbf{r}' = \frac{1}{\theta} \cdot \mathbf{r} \quad \rightarrow \quad \mathbf{r} = \theta \cdot \mathbf{r}' \quad \rightarrow \quad \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} = \theta \cdot \begin{bmatrix} r_x \\ r_y \\ r_z \end{bmatrix} \quad \text{με } \theta = \|\mathbf{r}\| \tag{3.109}$$

Η γωνία θ και το μοναδιαίο διάνυσμα **r′** μπορούν να υπολογιστούν από την ακόλουθη σχέση, με δεδομένο τον πίνακα στροφής **R**:

$$\sin\theta \cdot \begin{bmatrix} 0 & -r_z & r_y \\ r_z & 0 & -r_x \\ -r_y & r_x & 0 \end{bmatrix} = \frac{\mathbf{R} - \mathbf{R}^{\mathrm{T}}}{2} \tag{3.110}$$

ενώ στη συνέχεια το διάνυσμα $\mathbf{r} = \begin{bmatrix} r_1 & r_2 & r_3 \end{bmatrix}^{\mathrm{T}}$ που εμπεριέχει πληροφορίες για τον άξονα στροφής και τη γωνία γύρω από αυτόν μπορεί να υπολογιστεί από τη σχέση (3.109).

Αντίστοιχα, για την εξαγωγή του πίνακα στροφής **R** εκ του διανύσματος **r,** αυτό διαιρείται με τη νόρμα του και προκύπτει το κανονικοποιημένο διάνυσμα **r′**, όπως φαίνεται στη συνέχεια:

$$\theta = \|\mathbf{r}\|. \quad \mathbf{r}' = \frac{1}{\theta} \cdot \mathbf{r} \quad \rightarrow \quad \begin{bmatrix} r_x \\ r_y \\ r_z \end{bmatrix} = \frac{1}{\theta} \cdot \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} \tag{3.111}$$

Ο πίνακας στροφής **R** δίνεται από την ακόλουθη σχέση:

$$\mathbf{R} = \cos\theta \cdot \mathbf{I} + \left(1 - \cos\theta\right) \cdot \mathbf{r}' \cdot \mathbf{r}'^{\mathrm{T}} + \sin\theta \cdot \begin{bmatrix} 0 & -r_z & r_y \\ r_z & 0 & -r_x \\ -r_y & r_x & 0 \end{bmatrix} \tag{3.112}$$

Για την εξαγωγή καλύτερων προσεγγιστικών τιμών για τα στοιχεία του εξωτερικού προσανατολισμού, τα αντίστοιχα σημεία της εικόνας και του χώρου, τα προσεγγιστικά στοιχεία του εσωτερικού προσανατολισμού και τα προσεγγιστικά στοιχεία του εξωτερικού προσανατολισμού πρέπει να εξεταστούν σε μία ενιαία λύση [148]. Έτσι, μέσω της μεθόδου βελτιστοποίησης Levenberg-Marquardt, υπολογίζονται καλύτερες τιμές για τις προηγούμενες προσδιορισθείσες τιμές του εξωτερικού προσανατολισμού με κριτήριο την ελαχιστοποίηση του σφάλματος επαναπροβολής. Πρόκειται για το άθροισμα των τετραγώνων των αποστάσεων μεταξύ των προβολών των εσωτερικών γωνιών της σκακιέρας από το χώρο στο επίπεδο της εκάστοτε εικόνας – με βάση τις ορισθείσες επίγειες συντεταγμένες τους, τα στοιχεία του εσωτερικού προσανατολισμού της μηχανής και τα στοιχεία του εξωτερικού προσανατολισμού της αντίστοιχης εικόνας – και των αντίστοιχων σημείων που εντοπίστηκαν στο επίπεδο της εικόνας.

3.5.2.7 Υπολογισμός αρχικών τιμών για τους συντελεστές των πολυωνύμων της διαστροφής των φακών

Έως το σημείο αυτό, δεν έχει γίνει λόγος για υπολογισμό προσεγγιστικών συντελεστών των πολυωνύμων της διαστροφής των φακών, η οποία θεωρήθηκε μηδενική στην περίπτωση που δεν δίνονται αρχικές τιμές από το χρήστη. Ωστόσο, οι πραγματικοί φακοί συνήθως παρουσιάζουν διαστροφή, κυρίως ακτινική. Επομένως, αντί για τη θεώρηση μηδενικών τιμών για τη διαστροφή αν δεν δίνονται προσεγγιστικές τιμές από το χρήστη, μπορεί να μεσολαβήσει το βήμα αυτό για την εύρεση προσεγγιστικών τιμών για τους συντελεστές των πολυωνύμων της διαστροφής.

Συγκεκριμένα, τα σημεία της εικόνας βρίσκονται σε θέσεις οι οποίες παρεκκλίνουν από τις θέσεις που υπολογίζονται με βάση το μοντέλο της κεντρικής προβολής, λόγω της διαστροφής. Οι αρχικές τιμές για τα στοιχεία του εσωτερικού και του εξωτερικού προσανατολισμού, που υπολογίζονται με βάση την παραπάνω διαδικασία, σε συνδυασμό με τις συντεταγμένες χώρου των εσωτερικών γωνιών της σκακιέρας μπορούν να χρησιμοποιηθούν για τον υπολογισμό αρχικών τιμών για τις ιδανικές θέσεις των σημείων αυτών στο επίπεδο της εικόνας. Στη συνέχεια, οι παρατηρούμενες συντεταγμένες ($x_{dist.}$, $y_{dist.}$) όλων των σημείων σε συνδυασμό με τις ιδανικές τους συντεταγμένες (x_{ideal}, y_{ideal}) μπορούν να χρησιμοποιηθούν για τον υπολογισμό αρχικών τιμών για τους συντελεστές των πολυωνύμων της διαστροφής, σύμφωνα με τη σχέση (3.103) μέσω ελαχιστοτετραγωνικής επίλυσης [148], [151], [153].

Μολονότι το βήμα αυτό προτείνεται από το Zhang [148], μπορεί να παραλειφθεί και οι αρχικές τιμές για τους συντελεστές των πολυωνύμων της διαστροφής να θεωρηθούν μηδενικές.

3.5.2.8 Ενιαία βελτιστοποίηση των στοιχείων του εσωτερικού και εξωτερικού προσανατολισμού

Έχοντας υπολογίσει αρχικές τιμές για τα στοιχεία του εσωτερικού προσανατολισμού της φωτομηχανής και τα στοιχεία του εξωτερικού προσανατολισμού κάθε λήψης, είναι δυνατή η βελτιστοποίηση των στοιχείων αυτών μέσω ενός αλγορίθμου βελτιστοποίησης, όπως ο Levenberg-Marquardt.

Έστω ότι δίνονται n εικόνες της επίπεδης σκακιέρας και m εσωτερικές γωνίες της σκακιέρας. Η εύρεση των βέλτιστων τιμών για τα στοιχεία του εσωτερικού και εξωτερικού προσανατολισμού μπορεί να προκύψει από την ελαχιστοποίηση της ακόλουθης μη γραμμικής συνάρτησης, που συνιστά το σφάλμα επαναπροβολής [148]:

$$\sum_{i=1}^{n}\sum_{j=1}^{m}\left\|\mathbf{m}_{ij}-\mathbf{m}\left(\mathbf{K},k_0,k_1,k_2,p_1,p_2,\mathbf{r}_i,\mathbf{t}_i,\mathbf{M}_j\right)\right\|^2 \tag{3.113}$$

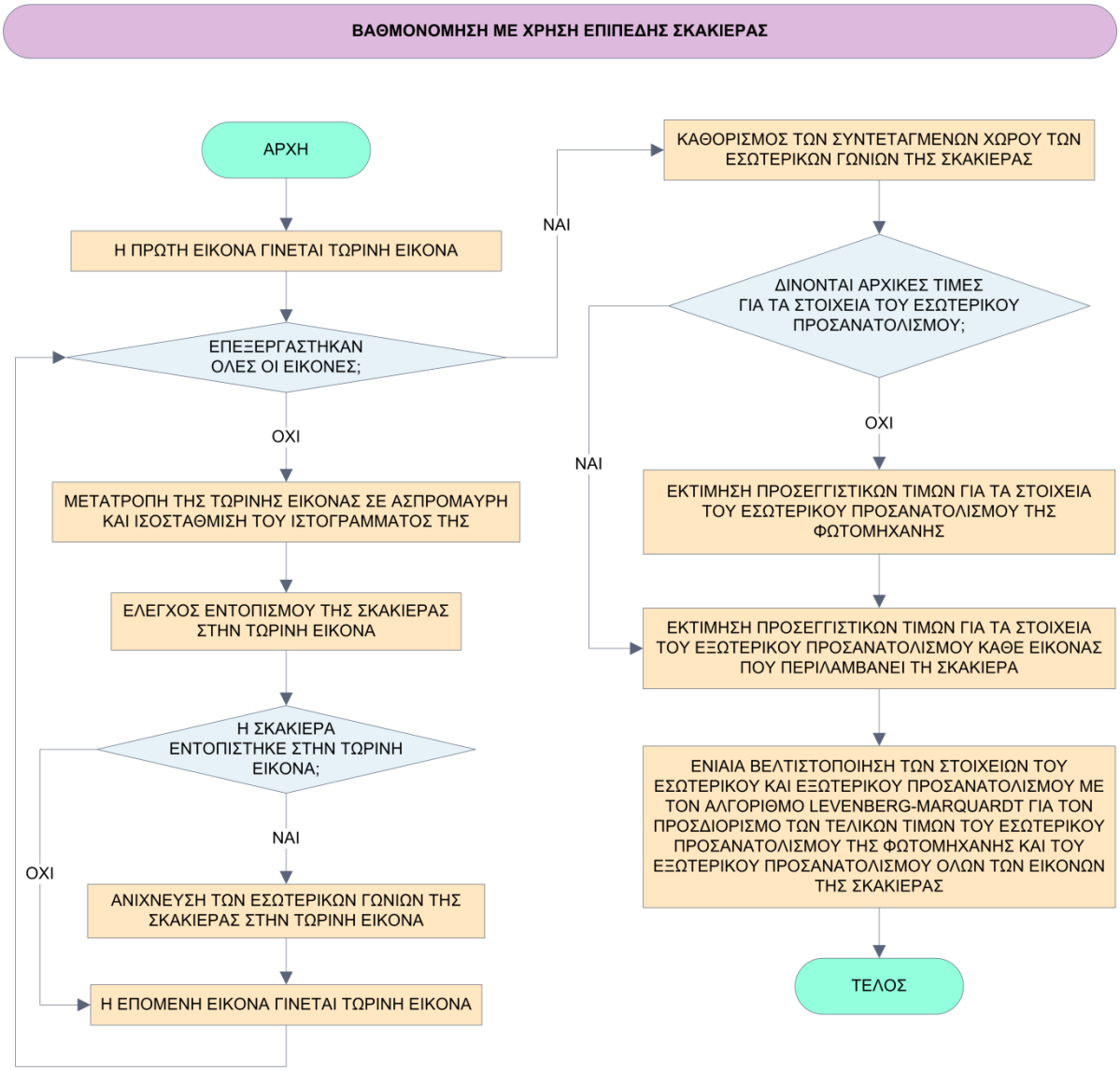
όπου $\mathbf{m}_{ij}$ είναι το παρατηρούμενο σημείο και $\mathbf{m}\left(\mathbf{K},k_0,k_1,k_2,p_1,p_2,\mathbf{r}_i,\mathbf{t}_i,\mathbf{M}_j\right)$ η προβολή του σημείου $\mathbf{M}_j$ στην εικόνα.

Τα αποτελέσματα, λοιπόν, της βαθμονόμησης από την παραπάνω επίλυση είναι ο πίνακας με τις εσωτερικές παραμέτρους της μηχανής (camera matrix) και οι συντελεστές της συμμετρικής ακτινικής διαστροφής k_1, k_2 και k_3, και της ασύμμετρης διαστροφής εκκεντρότητας p_1 και p_2. Εκτός των στοιχείων αυτών προσδιορίζονται και τα στοιχεία που ορίζουν το μετασχηματισμό του επίγειου συστήματος στο φωτογραμμετρικό και τα οποία είναι το διάνυσμα μετάθεσης $\mathbf{t}$ και το διάνυσμα στροφής $\mathbf{r}$.

Ένα απλοποιημένο διάγραμμα ροής της διαδικασίας που παρουσιάστηκε παρατίθεται στη Διαγραμματική παρουσίαση 3-1.

3.6 Προσδιορισμός των Στοιχείων του Εξωτερικού Προσανατολισμού

Στη φωτογραμμετρία, η θέση του προβολικού κέντρου και οι γωνίες κατά τις οποίες πρέπει να στραφεί το επίγειο σύστημα αναφοράς για να γίνει παράλληλο με το φωτογραμμετρικό καλούνται στοιχεία του εξωτερικού προσανατολισμού και προσδιορίζουν πλήρως τη θέση της δέσμης των ακτίνων στο χώρο. Σε φωτογραμμετρικά προβλήματα, η εύρεση των στοιχείων του εξωτερικού προσανατολισμού μίας λήψης γίνεται μέσω της συνθήκης συγγραμμικότητας, με δεδομένα τα στοιχεία του εσωτερικού προσανατολισμού και τις συντεταγμένες χώρου τουλάχιστον τριών σημείων και με μετρήσεις των προβολών τους στην εικόνα.



Διαγραμματική παρουσίαση 3-1. Διαδικασία βαθμονόμησης με χρήση επίπεδης σκακιέρας.

Στην όραση υπολογιστών, το πρόβλημα εύρεσης του εξωτερικού προσανατολισμού αναφέρεται συχνά ως «The Perspective-n-Point (PnP) Problem» και μπορεί να οριστεί ως η εκτίμηση της θέσης και του προσανατολισμού της κάμερας από τις γνωστές τρισδιάστατες συντεταγμένες n σημείων και των προβολών τους στο επίπεδο της εικόνας, με ταυτόχρονη γνώση του πίνακα της μηχανής και των συντελεστών των πολυωνύμων των διαστροφών [154].

Κατά τον προσδιορισμό του εξωτερικού προσανατολισμού, υπολογίζεται το διάνυσμα μετατόπισης **t** και το διάνυσμα στροφής **r**, το οποίο μετέπειτα μπορεί να μετατραπεί στον 3x3 πίνακα στροφής **R**. Ο προσδιορισμός των διανυσμάτων **r** και **t** γίνεται με παρόμοιο τρόπο με τον υπολογισμό τους κατά τη διαδικασία της βαθμονόμησης, με γνωστά πλέον τα στοιχεία του εσωτερικού προσανατολισμού, τις συντεταγμένες χώρου n σημείων του αντικειμένου και τις συντεταγμένες εικόνας των προβολών τους. Επιπρόσθετα, οι συντεταγμένες χώρου

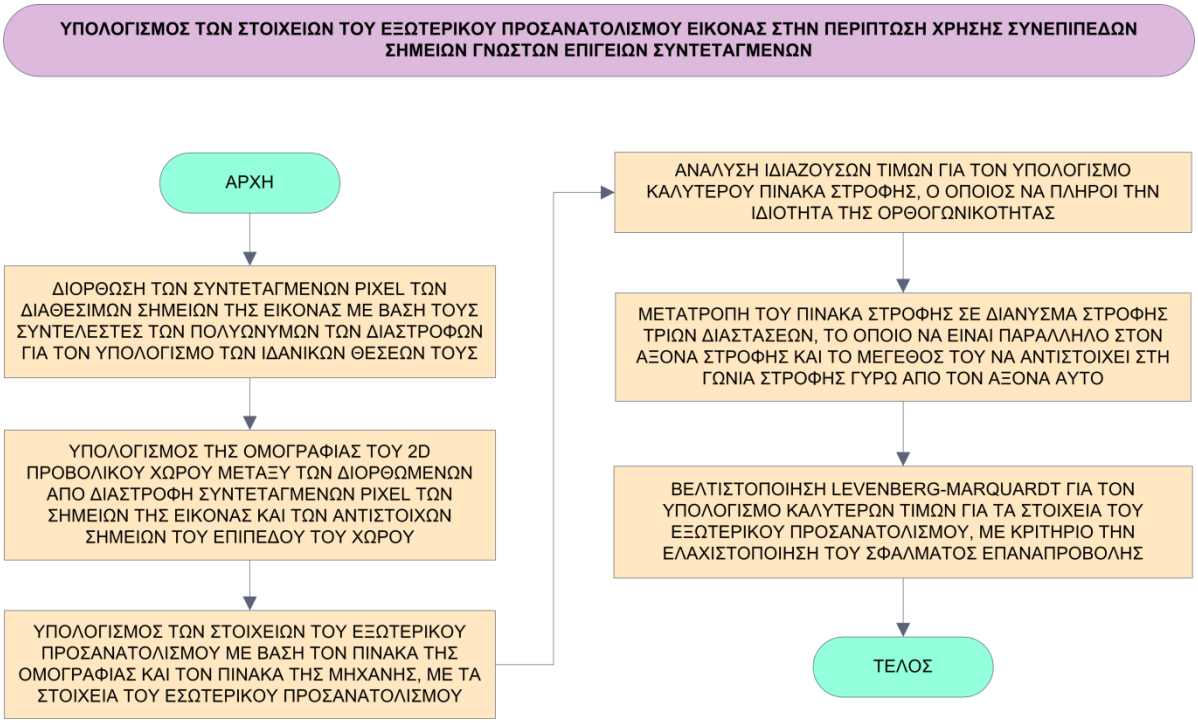
μπορεί να είναι τρισδιάστατες, αφού δεν αναφέρονται – αποκλειστικά – σε επίπεδο αντικείμενο. Στο πλαίσιο της παρούσας εργασίας, χρησιμοποιήθηκαν σημεία ενός επιπέδου του χώρου για την εύρεση του εξωτερικού προσανατολισμού και στην ενότητα αυτή παρουσιάζεται η διαδικασία αυτή, η οποία είναι ανάλογη με εκείνη που παρουσιάστηκε στην παράγραφο 3.5.2.6, με τη διαφορά ότι τα στοιχεία του εσωτερικού προσανατολισμού και τα υπολογιζόμενα στοιχεία του εξωτερικού προσανατολισμού δεν είναι πλέον προσεγγιστικά. Η εν λόγω διαδικασία συνοψίζεται στη συνέχεια (βλ. Διαγραμματική παρουσίαση 3-2).

Αρχικά υπολογίζονται οι συντεταγμένες των ιδανικών θέσεων των σημείων της εικόνας (x_{ideal}, y_{ideal}) για τα οποία είναι γνωστές οι επίγειες συντεταγμένες, με βάση τις παρατηρούμενες συντεταγμένες τους στην εικόνα ($x_{dist.}$, $y_{dist.}$) και τους συντελεστές των πολυωνύμων της διαστροφής των φακών, σύμφωνα με τη σχέση (3.103). Κατόπιν, υπολογίζεται ο πίνακας της ομογραφίας μεταξύ των ιδανικών θέσεων των σημείων αυτών στην εικόνα και των αντίστοιχων σημείων του χώρου. Ακολουθεί ο προσδιορισμός του εξωτερικού προσανατολισμού, δηλαδή του πίνακα στροφής $\mathbf{R}$ και του διανύσματος μετάθεσης $\mathbf{t}$. Το τελευταίο υπολογίζεται με βάση τη σχέση (3.80) σε συνδυασμό με τη σχέση (3.102), ενώ οι τρεις στήλες του πίνακα στροφής υπολογίζονται από τις σχέσεις (3.78), (3.79) και (3.101) σε συνδυασμό με τη σχέση (3.102). Ο υπολογισμός τους απαιτεί εκτός από τη γνώση του πίνακα του δισδιάστατου προβολικού μετασχηματισμού, τη γνώση του πίνακα $\mathbf{K}$ με τα στοιχεία του εσωτερικού προσανατολισμού της φωτομηχανής. Στη συνέχεια, μπορεί να μεσολαβήσει ο υπολογισμός καλύτερου πίνακα στροφής από τη σχέση (3.108), έτσι ώστε να πληροί την ιδιότητα της ορθογωνικότητας, αφού προηγηθεί ανάλυση ιδιαζουσών τιμών στον πίνακα $\mathbf{R}$ (σχέση (3.107)).

Ακολούθως, ο πίνακας στροφής $\mathbf{R}$ εκφράζεται μέσω ενός διανύσματος $\mathbf{r}$ τριών διαστάσεων, που είναι παράλληλο στον άξονα στροφής και το μέγεθός του είναι ίσο με τη γωνία στροφής. Η γωνία στροφής θ και το κανονικοποιημένο διάνυσμα $\mathbf{r}'$ του διανύσματος $\mathbf{r}$ μπορούν να υπολογιστούν από τη σχέση (3.110), με δεδομένο τον πίνακα στροφής $\mathbf{R}$, ενώ στη συνέχεια το διάνυσμα $\mathbf{r}$ μπορεί να υπολογιστεί από τη σχέση (3.109). Τέλος, για τον προσδιορισμό καλύτερων τιμών για τα στοιχεία του εξωτερικού προσανατολισμού, τα αντίστοιχα σημεία της εικόνας και του χώρου, τα στοιχεία του εσωτερικού προσανατολισμού και οι παράμετροι του εξωτερικού προσανατολισμού πρέπει να εξεταστούν σε μία ενιαία λύση. Έτσι, μέσω της μεθόδου βελτιστοποίησης Levenberg-Marquardt, υπολογίζονται καλύτερες τιμές για τις προηγούμενες προσδιορισθείσες τιμές του εξωτερικού προσανατολισμού με κριτήριο την ελαχιστοποίηση του σφάλματος επαναπροβολής.

Ωστόσο, στην περίπτωση κατά την οποία για την εύρεση του εξωτερικού προσανατολισμού χρησιμοποιούνται σημεία του χώρου τα οποία δεν βρίσκονται στο ίδιο επίπεδο, η σχέση που συνδέει τα σημεία του χώρου με τα αντίστοιχα εικονοσημεία δεν είναι η ομογραφία του δισδιάστατου προβολικού χώρου, αλλά ο άμεσος γραμμικός μετασχηματισμός (DLT). Η

παρουσίαση της διαδικασίας υπολογισμού των στοιχείων του εξωτερικού προσανατολισμού στην περίπτωση γνωστών επίγειων συντεταγμένων που ορίζουν τρισδιάστατο σημειοσύνολο ξεφεύγει από τους σκοπούς της παρούσας εργασίας.



Διαγραμματική παρουσίαση 3-2. Υπολογισμός των στοιχείων του εξωτερικού προσανατολισμού μίας εικόνας σε περίπτωση που διατίθενται ομολογίες σημείων ενός επίπεδου αντικειμένου.

ΚΕΦΑΛΑΙΟ 4

ΠΡΟΓΡΑΜΜΑΤΙΣΤΙΚΟ ΠΕΡΙΒΑΛΛΟΝ ΕΡΓΑΣΙΑΣ - ΜΕΤΑΣΧΗΜΑΤΙΣΜΟΙ ΣΥΣΤΗΜΑΤΩΝ

Προτού συνδυαστούν όσα αναφέρθηκαν στα κεφάλαια 2 και 3 για την παρουσίαση της βασικής διαδικασίας που ακολουθήθηκε κατά την ανάπτυξη των εφαρμογών επαυξημένης πραγματικότητας και για τη λεπτομερέστερη περιγραφή τους, είναι αναγκαία η αναφορά στο προγραμματιστικό περιβάλλον εργασίας. Ειδικότερα, στο κεφάλαιο αυτό γίνεται μία εισαγωγή στη γλώσσα προγραμματισμού C++, η οποία χρησιμοποιήθηκε για την ανάπτυξη των εφαρμογών, και περιγράφονται τα βασικά στοιχεία ενός περιβάλλοντος ανάπτυξης προγραμμάτων C++, αλλά και το συγκεκριμένο περιβάλλον που χρησιμοποιήθηκε στην παρούσα εργασία. Επίσης, παρουσιάζονται οι βιβλιοθήκες που χρησιμοποιήθηκαν, πέραν της τυπικής βιβλιοθήκης της C++. Αυτές είναι η OpenCV, οι βασικές βιβλιοθήκες της διεπαφής προγραμματισμού εφαρμογών OpenGL (OpenGL Core Library, GLU, freeglut, GLEW) και η βιβλιοθήκη GLM: An Alias Wavefront OBJ file Library. Ακόμη, περιγράφονται κάποια βασικά θεωρητικά στοιχεία για τις OpenCV και OpenGL, που κρίνονται απαραίτητα για την κατανόηση της διαδικασίας που ακολουθείται στις εφαρμογές. Πρόκειται για τα συστήματα συντεταγμένων της κάμερας που χρησιμοποιούνται από την OpenCV και OpenGL και τους μετασχηματισμούς που πραγματοποιούνται από την OpenGL προκειμένου ένα τρισδιάστατο μοντέλο να τοποθετηθεί στη σωστή θέση εντός του παραθύρου της συσκευής επαύξησης της πραγματικότητας.

4.1 Γλώσσα Προγραμματισμού C++

Η C++ είναι μία από τις πιο δημοφιλείς γλώσσες προγραμματισμού γενικού σκοπού. Αποτελεί γλώσσα υψηλού επιπέδου, η οποία όμως διαθέτει και δυνατότητες γλώσσας χαμηλού επιπέδου, και υποστηρίζει τόσο δομημένο όσο και αντικειμενοστραφή προγραμματισμό [155], [156]. Στην ενότητα αυτή παρουσιάζονται συνοπτικά κάποια ιστορικά στοιχεία για την εν λόγω γλώσσα προγραμματισμού, περιγράφονται τα κύρια

χαρακτηριστικά της και αιτιολογείται η επιλογή της ως γλώσσας προγραμματισμού των εφαρμογών που υλοποιήθηκαν στο πλαίσιο της παρούσας εργασίας.

4.1.1 Σύντομα ιστορικά στοιχεία

Η γλώσσα προγραμματισμού C++ αναπτύχθηκε από τον Bjarne Stroustrup το 1979 στα εργαστήρια Bell της εταιρίας AT&T (American Telephone & Telegraph company). Αρχικά ονομάστηκε «C with Classes», ενώ μετονομάστηκε σε C++ το 1983. Αποτελεί μία επέκταση της C, η οποία δημιουργήθηκε από τον Dennis Ritchie στις αρχές της δεκαετίας του 1970, στηρίχτηκε σε δύο προηγούμενες γλώσσες, την BCPL και την B, και έγινε γνωστή ως γλώσσα ανάπτυξης για το λειτουργικό σύστημα UNIX. Η C++ παρέχει διάφορες λειτουργίες που βελτιώνουν τη C, εκ των οποίων σημαντικότερη είναι η δυνατότητα που δίνει για αντικειμενοστραφή προγραμματισμό. Άλλες βελτιώσεις, που συνδέονται με την έννοια του αντικειμενοστραφούς προγραμματισμού, αποτελούν οι κλάσεις, οι εικονικές συναρτήσεις, η υπερφόρτωση τελεστών, η πολλαπλή κληρονομικότητα κ.ά.

Η C++ βασίζεται στη C και σε μία γλώσσα προγραμματισμού προσομοίωσης, την επομονομαζόμενη Simula. Σύμφωνα με το δημιουργό της C++, η εν λόγω γλώσσα σχεδιάστηκε για να παρέχει τις ευκολίες της Simula, σχετικά με την οργάνωση των προγραμμάτων, και την αποτελεσματικότητα και ευελιξία της C για τον προγραμματισμό συστημάτων. Ένας από τους βασικούς σκοπούς του Stroustrup, μέσω της ανάπτυξης της C++, ήταν ο σχεδιασμός μίας γλώσσας που θα διευκόλυνε την ανάπτυξη προγραμμάτων.

Το 1985 ο Stroustrup δημιοσίευσε το πρώτο εγχειρίδιο της C++, με τον τίτλο «The C++ Programming Language», και το ίδιο έτος η C++ έγινε πλέον εμπορικό προϊόν, ενώ το 1989 ανανεώθηκε για να συμπεριλάβει επιπλέον λειτουργίες. Το 1998, η C++ ορίστηκε παγκοσμίως με το πρότυπο ISO/IEC 14882:1998 (γνωστό και ως C++98), ενώ ο ορισμός της ανανεώθηκε το 2003, με το πρότυπο ISO/IEC 14882:2003 (C++03). Τέλος, στα μέσα του 2011 δημιουργήθηκε το νέο της πρότυπο, ονομάστηκε C++11 και αποτελεί την τρέχουσα έκδοση του προτύπου της C++ [155], [157], [158], [159], [160].

4.1.2 Βασικά χαρακτηριστικά της C++

Οι παλαιότερες από τη C++ γλώσσες προγραμματισμού, όπως η FORTAN, η Pascal, η Basic και η C, στηρίζονταν στις αρχές του δομημένου προγραμματισμού, δηλαδή έδιναν σημασία σε ενέργειες και όχι σε πράγματα ή αντικείμενα, δίνοντας έτσι βάρος στη διαδικασία που ακολουθείται για την επίτευξη ενός στόχου. Αντίθετα, η C++ στηρίζεται στις αρχές του αντικειμενοστραφούς προγραμματισμού[11], ο οποίος επικεντρώνεται σε πράγματα ή

[11] Η C++ είναι μία αντικειμενοστραφής γλώσσα αλλά τα προγράμματα C++ μπορούν να ακολουθήσουν και τις αρχές του δομημένου προγραμματισμού, αν το επιθυμεί ο προγραμματιστής.

αντικείμενα, και έτσι οι μονάδες των προγραμμάτων αντιπροσωπεύουν εύκολα οντότητες του πραγματικού κόσμου, ο οποίος είναι πιο κοντά στην έννοια του αντικειμενοστραφούς προγραμματισμού. Ο τελευταίος δίνει έμφαση στα δεδομένα και όχι στον κώδικα, στα «ουσιαστικά» και όχι στα «ρήματα», στα αντικείμενα και όχι στα στάδια που απαιτούνται για την ολοκλήρωση μίας διαδικασίας. Ένα από τα πλεονεκτήματά του είναι η δημιουργία καλύτερα οργανωμένων προγραμμάτων, ευκολότερων στη συντήρηση και τροποποίηση. Άλλο βασικό πλεονέκτημα που αποδεικνύει την ανωτερότητα του αντικειμενοστραφούς σε σχέση με το δομημένο προγραμματισμό αποτελεί η επαναχρησιμοποίηση του λογισμικού, καθώς οι οντότητες λογισμικού, οι οποίες είναι οι κλάσεις, μπορούν να είναι επαναχρησιμοποιήσιμες σε μελλοντικά έργα [157], [161].

Η C++ στηρίζεται σε αντικείμενα, τα οποία εμπεριέχουν δεδομένα και ανταλλάσσουν μεταξύ τους πληροφορίες και μηνύματα για την επίτευξη του στόχου του προγράμματος. Το τελευταίο αποτελείται από κλάσεις, οι οποίες είναι η κεντρική ιδέα του αντικειμενοστραφούς προγραμματισμού και αποτελούν τον κώδικα που ορίζει ένα αντικείμενο. Η κλάση χρησιμοποιείται για την κατασκευή πανομοιότυπων αντιγράφων του ίδιου αντικειμένου και περιλαμβάνει συναρτήσεις, οι οποίες εκτελούν εργασίες, και πεδία, τα οποία υλοποιούν χαρακτηριστικά. Έτσι, ένα αντικείμενο, το οποίο είναι «στιγμιότυπο» μίας κλάσης, περιέχει δεδομένα και διαθέτει μεθόδους για την επεξεργασία των δεδομένων αυτών αλλά και για την επικοινωνία με άλλα αντικείμενα [162], [163].

Επίσης, λόγω του αντικειμενοστραφούς χαρακτήρα της C++, μία κλάση μπορεί να κληρονομήσει συναρτήσεις και πεδία από άλλη κλάση, με αποτέλεσμα η κάθε μία να μπορεί να επεκταθεί σε νέα κλάση. Λόγω της ιδιότητας αυτής, η οποία ονομάζεται κληρονομικότητα, νέες κλάσεις απορρέουν από την απορρόφηση χαρακτηριστικών υπαρκτών κλάσεων και την προσθήκη μοναδικών χαρακτηριστικών σε αυτές [162]. Επιπλέον, οι κλάσεις έχουν την ιδιότητα να αποκρύπτουν τα ιδιωτικά τους δεδομένα από το υπόλοιπο πρόγραμμα και να καθορίζουν ποια δεδομένα – γνωστά ως δημόσια – μπορούν να προσπελαστούν από αυτό. Άλλα χαρακτηριστικά αποτελούν η υπερφόρτωση των συναρτήσεων, κατά την οποία μπορεί να υφίστανται συναρτήσεις με το ίδιο όνομα αλλά διαφορετικά ορίσματα, η έννοια της αφηρημένης κλάσης, που κληρονομείται σε άλλες υποκλάσεις και ορίζει ένα «συμβόλαιο» το οποίο θα πρέπει να ακολουθηθεί από αυτές, κ.ά. [164].

4.1.3 Επιλογή της C++ ως γλώσσας προγραμματισμού για την ανάπτυξη των εφαρμογών

Κύρια κριτήρια για την επιλογή της γλώσσας προγραμματισμού που επρόκειτο να χρησιμοποιηθεί για την πραγματοποίηση των εφαρμογών αποτέλεσαν η ταχύτητα κατά την εκτέλεση των προγραμμάτων, η διαχείριση της μνήμης και οι προϋπάρχουσες βιβλιοθήκες. Συγκεκριμένα, η ταχύτητα και η αποδοτική διαχείριση της μνήμης αποτελούν αναγκαίες

προϋποθέσεις για την ικανοποιητική εκτέλεση των εφαρμογών επαυξημένης πραγματικότητας, οι οποίες, λόγω των διαδικασιών που εκτελούνται για την διεκπεραίωση της επαύξησης του πραγματικού κόσμου, είναι απαιτητικές τόσο σε ταχύτητα όσο και σε μνήμη. Η C++ έχει σημαντικά πλεονεκτήματα σε σχέση με άλλες γλώσσες προγραμματισμού όσον αφορά στην ταχύτητα, τη διαχείριση μνήμης και την αποδοτικότητα σε απαιτητικές εφαρμογές. Επιπλέον, η βιβλιοθήκη OpenCV η οποία επρόκειτο να χρησιμοποιηθεί διαθέτει το κύριο interface της σε C++, ενώ και οι βιβλιοθήκες της OpenGL είναι συμβατές με τη γλώσσα αυτή.

4.2 Περιβάλλον Ανάπτυξης Προγραμμάτων C++

Βασικά στοιχεία ενός συστήματος C++ αποτελούν το περιβάλλον ανάπτυξης, η γλώσσα προγραμματισμού και η τυπική βιβλιοθήκη της C++ [157]. Στην ενότητα αυτή, αρχικά περιγράφονται τα βασικά στοιχεία ενός τυπικού περιβάλλοντος ανάπτυξης προγραμμάτων C++ και στη συνέχεια παρουσιάζεται το περιβάλλον που χρησιμοποιήθηκε για την ανάπτυξη των εφαρμογών στο πλαίσιο της παρούσας διπλωματικής εργασίας.

4.2.1 Βασικά στοιχεία ενός τυπικού περιβάλλοντος ανάπτυξης C++

Ένα τυπικό περιβάλλον ανάπτυξης προγραμμάτων C++ αποτελείται από έναν επεξεργαστή, έναν προεπεξεργαστή, ένα μεταγλωττιστή, ένα πρόγραμμα σύνδεσης και ένα πρόγραμμα φόρτωσης, τα οποία σχετίζονται με τις φάσεις από τις οποίες πρέπει να περάσει το πρόγραμμα προκειμένου να εκτελεστεί, δηλαδή την τροποποίηση, την προεπεξεργασία, τη μεταγλώττιση, τη σύνδεση, τη φόρτωση και τελικά την εκτέλεση (βλ. Διαγραμματική παρουσίαση 4-1).

Η τροποποίηση των αρχείων γίνεται μέσω ενός κειμενογράφου (επεξεργαστή), στον οποίο ο προγραμματιστής μπορεί να γράφει και να διορθώνει τον κώδικα. Τα ονόματα των αρχείων της C++ συνήθως τελειώνουν με τις επεκτάσεις .cpp, .cxx ή .cc και αποτελούν τα αρχεία πηγαίου κώδικα (source files), τα οποία στη συνέχεια πρέπει να προεπεξεργαστούν, να μεταγλωττιστούν και να συνδεθούν. Εκτός από αυτά, υπάρχουν και τα αρχεία επικεφαλίδων (header files), τα οποία δεν μεταγλωττίζονται, ούτε συνδέονται και έχουν συνήθως την κατάληξη .h, .hpp ή .hxx. Γενικά, χρησιμοποιούνται για τη δήλωση του «τι» κάνει μία κλάση, σε αντίθεση με τα αρχεία πηγαίου κώδικα, τα οποία ορίζουν «πώς» το κάνει. Ένα αρχείο επικεφαλίδων πρέπει να συμπεριληφθεί σε ένα άλλο αρχείο πηγαίου κώδικα ή στο αντίστοιχο αρχείο επικεφαλίδας, σε περίπτωση που αυτά χρησιμοποιούν στοιχεία που ορίζονται στο εν λόγω αρχείο επικεφαλίδας, όπως συναρτήσεις, τύπους δεδομένων κ.λπ.

Όταν ο προγραμμαστιστής δώσει την εντολή της μεταγλώττισης του προγράμματος σε γλώσσα μηχανής, εκτελείται αυτόματα ο προεπεξεργαστής, πριν την έναρξη της φάσης της

μεταγλώττισης. Πιθανές ενέργειες του προεπεξεργαστή αποτελούν η συμπερίληψη αρχείων στο μεταγλωττιζόμενο αρχείο, η υπό συνθήκη μεταγλώττιση τμήματος κώδικα ενός προγράμματος, ο ορισμός συμβολικών σταθερών και γενικά η «υπακοή» στις οδηγίες που ξεκινούν με το σύμβολο # και ονομάζονται οδηγίες προεπεξεργαστή.

Στη συνέχεια, καλείται ο μεταγλωττιστής (compiler), ο οποίος δημιουργεί τον αντικειμενικό κώδικα, μεταγλωττίζοντας το κάθε αρχείο πηγαίου κώδικα ξεχωριστά από τα υπόλοιπα, εάν φυσικά υπάρχουν άλλα. Αποτέλεσμα της διαδικασίας αυτής είναι ένα «object» αρχείο (με συνήθη κατάληξη .obj), για το κάθε αρχείο πηγαίου κώδικα.

Επόμενη φάση από την οποία περνάει το πρόγραμμα είναι η φάση της σύνδεσης. Γενικά, ένα πρόγραμμα C++ χρησιμοποιεί συναρτήσεις και δεδομένα που ορίζονται σε βιβλιοθήκες. Το πρόγραμμα σύνδεσης (linker) συνδέει όλα τα αρχεία αντικειμενικού κώδικα, που δημιουργήθηκαν από το μεταγλωττιστή, μεταξύ τους και με τις χρησιμοποιούμενες βιβλιοθήκες και παράγει ένα εκτελέσιμο αρχείο (αρχείο με κατάληξη .exe).

Έπειτα, το πρόγραμμα φόρτωσης τοποθετεί το πρόγραμμα στη μνήμη, διαδικασία απαραίτητη πριν την εκτέλεση του προγράμματος. Επίσης, φορτώνονται και επιπλέον συστατικά από κοινόχρηστες βιβλιοθήκες.

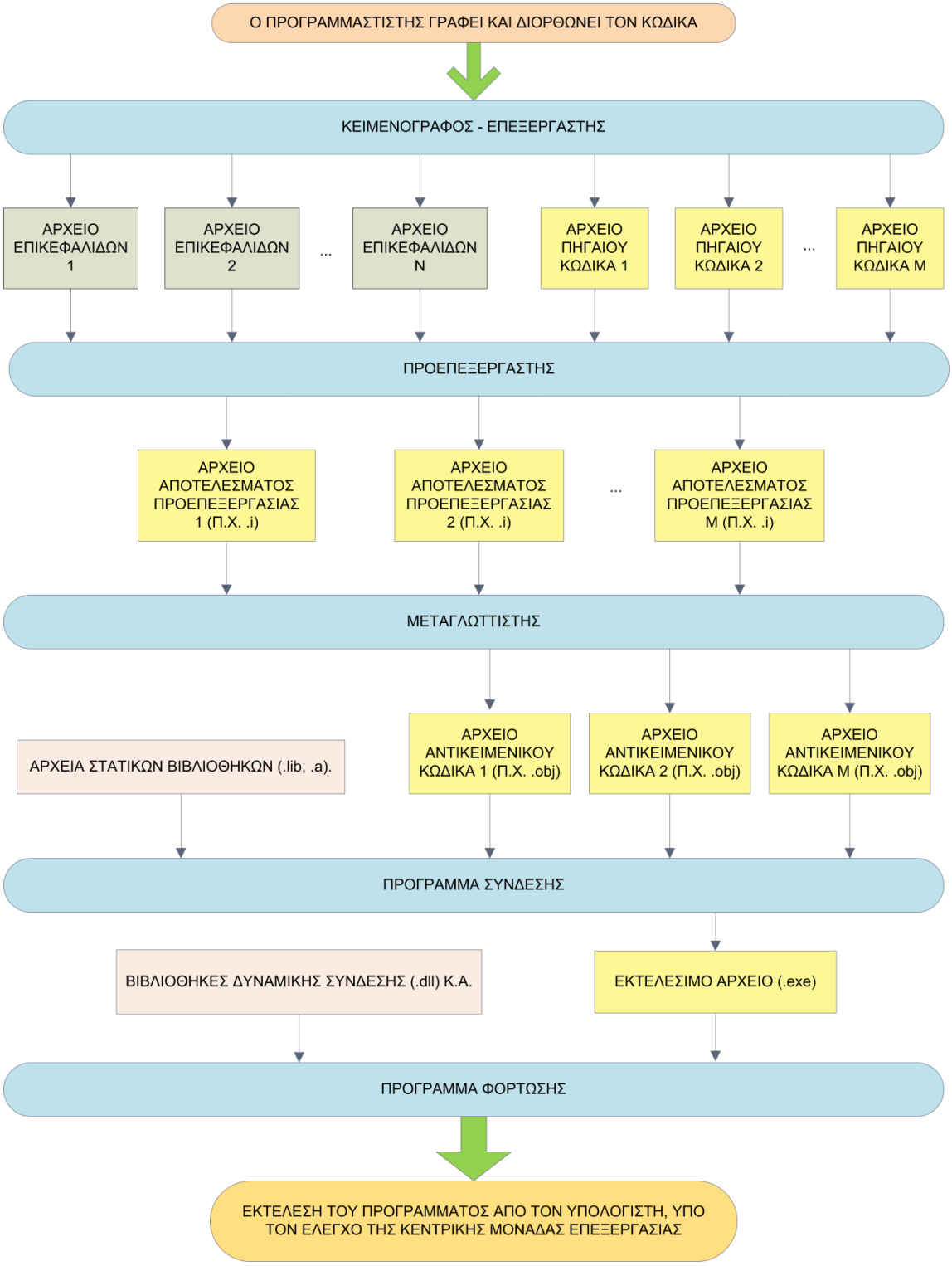
Τέλος, το πρόγραμμα εκτελείται από τον υπολογιστή, υπό τον έλεγχο της Κεντρικής Μονάδας Επεξεργασίας (CPU), η οποία εκτελεί κάθε φορά μία εντολή [157], [165], [166].

4.2.2 Microsoft Visual C++ 2010 Express

Το Microsoft Visual Studio Express (MVSE) είναι μία ομάδα από ολοκληρωμένα περιβάλλοντα ανάπτυξης (Integrated Development Environments - IDEs) που διατίθενται δωρεάν και αναπτύχθηκαν από τη Microsoft. Αποτελεί μία πιο «ελαφριά» έκδοση του Microsoft Visual Studio 2010 και παρέχει περιβάλλοντα ανάπτυξης για ερασιτέχνες, κυρίως, χρήστες ή φοιτητές και όχι τόσο για επαγγελματίες προγραμματιστές λογισμικών.

Η πρώτη του έκδοση ήταν το Microsoft Visual Studio Express 2005 και έκτοτε έχουν αναπτυχθεί άλλες τέσσερις (MVSE 2008, MVSE 2010, MVSE 2012 και MVSE 2013). Αποτελείται από πληθώρα προϊόντων, μεταξύ των οποίων είναι το Microsoft Visual C++, για προγραμματισμό σε γλώσσες C, C++ και C++/CLI[12]. Οι εφαρμογές της παρούσας εργασίας προγραμματίστηκαν σε περιβάλλον Microsoft Visual C++ Express 2010.

[12] Η C++/CLI (Common Language Infrastructure) είναι μία γλώσσα που δημιουργήθηκε από τη Microsoft με σκοπό να αντικαταστήσει ειδικές επεκτάσεις της C++ (Managed Extensions for C++), που σκοπό είχαν να φέρουν τη γλώσσα C++ στο .NET Framework. Το τελευταίο αποτελεί ένα πλαίσιο λογισμικού (software framework) που προορίζεται για τα Windows και αποτελείται από μία μεγάλη βιβλιοθήκη κλάσεων και από το περιβάλλον CLR, στο οποίο εκτελούνται τα προγράμματα. Σκοπός του είναι η απλοποίηση της ανάπτυξης εφαρμογών, υποστηρίζει πληθώρα γλωσσών προγραμματισμού, και παρέχει τη δυνατότητα στην κάθε γλώσσα να χρησιμοποιήσει κώδικα γραμμένο σε άλλη



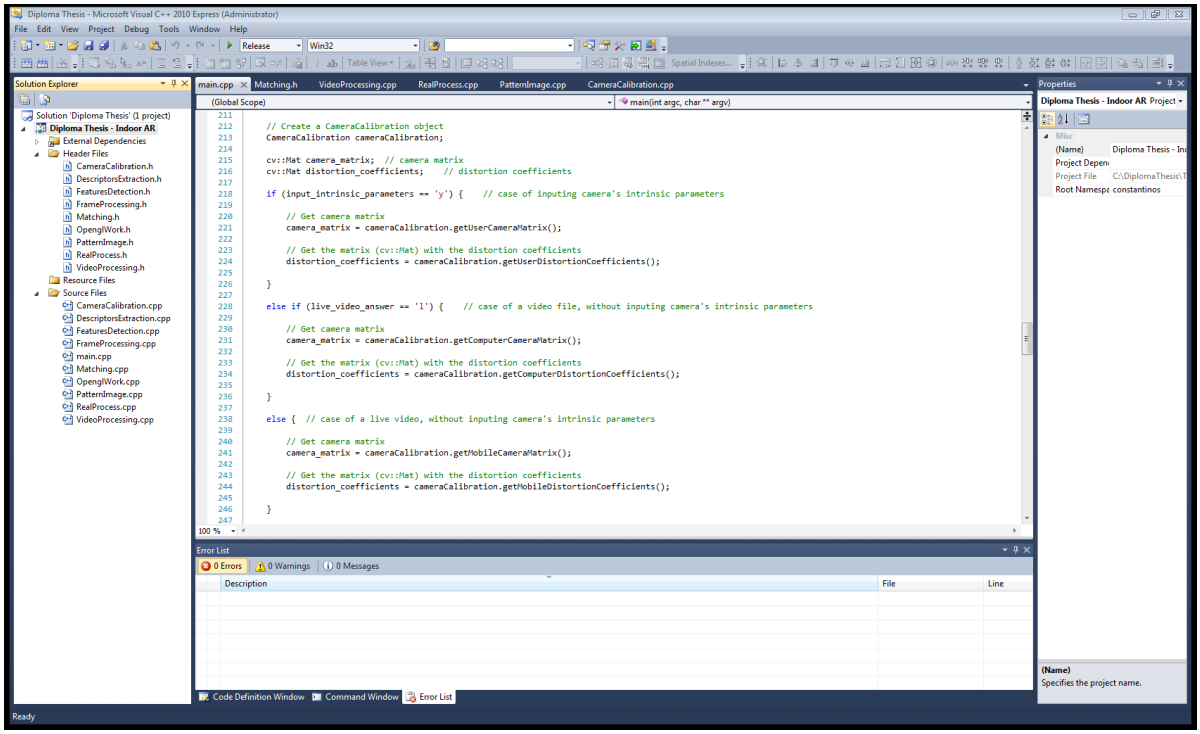
Διαγραμματική παρουσίαση 4-1. Απλοποιημένη διαδικασία που εκτελείται σε ένα τυπικό περιβάλλον ανάπτυξης προγραμμάτων C++.

Πηγές: http://www.tenouk.com/ModuleW.html, http://www.siafoo.net/help/reST, http://www.tenouk.com/Module1.html, https://secweb.cs.odu.edu/~zeil/cs250/website_s13/Lectures/cppProgramStructure/page/cppProgramStructure.html, Deitel & Deitel, 2003

γλώσσα. Μέσω του API (Application Programming Interface) του .NET παρέχονται διάφορες λειτουργίες, όπως για παράδειγμα γραφικές διεπαφές χρηστών, κ.ά., οι οποίες μπορούν να συνδυαστούν με τον κώδικα των προγραμματιστών για τη δημιουργία ολοκληρωμένων εφαρμογών [199], [200], [201], [202].

Το Microsoft Visual C++ Express 2010 χρησιμοποιείται για την εγγραφή και τροποποίηση κώδικα και την προεπεξεργασία, μεταγλώττιση, σύνδεση και φόρτωση εφαρμογών Win32 (όπως στην περίπτωση της παρούσας εργασίας) και .NET, ενώ με κάποιες διαμορφώσεις μπορούν να υποστηριχθούν και 64-bit εφαρμογές. Οι εφαρμογές που μεταγλωττίζονται με τη βοήθειά του μπορούν να τρέξουν σε Windows XP SP2 και σε νεότερες εκδόσεις των Windows. Επίσης, λόγω της χρησιμοποιούμενης τεχνολογίας Microsoft IntelliSense, επισημαίνονται σε πραγματικό χρόνο τα λάθη – τόσο συντακτικά όσο και λογικά – που κάνει ο προγραμματιστής κατά τη διάρκεια πληκτρολόγησης του κώδικα [167], [168], [169].

Ένα τυπικό περιβάλλον εργασίας του Microsoft Visual C++ Express 2010 απεικονίζεται στην Εικόνα 4-1.



Εικόνα 4-1. Περιβάλλον ανάπτυξης προγραμμάτων C++: Microsoft Visual C++ 2010 Express.

4.3 OpenCV και OpenGL

Για την ανάπτυξη των εφαρμογών που προγραμματίστηκαν για τις ανάγκες της παρούσας εργασίας έγινε χρήση της βιβλιοθήκης OpenCV, της διεπαφής προγραμματισμού εφαρμογών OpenGL και της βιβλιοθήκης GLM: An Alias Wavefront OBJ file Library, η οποία χρησιμοποιεί ομοίως την OpenGL. Στην παρούσα ενότητα παρουσιάζεται η βιβλιοθήκη OpenCV, οι βασικές βιβλιοθήκες που σχετίζονται με την OpenGL και χρησιμοποιήθηκαν στις εφαρμογές (OpenGL Core Library, GLU, freeglut, GLEW) και η βιβλιοθήκη GLM. Επίσης, γίνεται αναφορά στα διαφορετικά συστήματα συντεταγμένων κάμερας που χρησιμοποιούνται από την OpenCV και την OpenGL.

4.3.1 Βιβλιοθήκη OpenCV

Η OpenCV (Open Source Computer Vision Library) είναι μία βιβλιοθήκη ανοικτού κώδικα που διαθέτει πάνω από 500 αλγορίθμους σχετικούς κυρίως με το αντικείμενο της όρασης υπολογιστών, για ανάλυση εικόνων και βίντεο. Είναι γραμμένη σε C/C++, ενώ μπορεί να εκτελεστεί σε περισσότερες από μία υπολογιστικές πλατφόρμες, υποστηρίζοντας κυρίως τα λειτουργικά συστήματα Windows, Linux, Mac OS, iOS και Android, αλλά και BlackBerry 10, OpenBSD κ.ά. Το κύριο interface της OpenCV είναι σε C++, ενώ έχει επιπλέον και άλλα, όπως σε C, Python και Java. Ακόμη, η βιβλιοθήκη μπορεί να χρησιμοποιηθεί και σε MATLAB, ή OCTAVE και σε άλλες γλώσσες, όπως για παράδειγμα C# κ.ά.

Η αρχή της βιβλιοθήκης OpenCV χρονολογείται το 1999 και έγινε στην εταιρία Intel από μία ομάδα με επικεφαλής τον Gary Bradski. Σήμερα υποστηρίζεται από τις εταιρίες Willow Garage και Itseez. Η πρώτη «alpha» έκδοσή της διανεμήθηκε στο κοινό το 2000, ενώ πέντε «beta» εκδόσεις δόθηκαν μεταξύ του 2001 και του 2005. Τελικά, η πρώτη έκδοση της OpenCV, η OpenCV 1.0, δημοσιοποιήθηκε το 2006. Η δεύτερη σημαντική έκδοσή της, η OpenCV 2, χρονολογείται τον Οκτώβριο του 2009 και περιλαμβάνει σημαντικές αλλαγές στο C++ interface. Στο πλαίσιο των εφαρμογών, χρησιμοποιήθηκε η έκδοση 2.4.4 της βιβλιοθήκης αυτής, ενώ σήμερα, η τελευταία έκδοση είναι η 2.4.6. Τον Αύγουστο του 2012, ο μη κερδοσκοπικός οργανισμός OpenCV.org ξεκίνησε να υποστηρίζει την OpenCV [124], [170], [171], [172].

Η βιβλιοθήκη OpenCV διανέμεται υπό την άδεια BSD και έτσι είναι ελεύθερη τόσο για ακαδημαϊκή όσο και για εμπορική χρήση. Ωστόσο, κάποιοι αλγόριθμοι, όπως για παράδειγμα ο SURF, είναι ελεύθεροι μόνο για ακαδημαϊκή και όχι για εμπορική χρήση.

4.3.2 OpenGL και σχετικές βιβλιοθήκες

Η OpenGL (Open Graphics Library) είναι μία διεπαφή προγραμματισμού εφαρμογών (Application Programming Interface - API[13]) που χρησιμοποιείται για την απόδοση, δηλαδή την παραγωγή και εμφάνιση (rendering) δισδιάστατων και τρισδιάστατων γραφικών υπολογιστή. Συνεργάζεται με τον επεξεργαστή γραφικών της κάρτας γραφικών (GPU - Graphics Processing Unit), ώστε να επιτευχθεί η λήψη της εικόνας των διαφόρων «μοντέλων» και η οπτικοποίηση αυτών με γρηγορότερο τρόπο με τη βοήθεια του υλικού, διαδικασία η οποία συναντάται με τον όρο hardware-accelerated rendering. Όπως και η OpenCV, η OpenGL μπορεί να εκτελεστεί σε περισσότερες της μίας υπολογιστικές

[13] Το API (Application Programming Inteface) είναι μία προδιαγραφή σχεδιασμένη για να χρησιμοποιείται ως διεπαφή από τα συστατικά του λογισμικού προκειμένου να επικοινωνούν μεταξύ τους. Ένα API μπορεί να περιλαμβάνει προδιαγραφές για ρουτίνες, δομές δεδομένων, κλάσεις και μεταβλητές και διατυπώνει το σύνολο των λειτουργιών που μπορεί να παρέχει μία βιβλιοθήκη ή ένα λειτουργικό σύστημα σε άλλα προγράμματα, χωρίς να επιτρέπει πρόσβαση στον κώδικα που υλοποιεί αυτές τις λειτουργίες [203], [204].

πλατφόρμες και υποστηρίζει διάφορες γλώσσες προγραμματισμού, ενώ είναι σχεδιασμένη έτσι ώστε να είναι ανεξάρτητη από το υλικό (hardware). [173]. Με τον όρο OpenGL, λοιπόν, δεν νοείται κάποια βιβλιοθήκη σχεδίασης γραφικών, αλλά ένα πρότυπο υλοποίησης βιβλιοθηκών σχεδίασης γραφικών, που περιέχει τις συναρτήσεις που πρέπει να υλοποιούνται από μία βιβλιοθήκη, προκειμένου αυτή να είναι συμβατή με το πρότυπο αυτό [174]. Αναπτύχθηκε από την εταιρία Silicon Graphics Inc. το 1991 και δόθηκε στο κοινό το 1992, ενώ σήμερα η μη κερδοσκοπική κοινοπραξία Khronos Group διαχειρίζεται την OpenGL. Στο πλαίσιο των εφαρμογών, χρησιμοποιήθηκε η έκδοση 1.1 της OpenGL.

Οι κατηγορίες βιβλιοθηκών που συναντούνται συχνά σε διάφορες υλοποιήσεις της OpenGL είναι η βασική βιβλιοθήκη OpenGL Core Library, η βιβλιοθήκη GLU (OpenGL Utility Library), η GLUT (OpenGL Utility Toolkit) ή η freeglut αντί της τελευταίας – η οποία χρησιμοποιήθηκε στο πλαίσιο της εργασίας – και η βιβλιοθήκη GLEW (OpenGL Extension Wrangler Library).

Η βασική βιβλιοθήκη της OpenGL περιλαμβάνει τις κύριες εντολές σχεδίασης – που ξεκινούν με το πρόθεμα gl – και πολλές από αυτές δέχονται διάφορα ορίσματα, που αντιστοιχούν σε παραμέτρους ή σε διάφορες καταστάσεις λειτουργίας και ξεκινούν με το πρόθεμα GL_ [174].

Η βιβλιοθήκη GLU αποτελείται από συναρτήσεις που χρησιμοποιούν τη βασική βιβλιοθήκη της OpenGL για την παροχή υψηλότερου επιπέδου σχεδιαστικών ρουτινών από τις αντίστοιχες της βασικής βιβλιοθήκης. Οι συναρτήσεις της βιβλιοθήκης αυτής ξεκινούν με το πρόθεμα glu [175].

Το πρότυπο OpenGL είναι ανεξάρτητο πλατφόρμας και επομένως δεν περιλαμβάνει εντολές για την απεικόνιση των μοντέλων σε κάποιο παράθυρο της οθόνης, ούτε για είσοδο / έξοδο δεδομένων από το χρήστη κ.λπ. Για το λόγο αυτό απαιτείται μία βιβλιοθήκη, η οποία είναι υπεύθυνη για τα παραπάνω. Γενικά, η βιβλιοθήκη GLUT περιλαμβάνει εντολές εμφάνισης παραθύρων, συναρτήσεις που ορίζουν την είσοδο από το πληκτρολόγιο και το ποντίκι κ.ο.κ. Έτσι, δίνεται η δυνατότητα στον προγραμματιστή να δει το αποτέλεσμα των προγραμμάτων του και να μπορεί να αλληλεπιδράσει με αυτό. Όλες οι εντολές της βιβλιοθήκης αυτής ξεκινούν με το πρόθεμα glut [174], [176], [177]. Ωστόσο, η πιο πρόσφατη έκδοσή της χρονολογείται το 1998 και φαίνεται να έχει εγκαταλειφθεί. Αντί αυτής, στο πλαίσιο των εφαρμογών χρησιμοποιήθηκε η βιβλιοθήκη freeglut, η οποία αποτελεί μία αντικατάσταση της GLUT και βρίσκεται συνεχώς υπό ανάπτυξη [178]. Η έκδοση που χρησιμοποιήθηκε είναι η 2.8.1.

Τέλος, η GLEW είναι μία βιβλιοθήκη επεκτάσεων, ανοικτού κώδικα, γραμμένη σε C/C++ και υποστηρίζει διάφορες πλατφόρμες. Η βιβλιοθήκη αυτή καθορίζει ποιες επεκτάσεις της OpenGL υποστηρίζονται στην εκάστοτε πλατφόρμα [179].

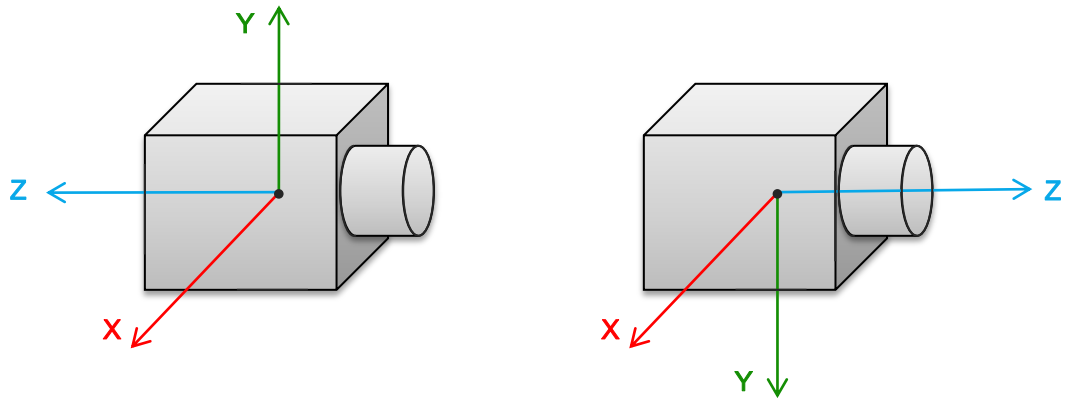
Για τον προγραμματισμό των εφαρμογών στο περιβάλλον Microsoft Visual C++ 2010 Express απαιτήθηκε η λήψη και εγκατάσταση των freeglut και GLEW και η συμπερίληψη

των αρχείων επικεφαλίδων τους στο πρόγραμμα. Τα αρχεία επικεφαλίδων της βασικής βιβλιοθήκης της OpenGL και της GLU συμπεριλαμβάνονται στα αρχεία επικεφαλίδων της freeglut και της GLEW και επομένως με τη συμπερίληψη στο πρόγραμμα των τελευταίων, υπονοείται και η χρήση της βασικής βιβλιοθήκης της OpenGL και της GLU. Οι τελευταίες υπάρχουν ήδη και δεν απαιτείται να γίνει η λήψη τους, καθώς αποτελούν μέρος του πακέτου ανάπτυξης λογισμικού Windows SDK (Software Development Kit).

Εκτός των παραπάνω βασικών βιβλιοθηκών για τον προγραμματισμό σε OpenGL, για τις ανάγκες των εφαρμογών χρησιμοποιήθηκε μία επιπλέον βιβλιοθήκη, η βιβλιοθήκη GLM: An Alias Wavefront OBJ File Library (http://devernay.free.fr/hacks/GLM) που στηρίζεται στην OpenGL. Η βιβλιοθήκη αυτή χρησιμοποιήθηκε για την εισαγωγή των τρισδιάστατων μοντέλων σε κάθε πρόγραμμα. Υπόκειται στη γενική άδεια δημόσιας χρήσης GNU (General Public License), και συγκεκριμένα στη δεύτερη έκδοσή της.

4.3.3 Συστήματα συντεταγμένων κάμερας των OpenCV και OpenGL

Το σύστημα συντεταγμένων της κάμερας που χρησιμοποιείται από την OpenGL είναι ένα δεξιόστροφο τρισορθογώνιο σύστημα με κέντρο το σημείο λήψης, άξονες X και Y παράλληλους με το επίπεδο της εικόνας και άξονα Z κάθετο σε αυτό, με θετική φορά αντίθετη από την κατεύθυνση λήψης (βλ. Εικόνα 4-2, αριστερά). Το σύστημα αυτό ταυτίζεται με το φωτογραμμετρικό σύστημα, το οποίο χρησιμοποιείται στη φωτογραμμετρία. Αντίθετα, το σύστημα συντεταγμένων της κάμερας που χρησιμοποιείται από την OpenCV είναι ένα αριστερόστροφο τρισορθογώνιο σύστημα με το ίδιο κέντρο, του οποίου οι άξονες Y και Z πρέπει να στραφούν κατά 180° γύρω από τον άξονα X προκειμένου να ταυτιστεί με το δεξιόστροφο σύστημα συντεταγμένων που χρησιμοποιείται από την OpenGL (βλ. Εικόνα 4-2, δεξιά). Η παρατήρηση αυτής της διαφοράς είναι πολύ σημαντική για το σωστό σχηματισμό ενός πίνακα μετασχηματισμού που απαιτείται από την OpenGL, με βάση στοιχεία που προέκυψαν με χρήση της βιβλιοθήκης OpenCV, όπως θα γίνει κατανοητό στο επόμενο κεφάλαιο.



Εικόνα 4-2. Συστήματα συντεταγμένων κάμερας της OpenGL [αριστερά] και της OpenCV [δεξιά].

4.4 Μετασχηματισμοί στην OpenGL

Η διαδικασία που ακολουθείται από την OpenGL για την παραγωγή της επιθυμητής σκηνής μέσω μίας σειράς μετασχηματισμών είναι ανάλογη με τα στάδια που πραγματοποιούνται κατά τη διαδικασία της φωτογράφησης, όπως περιγράφονται στη συνέχεια [177], [180].

1. Μετακίνηση της κάμερας στην επιθυμητή θέση και τοποθέτησή της έτσι ώστε να έχει τον επιθυμητό προσανατολισμό. Η διαδικασία αυτή υποδηλώνει το μετασχηματισμό κάμερας (viewing transformation).
2. Μετακίνηση του αντικειμένου στην κατάλληλη θέση εντός της σκηνής. Η διαδικασία αυτή υποδηλώνει το μετασχηματισμό του μοντέλου (modeling transformation).
3. Επιλογή του φακού, διαδικασία που υποδηλώνει το μετασχηματισμό προβολής (projection transformation).
4. Λήψη της φωτογραφίας, η οποία εμπεριέχει την εφαρμογή των παραπάνω μετασχηματισμών.
5. Καθορισμός του μεγέθους της εικόνας, δηλαδή μεγέθυνση ή σμίκρυνση αυτής, κατά τη μία ή τις δύο διευθύνσεις, διαδικασία η οποία υποδηλώνει το μετασχηματισμό παραθύρου παρατήρησης (viewport transformation). Για τρισδιάστατα γραφικά, στο βήμα αυτό περιλαμβάνεται επίσης και η κλιμάκωση του βάθους.

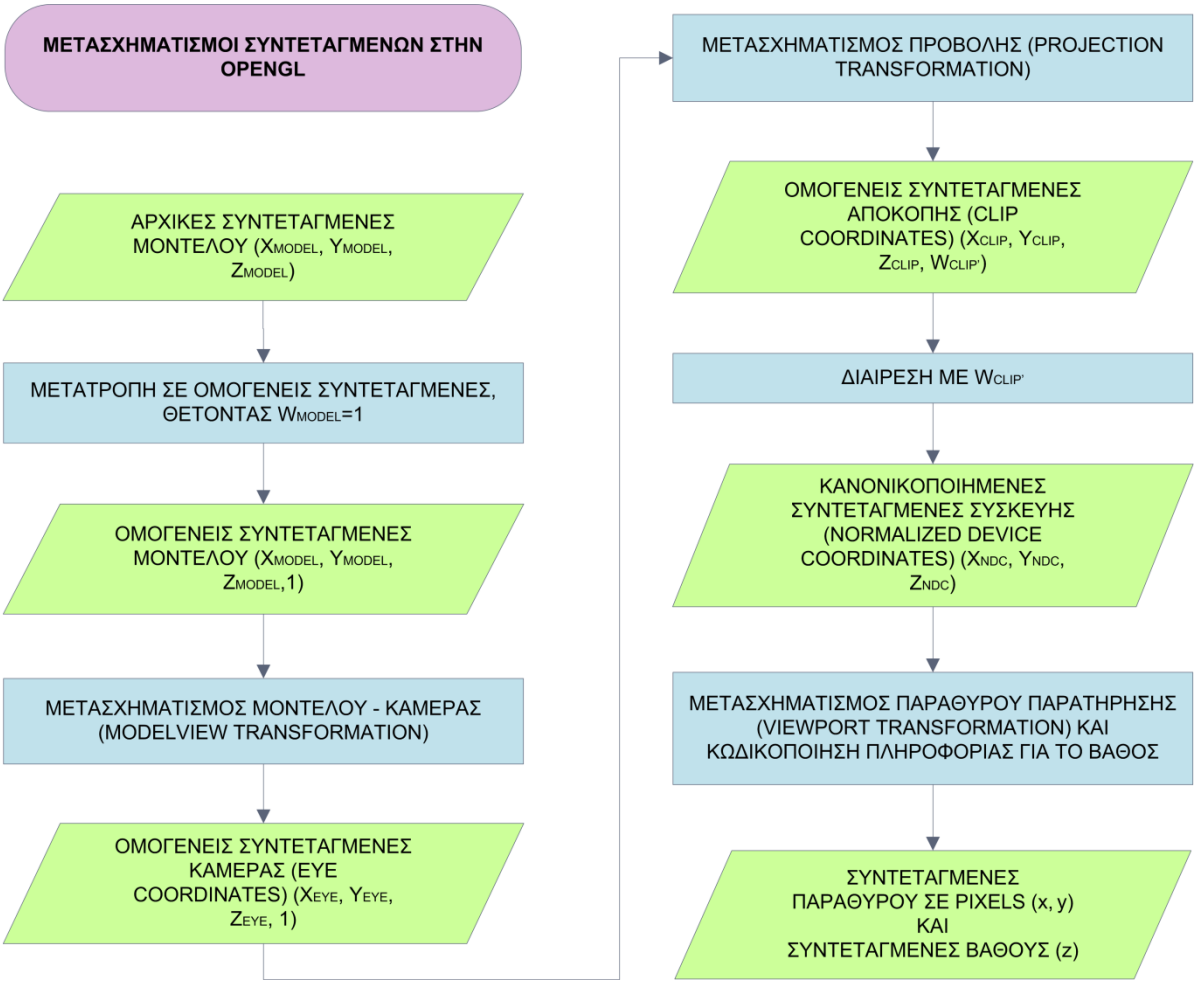
Ωστόσο, γίνεται εύκολα κατανοητό ότι το πρώτο και δεύτερο βήμα εκτελούν ουσιαστικά την ίδια διαδικασία. Έτσι, η μετακίνηση της κάμερας προς τα δεξιά ισοδυναμεί με μετακίνηση του αντικειμένου προς τα αριστερά, η δεξιόστροφη περιστροφή αυτής ισοδυναμεί με αριστερόστροφη περιστροφή του αντικειμένου κ.ο.κ. Συνεπώς, τα δύο αυτά βήματα συνιστούν έναν ενιαίο μετασχηματισμό, ο οποίος ονομάζεται μετασχηματισμός μοντέλου - κάμερας (modelview transformation) και μπορεί να περιλαμβάνει μετάθεση, στροφή και αλλαγή κλίμακας.

Στη Διαγραμματική παρουσίαση 4-2 φαίνεται η διαδικασία κατά την οποία η OpenGL μετασχηματίζει τις συντεταγμένες του αντικειμένου (μοντέλου) σε συντεταγμένες παραθύρου, ενώ στις υποενότητες που ακολουθούν περιγράφονται με περισσότερη λεπτομέρεια οι μετασχηματισμοί αυτοί.

4.4.1 Μετασχηματισμός μοντέλου - κάμερας

Όπως ήδη αναφέρθηκε, ο μετασχηματισμός αυτός συνίσταται από τους μετασχηματισμούς μοντέλου (modeling transformation) και κάμερας (viewing transformation). Ο πρώτος περιστρέφει και μετακινεί το αντικείμενο στο χώρο, μετατρέποντας έτσι τις συντεταγμένες μοντέλου σε συντεταγμένες στο επίγειο σύστημα αναφοράς. Οι συντεταγμένες μοντέλου εκφράζονται σε μονάδες μοντέλου (π.χ. μέτρα). Ο μετασχηματισμός κάμερας

προσανατολίζει τη σκηνή στα μάτια του παρατηρητή, μετατρέποντας έτσι τις συντεταγμένες του αντικειμένου από το επίγειο σύστημα αναφοράς στο σύστημα της κάμερας ή φωτογραμμετρικό σύστημα, όπως συνηθίζεται να ονομάζεται στη φωτογραμμετρία [177], [180].



Διαγραμματική παρουσίαση 4-2. Μετασχηματισμοί συντεταγμένων που πραγματοποιούνται από την OpenGL.

Συνεπώς, ο μετασχηματισμός μοντέλου - κάμερας μετατρέπει τις συντεταγμένες του αντικειμένου από το σύστημα αναφοράς του αντικειμένου στο φωτογραμμετρικό σύστημα. Επειδή εφαρμόζεται σε ομογενείς συντεταγμένες, στις αρχικές συντεταγμένες του αντικειμένου (X_{MODEL}, Y_{MODEL}, Z_{MODEL}) προστίθεται μία τέταρτη συντεταγμένη ίση με τη μονάδα (βλ. Διαγραμματική παρουσίαση 4-2). Ο πίνακας του μετασχηματισμού $\mathbf{M_{MODELVIEW}}$ είναι ένας τετραγωνικός πίνακας 4x4, ο οποίος προκύπτει από το συνδυασμό των πινάκων $\mathbf{M_{MODEL}}$ και $\mathbf{M_{VIEW}}$, δηλαδή $\mathbf{M_{MODELVIEW}} = \mathbf{M_{VIEW}} \cdot \mathbf{M_{MODEL}}$. Αν συμβολιστεί ως $[X_{EYE}\ Y_{EYE}\ Z_{EYE}\ 1]^T$ το διάνυσμα ομογενών συντεταγμένων του αντικειμένου στο σύστημα της κάμερας, ισχύει η ακόλουθη σχέση:

$$\begin{bmatrix} X_{EYE} \\ Y_{EYE} \\ Z_{EYE} \\ 1 \end{bmatrix} = \mathbf{M}_{\mathbf{MODELVIEW}} \cdot \begin{bmatrix} X_{MODEL} \\ Y_{MODEL} \\ Z_{MODEL} \\ 1 \end{bmatrix} \rightarrow \begin{bmatrix} X_{EYE} \\ Y_{EYE} \\ Z_{EYE} \\ 1 \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \\ m_{41} & m_{42} & m_{43} & m_{44} \end{bmatrix} \cdot \begin{bmatrix} X_{MODEL} \\ Y_{MODEL} \\ Z_{MODEL} \\ 1 \end{bmatrix} \rightarrow$$

$$\begin{bmatrix} X_{EYE} \\ Y_{EYE} \\ Z_{EYE} \\ 1 \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} X_{MODEL} \\ Y_{MODEL} \\ Z_{MODEL} \\ 1 \end{bmatrix} \quad (4.1)$$

Οι τρεις ομάδες στοιχείων του πίνακα $\mathbf{M}_{\mathbf{MODELVIEW}}$: (m_{11}, m_{21}, m_{31}), (m_{12}, m_{22}, m_{32}) και (m_{13}, m_{23}, m_{33}) αναπαριστούν τους μετασχηματισμούς στροφής και κλίμακας, ενώ τα στοιχεία (m_{14}, m_{24}, m_{34}) αναπαριστούν το μετασχηματισμό μετάθεσης.

Από τα παραπάνω γίνεται κατανοητό ότι ο πίνακας $\mathbf{M}_{\mathbf{MODELVIEW}}$ του μετασχηματισμού μοντέλου - κάμερας προκύπτει από τον πίνακα $\left[\mathbf{R} \mid \mathbf{t}\right]$ – που περιλαμβάνει τα στοιχεία του εξωτερικού προσανατολισμού, τα οποία ορίζουν το μετασχηματισμό μετάθεσης και στροφής που απαιτείται για να ταυτιστεί το επίγειο με το φωτογραμμετρικό σύστημα αναφοράς – υπό την προϋπόθεση ότι το σύστημα του μοντέλου και το επίγειο σύστημα αναφοράς ταυτίζονται. Αν δεν ταυτίζονται, πρέπει να γίνει πρώτα ο μετασχηματισμός των συντεταγμένων του μοντέλου από το σύστημα του μοντέλου στο επίγειο σύστημα αναφοράς. Στην περίπτωση αυτή, ο πίνακας $\mathbf{M}_{\mathbf{MODELVIEW}}$ θα εφαρμοστεί στις επίγειες συντεταγμένες των κορυφών του μοντέλου που προέκυψαν από τον παραπάνω μετασχηματισμό. Συνεπώς, ο πίνακας $\mathbf{M}_{\mathbf{MODELVIEW}}$ είναι ο εξής:

$$\mathbf{M}_{\mathbf{MODELVIEW}} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.2)$$

όπου r_{ij} τα στοιχεία του πίνακα στροφής και t_i τα στοιχεία του διανύσματος μετάθεσης. Ωστόσο, στην περίπτωση που τα στοιχεία αυτά αναφέρονται σε κάποιο διαφορετικά ορισμένο φωτογραμμετρικό σύστημα (π.χ. σε αυτό που ορίζεται στη βιβλιοθήκη OpenCV), πρέπει να μετασχηματιστούν, λαμβάνοντας υπ' όψιν τη σχέση που συνδέει τα δύο διαφορετικά ορισμένα τρισορθογώνια φωτογραμμετρικά συστήματα.

Ο μετασχηματισμός μοντέλου - κάμερας εφαρμόζεται όχι μόνο στις συντεταγμένες των κορυφών του μοντέλου, αλλά και σε κανονικά διανύσματα, για τον υπολογισμό του φωτισμού που δέχεται κάθε κορυφή ή επιφάνεια. Τα κανονικά διανύσματα αποτελούν διανύσματα τα οποία είναι κάθετα στο εφαπτόμενο επίπεδο του κάθε σημείου μίας επιφάνειας. Καθώς κάθε στοιχειώδης πολυγωνική επιφάνεια συμπίπτει με το εφαπτόμενο επίπεδό της, το κανονικό διάνυσμα μίας στοιχειώδους πολυγωνικής επιφάνειας είναι κάθετο

σε αυτή [181]. Ωστόσο, ο μετασχηματισμός αυτός εφαρμόζεται με διαφορετικό τρόπο σε σχέση με τις κορυφές και ο πολλαπλασιασμός του κανονικού διανύσματος γίνεται με τον ανάστροφο του αντίστροφου πίνακα $\mathbf{M_{MODELVIEW}}$, όπως θα αποδειχθεί στη συνέχεια [180], [182].

Για την απόδειξη της παραπάνω πρότασης, αντί το κανονικό διάνυσμα να θεωρηθεί ως διάνυσμα, είναι βολικό να θεωρηθεί ως ομογενές επίπεδο κάθετο στο διάνυσμα αυτό. Ένα σημείο, που ορίζεται από το διάνυσμα ομογενών συντεταγμένων $\begin{bmatrix} x & y & z & w \end{bmatrix}^T$, ανήκει στο ομογενές επίπεδο, που ορίζεται από το διάνυσμα $\begin{bmatrix} a & b & c & d \end{bmatrix}^T$, εφόσον το εσωτερικό γινόμενο των δύο διανυσμάτων είναι μηδενικό, δηλαδή αν ισχύει η ακόλουθη σχέση:

$$a \cdot x + b \cdot y + c \cdot z + d \cdot w = 0 \;\rightarrow\; \begin{bmatrix} a & b & c & d \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = 0 \qquad (4.3)$$

Έστω μία στοιχειώδης τριγωνική επιφάνεια ενός τρισδιάστατου μοντέλου με κανονικό διάνυσμα $\mathbf{n_{MODEL}} = \begin{bmatrix} n_{X_{MODEL}} & n_{Y_{MODEL}} & n_{Z_{MODEL}} & n_{W_{MODEL}} \end{bmatrix}^T$ και με συντεταγμένες μίας κορυφής της $\begin{bmatrix} X_{MODEL} & Y_{MODEL} & Z_{MODEL} & W_{MODEL} \end{bmatrix}^T$. Θεωρώντας ότι η κορυφή αυτή ανήκει στο ομογενές επίπεδο που είναι κάθετο στο κανονικό διάνυσμα, η σχέση (4.3) λαμβάνει την ακόλουθη μορφή:

$$\begin{bmatrix} n_{X_{MODEL}} & n_{Y_{MODEL}} & n_{Z_{MODEL}} & n_{W_{MODEL}} \end{bmatrix} \cdot \begin{bmatrix} X_{MODEL} \\ Y_{MODEL} \\ Z_{MODEL} \\ W_{MODEL} \end{bmatrix} = 0 \qquad (4.4)$$

Η εξίσωση αυτή γράφεται ισοδύναμα ως εξής:

$$\begin{bmatrix} n_{X_{MODEL}} & n_{Y_{MODEL}} & n_{Z_{MODEL}} & n_{W_{MODEL}} \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} X_{MODEL} \\ Y_{MODEL} \\ Z_{MODEL} \\ W_{MODEL} \end{bmatrix} = 0 \;\rightarrow$$

$$\begin{bmatrix} n_{X_{MODEL}} & n_{Y_{MODEL}} & n_{Z_{MODEL}} & n_{W_{MODEL}} \end{bmatrix} \cdot \mathbf{M}^{-1}_{\mathbf{MODELVIEW}} \cdot \mathbf{M}_{\mathbf{MODELVIEW}} \cdot \begin{bmatrix} X_{MODEL} \\ Y_{MODEL} \\ Z_{MODEL} \\ W_{MODEL} \end{bmatrix} = 0 \;\rightarrow$$

$$\left(\begin{bmatrix} n_{X_{MODEL}} & n_{Y_{MODEL}} & n_{Z_{MODEL}} & n_{W_{MODEL}} \end{bmatrix} \cdot \mathbf{M}^{-1}_{\mathbf{MODELVIEW}}\right) \cdot \left(\mathbf{M}_{\mathbf{MODELVIEW}} \cdot \begin{bmatrix} X_{MODEL} \\ Y_{MODEL} \\ Z_{MODEL} \\ W_{MODEL} \end{bmatrix}\right) = 0 \quad (4.5)$$

Το δεξί μέλος της εξίσωσης (4.5) μετασχηματίζει τις συντεταγμένες μοντέλου σε συντεταγμένες κάμερας. Αντίστοιχα, το αριστερό μέλος της ίδιας εξίσωσης $\mathbf{n}^{T}_{\mathbf{MODEL}} \cdot \mathbf{M}^{-1}_{\mathbf{MODELVIEW}}$ μετασχηματίζει το κανονικό διάνυσμα στο σύστημα συντεταγμένων της κάμερας. Επομένως, το κανονικό διάνυσμα στο σύστημα συντεταγμένων της κάμερας μπορεί να γραφεί ως εξής:

$$\mathbf{n}_{\mathbf{EYE}} = \mathbf{n}^{T}_{\mathbf{MODEL}} \cdot \mathbf{M}^{-1}_{\mathbf{MODELVIEW}} \quad (4.6)$$

Η σχέση (4.6) με βάση γνωστή ιδιότητα των πινάκων είναι ισοδύναμη με την παρακάτω σχέση:

$$\mathbf{n}_{\mathbf{EYE}} = \left(\mathbf{M}^{-1}_{\mathbf{MODELVIEW}}\right)^{T} \cdot \mathbf{n}_{\mathbf{MODEL}} \quad (4.7)$$

Επομένως, τα κανονικά διανύσματα του μοντέλου μετασχηματίζονται στο σύστημα της κάμερας με βάση την ακόλουθη σχέση:

$$\begin{bmatrix} n_{X_{EYE}} \\ n_{Y_{EYE}} \\ n_{Z_{EYE}} \\ n_{W_{EYE}} \end{bmatrix} = \left(\mathbf{M}^{-1}_{\mathbf{MODELVIEW}}\right)^{T} \cdot \begin{bmatrix} n_{X_{MODEL}} \\ n_{Y_{MODEL}} \\ n_{Z_{MODEL}} \\ n_{W_{MODEL}} \end{bmatrix} \quad (4.8)$$

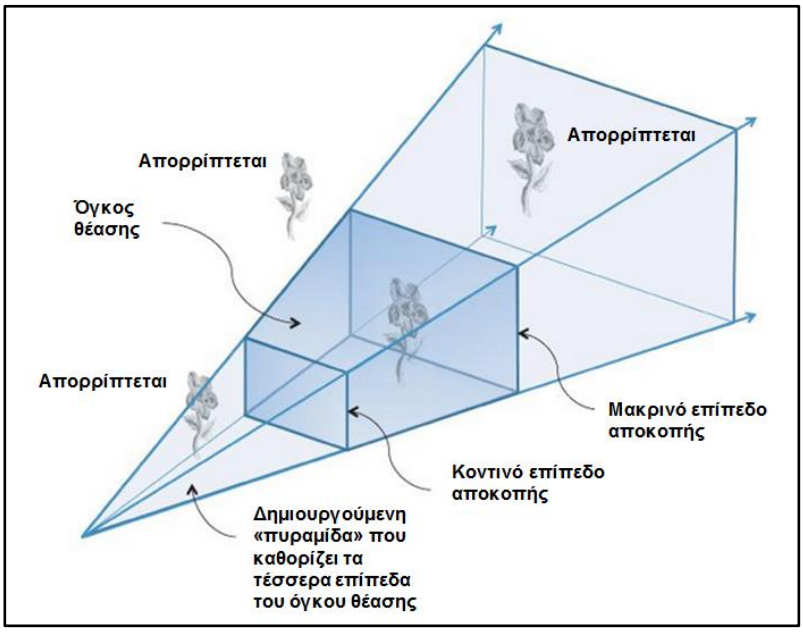
4.4.2 Μετασχηματισμός προβολής και κανονικοποίηση συντεταγμένων

Κατά τη φωτογράφηση, τα αντικείμενα που απεικονίζονται στην τελική εικόνα είναι εκείνα τα οποία βρίσκονται εντός του όγκου θέασης (viewing frustrum), ο οποίος ορίζεται από έξι επίπεδα, όπως φαίνεται στην Εικόνα 4-3, τα οποία συχνά αναφέρονται και ως επίπεδα αποκοπής (clipping planes). Συγκεκριμένα, δεν απεικονίζονται αντικείμενα που βρίσκονται έξω από τη δημιουργούμενη πυραμίδα, αλλά και αντικείμενα που βρίσκονται πολύ κοντά στη φωτομηχανή (δηλαδή μπροστά από το κοντινό επίπεδο αποκοπής ή near plane) και πολύ μακριά από αυτή (δηλαδή πίσω από το μακρινό επίπεδο αποκοπής ή far plane). Συνεπώς, ο όγκος θέασης αποτελεί μία κόλουρη ορθογωνική πυραμίδα[14]. Επίσης, αντικείμενα τα οποία βρίσκονται μερικώς εντός του όγκου θέασης πρέπει να αποκοπούν. Η διαδικασία της απόρριψης των αντικειμένων που βρίσκονται εκτός του όγκου θέασης και της αποκοπής αυτών που τέμνονται από ένα ή περισσότερα εκ των παραπάνω αναφερθέντων

[14] Αυτό ισχύει στην περίπτωση που εφαρμόζεται η προοπτική προβολή. Στην περίπτωση χρήσης της ορθογραφικής προβολής ο όγκος θέασης είναι ένα ορθογώνιο παραλληλεπίπεδο.

επιπέδων γίνεται αυτόματα από την OpenGL, με την προϋπόθεση να έχει οριστεί στο πρόγραμμα ο όγκος θέασης από τον προγραμματιστή. Ο καθορισμός του όγκου θέασης πραγματοποιείται με το τρίτο βήμα της διαδικασίας της φωτογράφησης, όπως ορίστηκε προηγουμένως, κατά το οποίο γίνεται η επιλογή των χαρακτηριστικών της φωτογραφικής μηχανής [180], [183].

Με την παραπάνω διαδικασία προσδιορίζονται οι ομογενείς συντεταγμένες αποκοπής (clipping coordinates). Στη συνέχεια θα παρουσιαστούν δύο είδη μετασχηματισμών προβολής που χρησιμοποιούνται από την OpenGL, η προοπτική και η ορθογραφική (ορθή) προβολή. Ο σκοπός των μετασχηματισμών αυτών είναι ο ορισμός του όγκου θέασης για τον καθορισμό του τρόπου προβολής των αντικειμένων στην οθόνη – χρησιμοποιώντας προοπτική ή ορθή προβολή – και ο ορισμός των αντικειμένων ή των τμημάτων τους τα οποία δεν απεικονίζονται στην τελική εικόνα [177].



Εικόνα 4-3. Όγκος θέασης στην περίπτωση επιλογής της προοπτικής προβολής.

Πηγή: Shreiner et al., 2013

Οι προοπτικοί μετασχηματισμοί τροποποιούν την τέταρτη ομογενή συντεταγμένη W, η οποία πριν την επιβολή του μετασχηματισμού ήταν μονάδα. Η αύξηση της τιμής του W ενός σημείου ισοδυναμεί με απεικόνισή του πιο μακριά. Αυτό συμβαίνει επειδή η μετατροπή των συντεταγμένων αποκοπής του αντικειμένου σε καρτεσιανές συντεταγμένες, δηλαδή κανονικοποιημένες συντεταγμένες συσκευής (normalized device coordinates), ισοδυναμεί με διαίρεση με την τέταρτη συντεταγμένη W (βλ. Διαγραμματική παρουσίαση 4-2). Επομένως, όσο μεγαλύτερη είναι η τιμή αυτή, τόσο μικρότερες είναι οι καρτεσιανές συντεταγμένες του αντικειμένου, με αποτέλεσμα η απεικόνισή του να γίνεται σε μικρότερη κλίμακα. Η ορθή προβολή χρησιμοποιείται όταν δεν είναι επιθυμητή μία προοπτική θέαση της σκηνής, αλλά επιδιώκεται η διατήρηση των πραγματικών μεγεθών των αντικειμένων και των

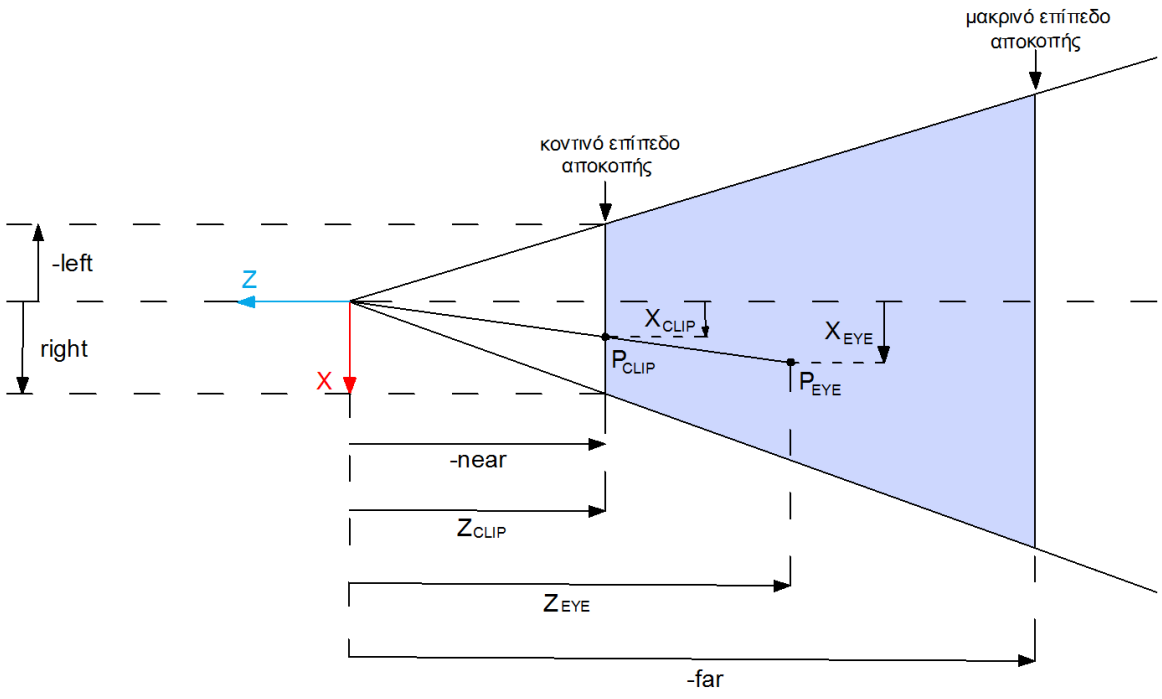
σχηματιζόμενων γωνιών [180]. Στο πλαίσιο της ορθής προβολής ο όγκος θέασης δεν αποτελεί πλέον μία κόλουρη πυραμίδα, αλλά θεωρείται ότι έχει σχήμα παραλληλεπιπέδου.

4.4.2.1 Προοπτική προβολή

Όσο πιο μακριά βρίσκεται ένα αντικείμενο από την κάμερα, τόσο πιο μικρό απεικονίζεται στην τελική εικόνα μέσω προοπτικής προβολής, γεγονός το οποίο οφείλεται στην κόλουρη πυραμίδα που αποτελεί τον όγκο θέασης. Με την εφαρμογή του μετασχηματισμού αυτού – όπως και στην περίπτωση της ορθογραφικής προβολής – μετατρέπονται οι ομογενείς συντεταγμένες κάμερας σε ομογενείς συντεταγμένες αποκοπής. Αν συμβολιστεί ως $\mathbf{P}_{\mathbf{PERSPECTIVE_PROJ.}}$ ο πίνακας του προοπτικού αυτού μετασχηματισμού, ισχύει:

$$\begin{bmatrix} X_{CLIP} \\ Y_{CLIP} \\ Z_{CLIP} \\ W_{CLIP} \end{bmatrix} = \mathbf{P}_{\mathbf{PERSPECTIVE_PROJ.}} \cdot \begin{bmatrix} X_{EYE} \\ Y_{EYE} \\ Z_{EYE} \\ 1 \end{bmatrix} \rightarrow \begin{bmatrix} X_{CLIP} \\ Y_{CLIP} \\ Z_{CLIP} \\ W_{CLIP} \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \\ p_{41} & p_{42} & p_{43} & p_{44} \end{bmatrix} \cdot \begin{bmatrix} X_{EYE} \\ Y_{EYE} \\ Z_{EYE} \\ 1 \end{bmatrix} \quad (4.9)$$

Στη συνέχεια, υπολογίζονται τα στοιχεία του πίνακα του προβολικού μετασχηματισμού [184] θεωρώντας ότι η κάμερα βρίσκεται στο σημείο (0,0,0) και «κοιτάει» προς τον άξονα –Z. Εξετάζεται η γενική περίπτωση του μη συμμετρικού όγκου θέασης, κατά την οποία ο άξονας Z δεν διέρχεται από το κέντρο της πυραμίδας, δηλαδή το πρωτεύον σημείο δεν συμπίπτει με το κέντρο της εικόνας.



Εικόνα 4-4. Τομή στο επίπεδο XZ του όγκου θέασης.

Έστω ένα σημείο P_{EYE}, με διάνυσμα ομογενών συντεταγμένων $\begin{bmatrix} X_{EYE} & Y_{EYE} & Z_{EYE} & 1 \end{bmatrix}^T$ στο σύστημα συντεταγμένων της κάμερας. Η προβολή του σημείου αυτού στο κοντινό επίπεδο

αποκοπής είναι το σημείο P_{CLIP}, με διάνυσμα ομογενών συντεταγμένων $\left[X_{CLIP}\ Y_{CLIP}\ Z_{CLIP}\ W_{CLIP}\right]^T$. Η Εικόνα 4-4 δείχνει την τομή στο επίπεδο XZ του όγκου θέασης. Ως near, far, right και left στην εικόνα αυτή έχουν συμβολιστεί οι εικονιζόμενες αποστάσεις (απόλυτες τιμές). Η εξίσωση της ευθείας που συνδέει τις προβολές στο επίπεδο XZ των σημείων P_{EYE} και P_{CLIP} είναι η ακόλουθη:

$$X = \lambda \cdot Z \quad \text{με } \lambda = \frac{X_{EYE}}{Z_{EYE}} \tag{4.10}$$

Για $X=X_{CLIP}$ και $Z=Z_{CLIP}$=-near (βλ. Εικόνα 4-4) η σχέση (4.10) γράφεται ως εξής:

$$X_{CLIP} = \frac{X_{EYE}}{Z_{EYE}} \cdot Z_{CLIP} \rightarrow X_{CLIP} = \frac{X_{EYE}}{-Z_{EYE}} \cdot \text{near} \tag{4.11}$$

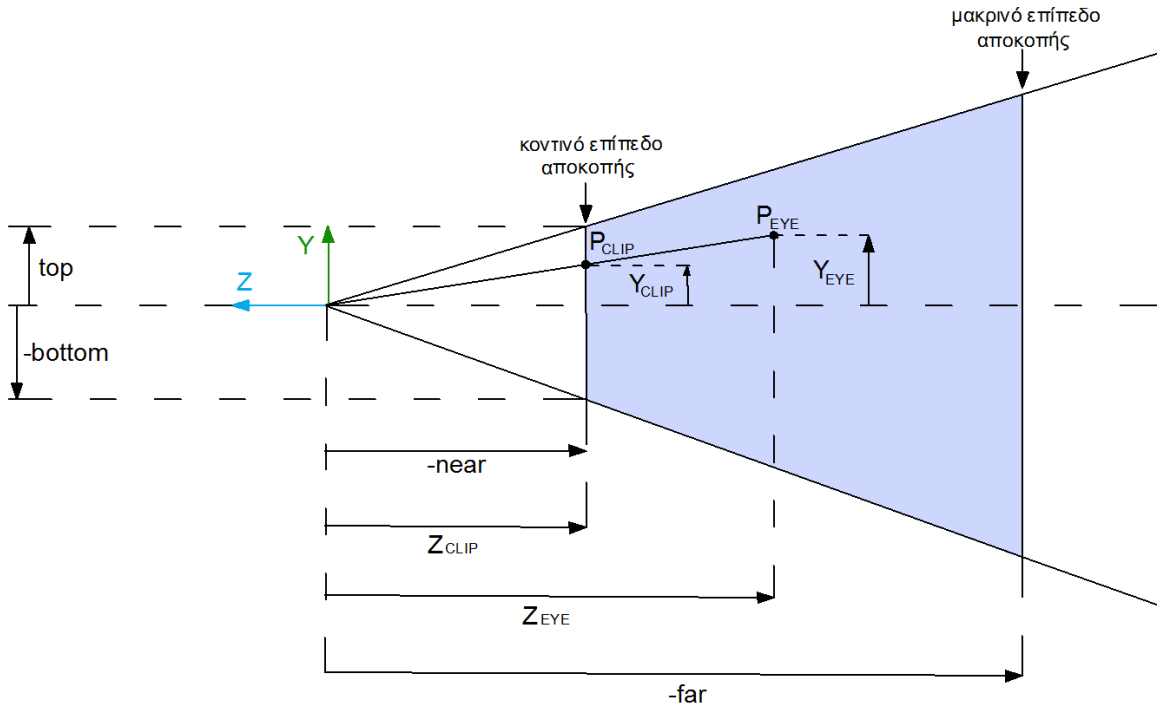
Η Εικόνα 4-5 δείχνει την τομή στο επίπεδο YZ του όγκου θέασης. Η ευθεία που συνδέει τις προβολές των σημείων P_{EYE} και P_{CLIP} στο επίπεδο YZ δίνεται από την παρακάτω εξίσωση:

$$Y = \lambda \cdot Z \quad \text{με } \lambda = \frac{Y_{EYE}}{Z_{EYE}} \tag{4.12}$$

Για $Y=Y_{CLIP}$ και Z=-near (βλ. Εικόνα 4-5) η σχέση (4.12) γράφεται ως εξής:

$$Y_{CLIP} = \frac{Y_{EYE}}{-Z_{EYE}} \cdot \text{near} \tag{4.13}$$

Συνεπώς, τόσο η συντεταγμένη X_{CLIP} όσο και η Y_{CLIP} εξαρτώνται από την απόσταση του σημείου από το επίπεδο της κάμερας, καθώς διαιρούνται με το $-Z_{EYE}$.



Εικόνα 4-5. Τομή στο επίπεδο YZ του όγκου θέασης.

Οι κανονικοποιημένες συντεταγμένες συσκευής $\left[X_{NDC} \; Y_{NDC} \; Z_{NDC} \right]^T$ υπολογίζονται ως εξής:

$$X_{NDC} = X_{CLIP} / W_{CLIP} \tag{4.14}$$

$$Y_{NDC} = Y_{CLIP} / W_{CLIP} \tag{4.15}$$

$$Z_{NDC} = Z_{CLIP} / W_{CLIP} \tag{4.16}$$

Θεωρείται ότι $W_{CLIP} = -Z_{EYE}$, ώστε τα αντικείμενα με μεγαλύτερη τιμή του Z_{EYE} να φαίνονται μικρότερα και έτσι να γίνεται φανερό ότι βρίσκονται πιο μακριά, δημιουργώντας ένα προοπτικό αποτέλεσμα. Με βάση τη θεώρηση αυτή, προκύπτει η τέταρτη σειρά του πίνακα του προοπτικού μετασχηματισμού, ο οποίος λαμβάνει τη μορφή:

$$\mathbf{P}_{\mathbf{PERSPECTIVE_PROJ.}} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \\ 0 & 0 & -1 & 0 \end{bmatrix} \tag{4.17}$$

καθώς ισχύει:

$$\begin{bmatrix} X_{CLIP} \\ Y_{CLIP} \\ Z_{CLIP} \\ W_{CLIP} \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \\ 0 & 0 & -1 & 0 \end{bmatrix} \cdot \begin{bmatrix} X_{EYE} \\ Y_{EYE} \\ Z_{EYE} \\ 1 \end{bmatrix} \rightarrow \; W_{CLIP} = -Z_{EYE} \tag{4.18}$$

Ένα σημείο εντός της κόλουρης πυραμίδας (σε συντεταγμένες κάμερας) μετασχηματίζεται σε σημείο εντός ενός κύβου (σε κανονικοποιημένες συντεταγμένες συσκευής), όπως φαίνεται στην Εικόνα 4-6. Από την ίδια εικόνα γίνεται κατανοητό ότι ο άξονας Z στο σύστημα των κανονικοποιημένων συντεταγμένων συσκευής έχει αλλάξει φορά και δεν αυξάνεται όσο πλησιάζει τον παρατηρητή.

Η συντεταγμένη X_{CLIP} μπορεί να λάβει τιμές στο εύρος [-left, right]. Αντίστοιχα, η κανονικοποιημένη συντεταγμένη συσκευής X_{NDC} λαμβάνει τιμές στο διάστημα [-1,1] (βλ. Εικόνα 4-6). Η γραμμική σχέση τους υπολογίζεται ως εξής (βλ. Εικόνα 4-7, αριστερά):

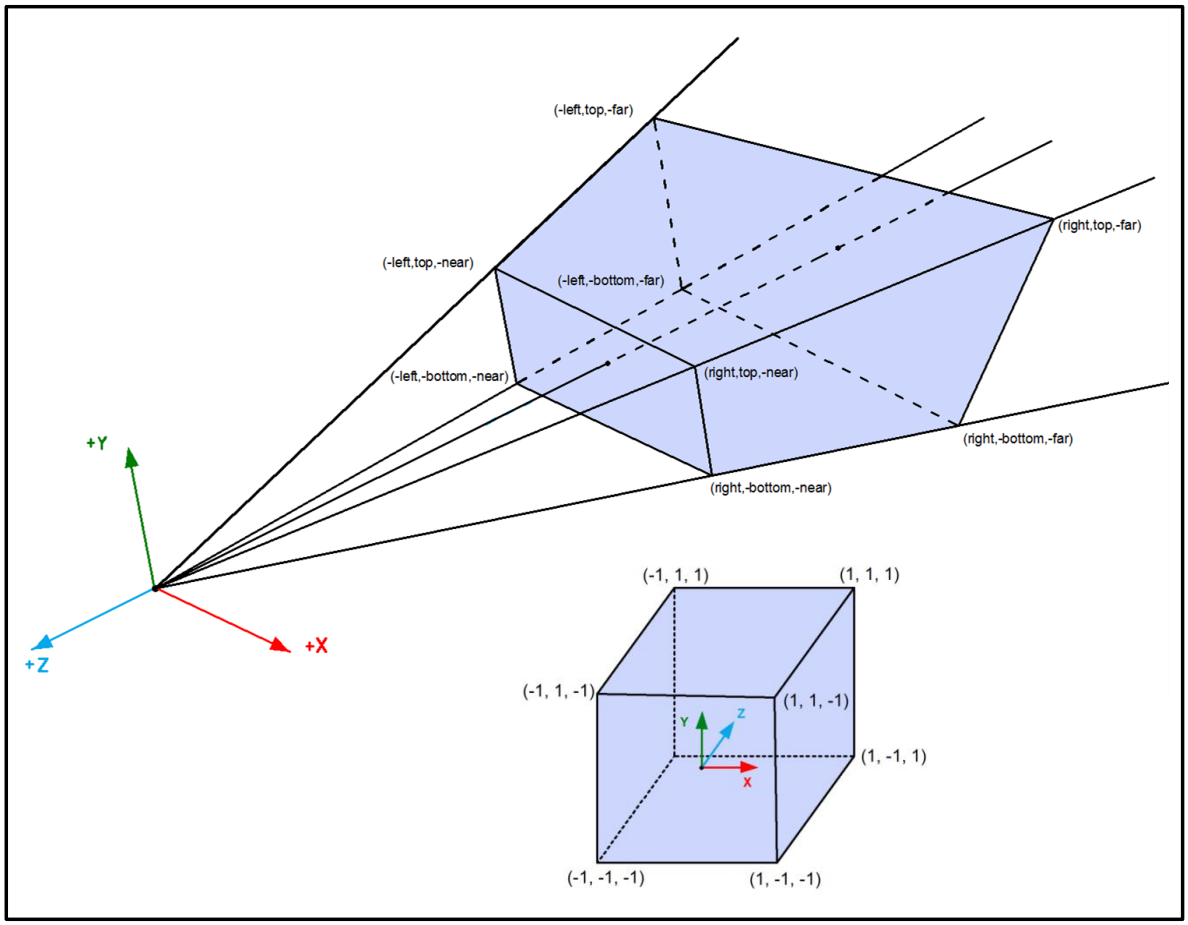
$$X_{NDC} = a \cdot X_{CLIP} + b \; \rightarrow \; \left\{ \begin{array}{l} 1 = a \cdot right + b \\ -1 = a \cdot \left(-left \right) + b \end{array} \right\} \; \rightarrow \; \left\{ \begin{array}{l} a = \dfrac{2}{right + left} \\ b = \dfrac{left - right}{right + left} \end{array} \right\} \; \rightarrow$$

$$X_{NDC} = \frac{2}{right + left} \cdot X_{CLIP} + \frac{left - right}{right + left} \tag{4.19}$$

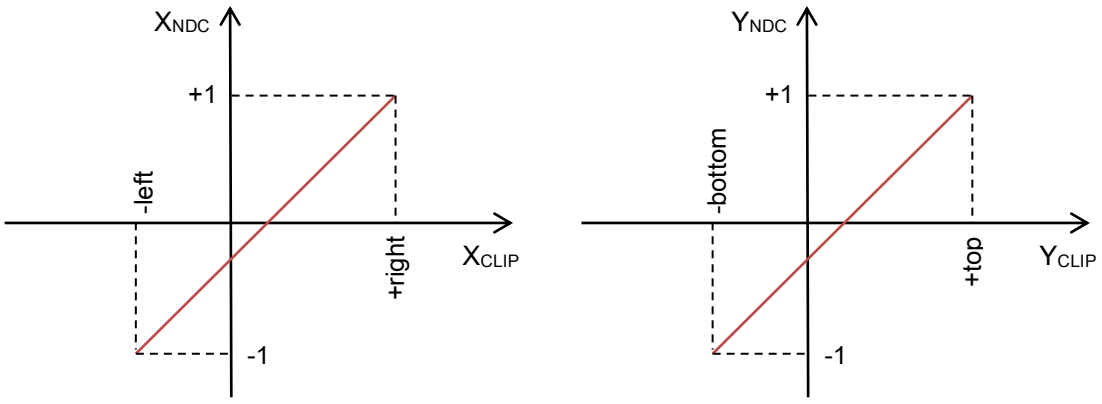
Η συντεταγμένη Y_{CLIP} μπορεί να λάβει τιμές στο εύρος [-bottom, top]. Αντίστοιχα, η κανονικοποιημένη συντεταγμένη συσκευής Y_{NDC} λαμβάνει τιμές στο διάστημα [-1,1] (βλ. Εικόνα 4-6). Η γραμμική σχέση τους υπολογίζεται ως εξής (βλ. Εικόνα 4-7, δεξιά):

$$Y_{NDC} = c \cdot X_{CLIP} + d \rightarrow \left\{ \begin{array}{c} 1 = c \cdot top + d \\ -1 = c \cdot (-bottom) + d \end{array} \right\} \rightarrow \left\{ \begin{array}{l} c = \dfrac{2}{top + bottom} \\ d = \dfrac{bottom - top}{top + bottom} \end{array} \right\} \rightarrow$$

$$Y_{NDC} = \frac{2}{top + bottom} \cdot Y_{CLIP} + \frac{bottom - top}{top + bottom} \qquad (4.20)$$



Εικόνα 4-6. Ένα σημείο εντός της κόλουρης ορθογωνικής πυραμίδας (σε συντεταγμένες κάμερας) μετασχηματίζεται σε σημείο εντός ενός κύβου (σε κανονικοποιημένες συντεταγμένες συσκευής).



Εικόνα 4-7. Γραμμική σχέση μεταξύ των X_{NDC} και X_{CLIP} και μεταξύ των Y_{NDC} και Y_{CLIP}.

Με αντικατάσταση της σχέσης (4.11) στην (4.19) προκύπτει:

$$X_{NDC} = \frac{2}{right + left} \cdot \frac{X_{EYE}}{-Z_{EYE}} \cdot near + \frac{left - right}{right + left} \quad \rightarrow$$

$$X_{NDC} = \frac{\frac{2 \cdot near}{right + left} \cdot X_{EYE} + \frac{right - left}{right + left} \cdot Z_{EYE}}{-Z_{EYE}} \tag{4.21}$$

Με αντικατάσταση της σχέσης (4.13) στην (4.20) προκύπτει:

$$Y_{NDC} = \frac{2}{top + bottom} \cdot \frac{Y_{EYE}}{-Z_{EYE}} \cdot near + \frac{bottom - top}{top + bottom} \quad \rightarrow$$

$$Y_{NDC} = \frac{\frac{2 \cdot near}{top + bottom} \cdot Y_{EYE} + \frac{top - bottom}{top + bottom} \cdot Z_{EYE}}{-Z_{EYE}} \tag{4.22}$$

Με βάση τις σχέσεις (4.14) και (4.15) και δεδομένου ότι $W_{CLIP} = -Z_{EYE}$, ισχύει:

$$X_{NDC} = X_{CLIP} / \left(-Z_{EYE}\right) \tag{4.23}$$

$$Y_{NDC} = Y_{CLIP} / \left(-Z_{EYE}\right) \tag{4.24}$$

Από τις σχέσεις (4.21), (4.22), (4.23) και (4.24) προκύπτουν οι σχέσεις που συνδέουν τις συντεταγμένες X_{CLIP} και Y_{CLIP} με τις συντεταγμένες κάμερας:

$$X_{CLIP} = \frac{2 \cdot near}{right + left} \cdot X_{EYE} + \frac{right - left}{right + left} \cdot Z_{EYE} \tag{4.25}$$

$$Y_{CLIP} = \frac{2 \cdot near}{top + bottom} \cdot Y_{EYE} + \frac{top - bottom}{top + bottom} \cdot Z_{EYE} \tag{4.26}$$

Με βάση τις σχέσεις (4.25) και (4.26) προκύπτουν οι δύο πρώτες γραμμές του πίνακα προοπτικής προβολής, ο οποίος λαμβάνει πλέον την ακόλουθη μορφή:

$$\mathbf{P}_{\mathbf{PERSPECTIVE_PROJ.}} = \begin{bmatrix} \frac{2 \cdot near}{right + left} & 0 & \frac{right - left}{right + left} & 0 \\ 0 & \frac{2 \cdot near}{top + bottom} & \frac{top - bottom}{top + bottom} & 0 \\ p_{31} & p_{32} & p_{33} & p_{34} \\ 0 & 0 & -1 & 0 \end{bmatrix} \tag{4.27}$$

Τέλος, απομένει ο υπολογισμός των στοιχείων της τρίτης γραμμής του πίνακα προοπτικής προβολής, δηλαδή η εύρεση της σχέσης του Z_{CLIP} με τις συντεταγμένες στο σύστημα της κάμερας (X_{EYE}, Y_{EYE}, Z_{EYE}, 1). Καθώς η συντεταγμένη Z_{CLIP} δεν εξαρτάται από τις X_{EYE} και Y_{EYE}, τα δύο πρώτα στοιχεία της τρίτης σειράς είναι μηδενικά. Επομένως η σχέση (4.9) λαμβάνει την παρακάτω μορφή:

$$\begin{bmatrix} X_{CLIP} \\ Y_{CLIP} \\ Z_{CLIP} \\ W_{CLIP} \end{bmatrix} = \begin{bmatrix} \frac{2\cdot near}{right+left} & 0 & \frac{right-left}{right+left} & 0 \\ 0 & \frac{2\cdot near}{top+bottom} & \frac{top-bottom}{top+bottom} & 0 \\ 0 & 0 & p_{33} & p_{34} \\ 0 & 0 & -1 & 0 \end{bmatrix} \cdot \begin{bmatrix} X_{EYE} \\ Y_{EYE} \\ Z_{EYE} \\ 1 \end{bmatrix} \tag{4.28}$$

Με βάση τη σχέση (4.28) προκύπτει:

$$Z_{CLIP} = p_{33} \cdot Z_{EYE} + p_{34} \tag{4.29}$$

Από τη σχέση (4.16) και δεδομένου ότι $W_{CLIP} = -Z_{EYE}$, ισχύει η σχέση:

$$Z_{NDC} = Z_{CLIP} / \left(-Z_{EYE}\right) \tag{4.30}$$

η οποία σύμφωνα με την (4.29) λαμβάνει την ακόλουθη μορφή:

$$Z_{NDC} = \frac{p_{33} \cdot Z_{EYE} + p_{34}}{-Z_{EYE}} \tag{4.31}$$

Η συντεταγμένη Z_{EYE} μπορεί να λάβει τιμές στο εύρος [-near, -far], το οποίο αντιστοιχεί στο εύρος [-1,1] των κανονικοποιημένων συντεταγμένων συσκευής Z_{NDC} (βλ. Εικόνα 4-6). Αντικαθιστώντας τα δύο αυτά ζεύγη τιμών στη σχέση (4.31), προκύπτει το ακόλουθο σύστημα δύο εξισώσεων:

$$\left\{ \begin{array}{c} -1 = \frac{p_{33}\cdot(-near)+p_{34}}{near} \\ 1 = \frac{p_{33}\cdot(-far)+p_{34}}{far} \end{array} \right\} \rightarrow \left\{ \begin{array}{l} near = p_{33}\cdot near - p_{34} \\ far = -p_{33}\cdot far + p_{34} \end{array} \right\} \tag{4.32}$$

από την επίλυση του οποίου υπολογίζονται οι συντελεστές p_{33} και p_{34}.

$$p_{33} = \frac{near + far}{near - far} \tag{4.33}$$

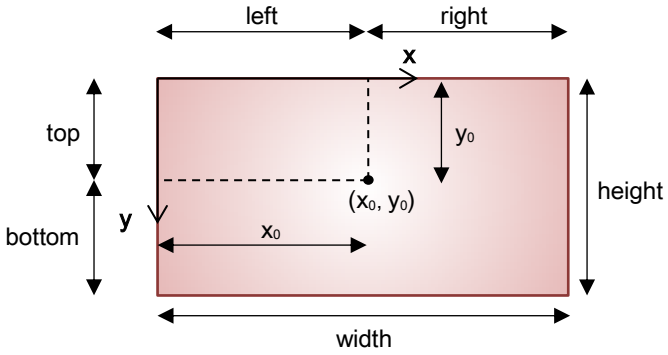
$$p_{34} = \frac{-2\cdot far \cdot near}{far - near} \tag{4.34}$$

Τελικά, σύμφωνα με τα παραπάνω, ο πίνακας προοπτικής προβολής είναι ο ακόλουθος:

$$\mathbf{P}_{\mathbf{PERSPECTIVE_PROJ.}} = \begin{bmatrix} \frac{2\cdot near}{right+left} & 0 & \frac{right-left}{right+left} & 0 \\ 0 & \frac{2\cdot near}{top+bottom} & \frac{top-bottom}{top+bottom} & 0 \\ 0 & 0 & \frac{near+far}{near-far} & \frac{-2\cdot far\cdot near}{far-near} \\ 0 & 0 & -1 & 0 \end{bmatrix} \tag{4.35}$$

Ο παραπάνω πίνακας προέκυψε θεωρώντας ότι το επίπεδο προβολής είναι το κοντινό επίπεδο αποκοπής και είναι ο πίνακας ο οποίος χρησιμοποιείται από την OpenGL για την εφαρμογή του μετασχηματισμού προοπτικής προβολής. Ωστόσο, στην περίπτωση γνώσης της σταθεράς της μηχανής είναι γνωστό και το επίπεδο προβολής, το οποίο γενικά δεν ταυτίζεται με το κοντινό επίπεδο αποκοπής. Έτσι, ο παραπάνω πίνακας μπορεί να εκφραστεί συναρτήσει της σταθεράς της μηχανής, αλλά και των διαστάσεων της εικόνας, των συντεταγμένων του πρωτεύοντος σημείου και των αποστάσεων των κοντινών και μακρινών επιπέδων αποκοπής από το σημείο λήψης.

Ως αρχή του συστήματος συντεταγμένων της εικόνας λαμβάνεται το επάνω αριστερά pixel, όπως θεωρείται και από τη βιβλιοθήκη OpenCV (βλ. Εικόνα 4-8). Οι συντεταγμένες pixel του πρωτεύοντος σημείου συμβολίζονται ως x_0 και y_0 και είναι θετικοί αριθμοί, όπως προκύπτει από το σύστημα αναφοράς που ορίστηκε. Συγκεκριμένα, η τετμημένη x_0 συμπίπτει με την απόσταση left, ενώ η τεταγμένη του πρωτεύοντος σημείου συμπίπτει με την απόσταση top. Επίσης, το πλάτος της εικόνας (width) σε pixels είναι ίσο με το άθροισμα των αποστάσεων left και right, ενώ το ύψος της (height), ομοίως εκφρασμένο σε pixels, ισούται με το άθροισμα των αποστάσεων top και bottom.



Εικόνα 4-8. Σχέση μεταξύ των συντεταγμένων του πρωτεύοντος σημείου και των διαστάσεων της εικόνας με τις αποστάσεις των τεσσάρων επιπέδων αποκοπής (top, bottom, left, right) από το σημείο λήψης.

Στα στοιχεία p_{11} και p_{22} του πίνακα του προοπτικού μετασχηματισμού, η απόσταση near αντικαθίσταται από τη σταθερά της μηχανής, εκφρασμένη σε pixels, αφού θεωρείται γνωστή η απόσταση σχηματισμού της εικόνας από το προβολικό κέντρο κατά τον άξονα λήψης. Όπως αναφέρθηκε στην υποενότητα 3.3.2, η σταθερά της μηχανής εκφρασμένη σε pixels μεγέθους ίσου με το μήκος της οριζόντιας πλευράς τους αναφέρεται ως c_x, ενώ η σταθερά της μηχανής εκφρασμένη σε pixels μεγέθους ίσου με το μήκος της κατακόρυφης πλευράς τους αναφέρεται ως c_y. Έτσι, στο στοιχείο p_{11} η απόσταση near αντικαθίσταται από τη c_x, ενώ στο στοιχείο p_{12} η ίδια απόσταση αντικαθίσταται από τη c_y, έτσι ώστε να ληφθεί υπ'όψιν η ενδεχόμενη διαφορά στα μεγέθη των πλευρών των εικονοστοιχείων. Ωστόσο, στα στοιχεία p_{33} και p_{34} η απόσταση near δεν αντικαθίσταται από τη σταθερά της μηχανής, καθώς δεν αναφέρεται στην απόσταση σχηματισμού της εικόνας, αλλά στο δυνατό εύρος τιμών των συντεταγμένων Z_{EYE}, από –near έως –far (βλ. σχέσεις (4.31) και (4.32)).

Με βάση τα παραπάνω, μπορούν να υπολογιστούν τα στοιχεία του πίνακα προοπτικής προβολής:

$$p_{11} = \frac{2 \cdot \text{near}}{\text{right} + \text{left}} = \frac{2 \cdot c_x}{\text{width}} \tag{4.36}$$

$$p_{13} = \frac{\text{right} - \text{left}}{\text{right} + \text{left}} = \frac{\text{right} + \text{left} - \text{left} - \text{left}}{\text{right} + \text{left}} = \frac{\text{right} + \text{left}}{\text{right} + \text{left}} - \frac{2 \cdot \text{left}}{\text{right} + \text{left}} = 1 - \frac{2 \cdot x_0}{\text{width}} \tag{4.37}$$

$$p_{22} = \frac{2 \cdot \text{near}}{\text{top} + \text{bottom}} = \frac{2 \cdot c_y}{\text{height}} \tag{4.38}$$

$$p_{23} = \frac{\text{top} - \text{bottom}}{\text{top} + \text{bottom}} = \frac{\text{top} + \text{bottom} - \text{bottom} - \text{bottom}}{\text{top} + \text{bottom}} = 1 - \frac{2 \cdot \text{bottom}}{\text{top} + \text{bottom}} =$$
$$1 - \frac{2 \cdot (\text{height} - \text{top})}{\text{height}} = 1 - 2 + \frac{2 \cdot \text{top}}{\text{height}} = -1 + \frac{2 \cdot y_0}{\text{height}} \tag{4.39}$$

Τα υπόλοιπα στοιχεία παραμένουν αμετάβλητα. Συνεπώς, ο πίνακας προοπτικού μετασχηματισμού, στην περίπτωση που είναι γνωστά τα στοιχεία του εσωτερικού προσανατολισμού, λαμβάνει την ακόλουθη μορφή:

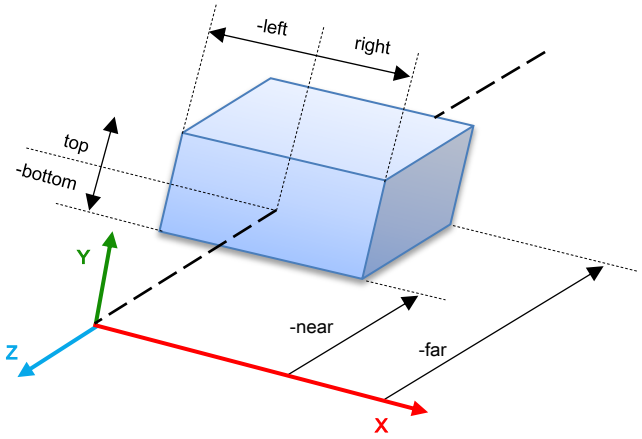
$$\mathbf{P}_{\text{PERSPECTIVE_PROJ.}} = \begin{bmatrix} \frac{2 \cdot c_x}{\text{width}} & 0 & 1 - \frac{2 \cdot x_0}{\text{width}} & 0 \\ 0 & \frac{2 \cdot c_y}{\text{height}} & -1 + \frac{2 \cdot y_0}{\text{height}} & 0 \\ 0 & 0 & \frac{\text{near} + \text{far}}{\text{near} - \text{far}} & \frac{-2 \cdot \text{far} \cdot \text{near}}{\text{far} - \text{near}} \\ 0 & 0 & -1 & 0 \end{bmatrix} \tag{4.40}$$

Στον πίνακα αυτό, οι αποστάσεις near και far ορίζονται από τον προγραμματιστή, ενώ τα υπόλοιπα στοιχεία είτε προκύπτουν από τη βαθμονόμηση (σταθερά μηχανής, πρωτεύον σημείο), είτε είναι γνωστά (μήκος και πλάτος εικόνας). Τέλος, στην περίπτωση που θεωρείται ότι ο όγκος θέασης είναι συμμετρικός, με αποτέλεσμα το πρωτεύον σημείο να ταυτίζεται με το κέντρο της εικόνας, μηδενίζονται τα στοιχεία p_{13} και p_{23} καθώς $2x_0$=width και $2y_0$=height. Συνεπώς, ο πίνακας της προοπτικής προβολής στην περίπτωση του συμμετρικού όγκου θέασης είναι ο εξής:

$$\mathbf{P}_{\text{PERSPECTIVE_PROJ.}} = \begin{bmatrix} \frac{2 \cdot c_x}{\text{width}} & 0 & 0 & 0 \\ 0 & \frac{2 \cdot c_y}{\text{height}} & 0 & 0 \\ 0 & 0 & \frac{\text{near} + \text{far}}{\text{near} - \text{far}} & \frac{-2 \cdot \text{far} \cdot \text{near}}{\text{far} - \text{near}} \\ 0 & 0 & -1 & 0 \end{bmatrix} \tag{4.41}$$

4.4.2.2 Ορθογραφική προβολή

Στο πλαίσιο της ορθογραφικής προβολής (βλ. Εικόνα 4-9), ο όγκος θέασης είναι ένα ορθογώνιο παραλληλεπίπεδο και η απόσταση των αντικειμένων από το σημείο λήψης δεν επηρεάζει το μέγεθος με το οποίο απεικονίζονται στην εικόνα. Στην περίπτωση αυτή όλες οι γραμμές προβολής είναι κάθετες στο επίπεδο προβολής.



Εικόνα 4-9. Θεωρούμενος όγκος θέασης στην ορθογραφική προβολή.

Για την εφαρμογή της ορθογραφικής προβολής, η συντεταγμένη W_{CLIP} τίθεται ίση με τη μονάδα, αφού δεν φέρει κάποια χρήσιμη πληροφορία [184]. Επομένως, η σχέση που συνδέει τις συντεταγμένες αποκοπής με τις συντεταγμένες στο φωτογραμμετρικό σύστημα, μέσω της ορθογραφικής προβολής:

$$\begin{bmatrix} X_{CLIP} \\ Y_{CLIP} \\ Z_{CLIP} \\ W_{CLIP} \end{bmatrix} = \mathbf{P}_{ORTHOGRAPHIC_PROJ.} \cdot \begin{bmatrix} X_{EYE} \\ Y_{EYE} \\ Z_{EYE} \\ 1 \end{bmatrix} \rightarrow \begin{bmatrix} X_{CLIP} \\ Y_{CLIP} \\ Z_{CLIP} \\ W_{CLIP} \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \\ p_{41} & p_{42} & p_{43} & p_{44} \end{bmatrix} \cdot \begin{bmatrix} X_{EYE} \\ Y_{EYE} \\ Z_{EYE} \\ 1 \end{bmatrix} \quad (4.42)$$

μπορεί να γραφεί ως εξής:

$$\begin{bmatrix} X_{CLIP} \\ Y_{CLIP} \\ Z_{CLIP} \\ W_{CLIP} \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} X_{EYE} \\ Y_{EYE} \\ Z_{EYE} \\ 1 \end{bmatrix} \quad (4.43)$$

Με δεδομένο ότι $W_{CLIP}=1$ και με βάση τις σχέσεις (4.14), (4.15) και (4.16), οι κανονικοποιημένες συντεταγμένες συσκευής είναι ίσες με τις συντεταγμένες αποκοπής:

$$X_{NDC} = X_{CLIP} \quad (4.44)$$

$$Y_{NDC} = Y_{CLIP} \quad (4.45)$$

$$Z_{NDC} = Z_{CLIP} \quad (4.46)$$

Συνεπώς, αντί να βρεθεί η σχέση που συνδέει τις συντεταγμένες αποκοπής X_{CLIP}, Y_{CLIP}, Z_{CLIP} με τις συντεταγμένες στο σύστημα της κάμερας X_{EYE}, Y_{EYE}, Z_{EYE} υπολογίζεται η σχέση που συνδέει τις κανονικοποιημένες συντεταγμένες συσκευής X_{NDC}, Y_{NDC}, Z_{NDC} με τις συντεταγμένες στο φωτογραμμετρικό σύστημα X_{EYE}, Y_{EYE}, Z_{EYE}.

Η συντεταγμένη X_{EYE} μπορεί να λάβει τιμές στο εύρος [-left, right]. Αντίστοιχα, η κανονικοποιημένη συντεταγμένη συσκευής X_{NDC} μπορεί να λάβει τιμές στο διάστημα [-1,1]. Η γραμμική σχέση τους υπολογίζεται ως εξής:

$$X_{NDC} = a \cdot X_{EYE} + b \rightarrow \left\{ \begin{matrix} 1 = a \cdot right + b \\ -1 = a \cdot \left(-left\right) + b \end{matrix} \right\} \rightarrow \left\{ \begin{matrix} a = \dfrac{2}{right + left} \\ b = \dfrac{left - right}{right + left} \end{matrix} \right\} \rightarrow$$

$$X_{NDC} = \frac{2}{right + left} \cdot X_{EYE} + \frac{left - right}{right + left} \tag{4.47}$$

Η συντεταγμένη Y_{EYE} μπορεί να λάβει τιμές στο εύρος [-bottom, top], ενώ η κανονικοποιημένη συντεταγμένη συσκευής Y_{NDC} μπορεί να λάβει τιμές στο διάστημα [-1,1]. Η γραμμική σχέση τους υπολογίζεται ως εξής:

$$Y_{NDC} = c \cdot X_{EYE} + d \rightarrow \left\{ \begin{matrix} 1 = c \cdot top + d \\ -1 = c \cdot \left(-bottom\right) + d \end{matrix} \right\} \rightarrow \left\{ \begin{matrix} c = \dfrac{2}{top + bottom} \\ d = \dfrac{bottom - top}{top + bottom} \end{matrix} \right\} \rightarrow$$

$$Y_{NDC} = \frac{2}{top + bottom} \cdot Y_{EYE} + \frac{bottom - top}{top + bottom} \tag{4.48}$$

Τέλος, η συντεταγμένη Z_{EYE} λαμβάνει τιμές στο εύρος [-near, -far], οι οποίες αντιστοιχούν στο εύρος τιμών [-1,1] της κανονικοποιημένης συντεταγμένης συσκευής Z_{NDC}. Η γραμμική σχέση τους υπολογίζεται ως εξής:

$$Z_{NDC} = e \cdot Z_{EYE} + f \rightarrow \left\{ \begin{matrix} -1 = e \cdot \left(-near\right) + f \\ 1 = e \cdot \left(-far\right) + f \end{matrix} \right\} \rightarrow \left\{ \begin{matrix} e = \dfrac{2}{near - far} \\ f = \dfrac{near + far}{near - far} \end{matrix} \right\} \rightarrow$$

$$Z_{NDC} = \frac{2}{near - far} \cdot Z_{EYE} + \frac{near + far}{near - far} \tag{4.49}$$

Επομένως, σύμφωνα με τις σχέσεις (4.44) έως (4.49), η σχέση (4.43) μεταξύ των συντεταγμένων αποκοπής και των συντεταγμένων στο σύστημα της κάμερας με χρήση της ορθογραφικής προβολής γράφεται ως εξής:

$$\begin{bmatrix} X_{CLIP} \\ Y_{CLIP} \\ Z_{CLIP} \\ W_{CLIP} \end{bmatrix} = \begin{bmatrix} \frac{2}{right+left} & 0 & 0 & \frac{left-right}{right+left} \\ 0 & \frac{2}{top+bottom} & 0 & \frac{bottom-top}{top+bottom} \\ 0 & 0 & \frac{2}{near-far} & \frac{near+far}{near-far} \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} X_{EYE} \\ Y_{EYE} \\ Z_{EYE} \\ 1 \end{bmatrix} \quad (4.50)$$

η οποία είναι ισοδύναμη με την ακόλουθη

$$\begin{bmatrix} X_{CLIP} \\ Y_{CLIP} \\ Z_{CLIP} \\ W_{CLIP} \end{bmatrix} = \begin{bmatrix} \frac{2}{width} & 0 & 0 & \frac{left-right}{width} \\ 0 & \frac{2}{height} & 0 & \frac{bottom-top}{height} \\ 0 & 0 & \frac{2}{near-far} & \frac{near+far}{near-far} \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} X_{EYE} \\ Y_{EYE} \\ Z_{EYE} \\ 1 \end{bmatrix} \quad (4.51)$$

Ειδικότερα, στην περίπτωση συμμετρικού όγκου θέασης, η μετατροπή από τις συντεταγμένες στο φωτογραμμετρικό σύστημα, που ορίζεται από την OpenGL, σε συντεταγμένες αποκοπής μέσω ορθογραφικής προβολής λαμβάνει την ακόλουθη μορφή:

$$\begin{bmatrix} X_{CLIP} \\ Y_{CLIP} \\ Z_{CLIP} \\ W_{CLIP} \end{bmatrix} = \begin{bmatrix} \frac{1}{right} & 0 & 0 & 0 \\ 0 & \frac{1}{top} & 0 & 0 \\ 0 & 0 & \frac{2}{near-far} & \frac{near+far}{near-far} \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} X_{EYE} \\ Y_{EYE} \\ Z_{EYE} \\ 1 \end{bmatrix} \quad (4.52)$$

ή ισοδύναμα:

$$\begin{bmatrix} X_{CLIP} \\ Y_{CLIP} \\ Z_{CLIP} \\ W_{CLIP} \end{bmatrix} = \begin{bmatrix} \frac{2}{width} & 0 & 0 & 0 \\ 0 & \frac{2}{height} & 0 & 0 \\ 0 & 0 & \frac{2}{near-far} & \frac{near+far}{near-far} \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} X_{EYE} \\ Y_{EYE} \\ Z_{EYE} \\ 1 \end{bmatrix} \quad (4.53)$$

4.4.2.3 Κανονικοποίηση συντεταγμένων

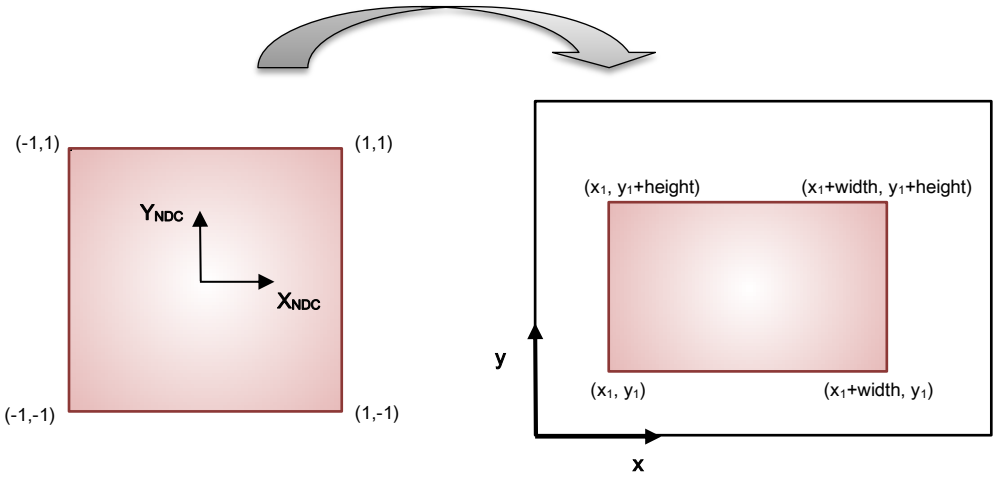
Όπως ήδη αναφέρθηκε, μετά την εφαρμογή του μετασχηματισμού προβολής και την αποκοπή πραγματοποιείται μία διαδικασία κανονικοποίησης των συντεταγμένων σε ένα καθορισμένο εύρος τιμών. Αυτό επιτυγχάνεται μέσω διαίρεσης των συντεταγμένων X_{CLIP}, Y_{CLIP} και Z_{CLIP} με την τέταρτη συντεταγμένη W_{CLIP}, έτσι ώστε οι κανονικοποιημένες

συντεταγμένες να βρίσκονται στο εύρος τιμών [-1,1]. Οι συντεταγμένες αυτές είναι καρτεσιανές και όχι ομογενείς, όπως οι προηγούμενες.

4.4.3 Μετασχηματισμός παραθύρου παρατήρησης και κωδικοποίηση της πληροφορίας βάθους

Το παράθυρο παρατήρησης είναι η ορθογώνια περιοχή στην οποία θα σχεδιαστεί η εικόνα και βρίσκεται εντός του παραθύρου που έχει οριστεί για το σκοπό αυτό. Έτσι, ο μετασχηματισμός παραθύρου παρατήρησης (viewport transformation) έγκειται στον ορισμό της επιφάνειας σχεδίασης εντός ολόκληρης της διαθέσιμης επιφάνειας του παραθύρου και καθορίζει το τελικό μέγεθος του αντικειμένου που θα σχεδιαστεί, ελέγχοντας τη μετατροπή από τις τρισδιάστατες κανονικοποιημένες συντεταγμένες σε συντεταγμένες παραθύρου, σε pixels [177], [185], [186].

Ο μετασχηματισμός μεταξύ των κανονικοποιημένων συντεταγμένων συσκευής και των συντεταγμένων παραθύρου υπολογίζεται με τη θεώρηση της γραμμικής σχέσης με την οποία συνδέονται [187]. Στην Εικόνα 4-10 παρουσιάζεται η έμπροσθεν όψη του κύβου κανονικοποιημένων συντεταγμένων (αριστερά) και το παράθυρο παρατήρησης (δεξιά) και απεικονίζεται η σχέση μεταξύ των κανονικοποιημένων συντεταγμένων συσκευής X_{NDC} και Y_{NDC} με τις συντεταγμένες παραθύρου (x,y).



Εικόνα 4-10. Σχέση μεταξύ συστήματος κανονικοποιημένων συντεταγμένων συσκευής και συστήματος συντεταγμένων παραθύρου.

Τα δύο συστήματα συντεταγμένων που φαίνονται στην Εικόνα 4-10 συνδέονται μεταξύ τους μέσω δύο μεταθέσεων και δύο κλιμάκων κατά τους δύο άξονες. Η μεταξύ τους σχέση μπορεί να γραφεί ως εξής:

$$x = s_x \cdot X_{NDC} + t_x \quad (4.54)$$

$$y = s_y \cdot Y_{NDC} + t_y \quad (4.55)$$

Με δεδομένες τις τέσσερις αντιστοιχίσεις συντεταγμένων στα δύο συστήματα: $X_{NDC} = -1 \longleftrightarrow x = x_1$, $X_{NDC} = 1 \longleftrightarrow x = x_1 + \text{width}$, $Y_{NDC} = -1 \longleftrightarrow y = y_1$ και $Y_{NDC} = 1 \longleftrightarrow y = y_1 + \text{height}$ είναι δυνατή η εύρεση των τεσσάρων συντελεστών. Έτσι, οι σχέσεις (4.54) και (4.55) μετά από την επίλυση του συστήματος των τεσσάρων εξισώσεων λαμβάνουν την ακόλουθη μορφή:

$$x = \frac{\text{width}}{2} \cdot X_{NDC} + x_1 + \frac{\text{width}}{2} \tag{4.56}$$

$$y = \frac{\text{height}}{2} \cdot Y_{NDC} + y_1 + \frac{\text{height}}{2} \tag{4.57}$$

Αν δεν οριστεί από τον προγραμματιστή ο παραπάνω μετασχηματισμός, θεωρείται από την OpenGL ότι το παράθυρο παρατήρησης καταλαμβάνει ολόκληρο το παράθυρο, δηλαδή x_1=0, y_1=0 ενώ width και height είναι οι διαστάσεις του παραθύρου.

Η τρίτη συντεταγμένη (z) κατά το μετασχηματισμό παραθύρου παρατήρησης κωδικοποιείται και φέρει πληροφορία για το βάθος [177]. Εκ προεπιλογής, το εύρος τιμών της συντεταγμένης z θεωρείται από την OpenGL το [0,1]. Προαιρετικά, ο προγραμματιστής μπορεί να ορίσει άλλο εύρος τιμών, το οποίο συνήθως είναι το [near, far]. Θεωρώντας ότι η συντεταγμένη βάθους z κυμαίνεται στο προεπιλεγμένο εύρος τιμών [0,1] και το αντίστοιχο εύρος τιμών της κανονικοποιημένης συντεταγμένης Z_{NDC} είναι το [-1,1], μπορεί να βρεθεί η μεταξύ τους σχέση:

$$z = a \cdot Z_{NDC} + b \;\rightarrow\; \begin{Bmatrix} 0 = a \cdot (-1) + b \\ 1 = a \cdot 1 + b \end{Bmatrix} \rightarrow \begin{Bmatrix} a = 1/2 \\ b = 1/2 \end{Bmatrix} \rightarrow$$

$$z = \frac{1}{2} \cdot Z_{NDC} + \frac{1}{2} \tag{4.58}$$

Θεωρώντας ότι η συντεταγμένη z κυμαίνεται στο ορισμένο από τον προγραμματιστή εύρος τιμών [near, far] και το αντίστοιχο εύρος τιμών της κανονικοποιημένης συντεταγμένης Z_{NDC} είναι το [-1,1] η μεταξύ τους σχέση υπολογίζεται ως εξής:

$$z = a \cdot Z_{NDC} + b \;\rightarrow\; \begin{Bmatrix} \text{near} = a \cdot (-1) + b \\ \text{far} = a \cdot 1 + b \end{Bmatrix} \rightarrow \begin{Bmatrix} a = \dfrac{\text{far} - \text{near}}{2} \\ b = \dfrac{\text{far} + \text{near}}{2} \end{Bmatrix} \rightarrow$$

$$z = \frac{\text{far} - \text{near}}{2} \cdot Z_{NDC} + \frac{\text{far} + \text{near}}{2} \tag{4.59}$$

ΚΕΦΑΛΑΙΟ 5

ΕΦΑΡΜΟΓΕΣ ΕΠΑΥΞΗΜΕΝΗΣ ΠΡΑΓΜΑΤΙΚΟΤΗΤΑΣ

Το περιεχόμενο των προηγούμενων κεφαλαίων σχετίζεται άμεσα με το πέμπτο και τελευταίο κεφάλαιο της παρούσας εργασίας, στο οποίο παρουσιάζονται οι εφαρμογές επαυξημένης πραγματικότητας βάσει επίπεδου προτύπου που αναπτύχθηκαν για τις ανάγκες της διπλωματικής εργασίας. Όλα όσα αναφέρθηκαν ήταν αναγκαία για την κατανόηση της έννοιας της επαυξημένης πραγματικότητας και του θεωρητικού υποβάθρου των σχέσεων και των αλγορίθμων που εφαρμόστηκαν στις εφαρμογές, αλλά και για την ενημέρωση για το περιβάλλον εργασίας, τη γλώσσα προγραμματισμού, τις βιβλιοθήκες που χρησιμοποιήθηκαν και για κάποια σημαντικά για τις εφαρμογές στοιχεία των OpenCV και OpenGL. Έχοντας ολοκληρώσει την παρουσίαση των παραπάνω στοιχείων, ο αναγνώστης κατά τη μελέτη του κεφαλαίου αυτού μπορεί να κατανοήσει σε βάθος τον τρόπο εργασίας κατά τον προγραμματισμό των εφαρμογών, τη μεθοδολογία που εφαρμόστηκε και τη θεωρία και τις μαθηματικές σχέσεις που ενυπάρχουν πίσω από τη διαδικασία που ακολουθήθηκε.

Συγκεκριμένα, στο κεφάλαιο αυτό παρουσιάζονται οι εφαρμογές που αποτέλεσαν το αντικείμενο της παρούσας εργασίας, εκ των οποίων η πρώτη αφορά στην επαύξηση της πραγματικότητας σε εσωτερικό χώρο, ενώ η δεύτερη αφορά στην επαύξηση της πραγματικότητας σε εξωτερικό χώρο. Στην πρώτη ενότητα περιγράφεται η γενική μεθοδολογία η οποία ακολουθήθηκε κατά την ανάπτυξη των εφαρμογών – χωρίς λεπτομέρειες για τον τρόπο περάτωσής τους – ώστε να γίνει αντιληπτή η σχέση τους με τα θεωρητικά στοιχεία που παρουσιάστηκαν στα κεφάλαια 2, 3 και 4 και να δοθεί το γενικό πλάνο που ακολουθήθηκε κατά την υλοποίηση των εφαρμογών. Η δεύτερη ενότητα του κεφαλαίου αυτού αναφέρεται στις βαθμονομήσεις που έγιναν για την κάλυψη των απαιτήσεων των εφαρμογών. Στην τρίτη ενότητα γίνεται αναφορά στο μορφότυπο OBJ, στο οποίο διατέθηκαν τα τρισδιάστατα μοντέλα που συμπληρώνουν τον πραγματικό κόσμο, και ακολουθεί η σύντομη παρουσίαση των μοντέλων αυτών. Τέλος, στην τέταρτη και στην πέμπτη ενότητα, περιγράφονται αναλυτικά οι εφαρμογές που αναπτύχθηκαν, χωρίς έμφαση στις προγραμματιστικές λεπτομέρειες κατά την ανάπτυξή τους, καθώς ο προγραμματισμός

δεν ήταν ο σκοπός της παρούσας εργασίας, αλλά το μέσο συνδυασμού όλων των φωτογραμμετρικών γνώσεων αλλά και μεθοδολογιών επαύξησης της πραγματικότητας για την επίτευξη του τελικού αποτελέσματος.

5.1 Γενική Μεθοδολογία

Οι εφαρμογές επαυξημένης πραγματικότητας που αναπτύχθηκαν στηρίζονται στην αναγνώριση κάποιου χαρακτηριστικού επίπεδου πρότυπου αντικειμένου. Το τελευταίο, ωστόσο, δεν αποτελεί επίπεδο στόχο (marker), αλλά είναι αναπόσπαστο τμήμα της σκηνής που πρόκειται να επαυξηθεί με επιπρόσθετες πληροφορίες, οι οποίες στην προκειμένη περίπτωση είναι τρισδιάστατα μοντέλα. Η γενική μεθοδολογία που ακολουθήθηκε για την υλοποίηση των εφαρμογών επαυξημένης πραγματικότητας βάσει επίπεδου προτύπου περιγράφεται στη συνέχεια.

Πρώτο βήμα αποτελεί ο καθορισμός του τρισδιάστατου μοντέλου και η εισαγωγή του στο πρόγραμμα προκειμένου να επαυξήσει τη θέαση του πραγματικού κόσμου. Εκτός από το μοντέλο, καθορίζεται και το επίπεδο αντικείμενο το οποίο πρέπει να υπάρχει στη σκηνή προκειμένου αυτή να επαυξηθεί με το μοντέλο, το οποίο μπορεί να τοποθετηθεί πάνω, μπροστά και γενικά κοντά στο χαρακτηριστικό αυτό επίπεδο αντικείμενο. Συνεπώς, η εικόνα αυτού του αντικειμένου εισάγεται στο πρόγραμμα. Στο εξής η εικόνα αυτή θα αναφέρεται ως εικόνα - πρότυπο και θα θεωρείται ότι απεικονίζει πλήρως το πρότυπο αντικείμενο και μόνο αυτό. Με βάση αυτή τη θεώρηση, το πρότυπο αντικείμενο είναι τετράπλευρο και οι συντεταγμένες pixel των τεσσάρων κορυφών της εικόνας - πρότυπο αντιστοιχούν στις τέσσερις κορυφές του πρότυπου αντικειμένου. Επιπλέον, τα στοιχεία του εσωτερικού προσανατολισμού της κάμερας, με την οποία γίνεται η λήψη των στιγμιότυπων του πραγματικού κόσμου, αποτελούν επίσης δεδομένα που εισάγονται στο πρόγραμμα. Η είσοδος των τριών αυτών βασικών στοιχείων σε συνδυασμό με την εκτέλεση κάποιων βημάτων, τα οποία θα αναφερθούν στη συνέχεια, αποτελούν τη βάση για την ανάπτυξη εφαρμογών επαυξημένης πραγματικότητας, ανάλογων με αυτών που προγραμματίστηκαν για τις ανάγκες της παρούσας εργασίας.

Στην εικόνα - πρότυπο που εισήχθη στο πρόγραμμα πραγματοποιείται ανίχνευση και περιγραφή σημείων ενδιαφέροντος με κατάλληλο αλγόριθμο. Βασικό συστατικό μίας εφαρμογής επαυξημένης πραγματικότητας είναι ένα μέσο λήψης σκηνών του πραγματικού κόσμου, το οποίο μπορεί να είναι για παράδειγμα η κάμερα του ηλεκτρονικού υπολογιστή. Για κάθε στιγμιότυπο της κάμερας πραγματοποιείται ομοίως ανίχνευση και περιγραφή των σημείων ενδιαφέροντος. Για το στάδιο αυτό, τόσο στην εικόνα - πρότυπο όσο και σε κάθε στιγμιότυπο της κάμερας στο πλαίσιο των εφαρμογών χρησιμοποιήθηκε ο αλγόριθμος SURF, λόγω των καλών επιδόσεών του, τόσο από άποψη ακρίβειας όσο και από άποψη υπολογιστικού χρόνου. Στην επιλογή του SURF συνέβαλε και η ανεξαρτησία του ανιχνευτή

και του περιγραφέα του από μετασχηματισμούς περιστροφής και κλίμακας. Το γεγονός αυτό οδηγεί στον εντοπισμό των ίδιων, κατά το δυνατόν, σημείων ενδιαφέροντος στην εικόνα - πρότυπο και σε κάθε στιγμιότυπο, ανεξάρτητα από την ενδεχόμενη στροφή του τελευταίου σε σχέση με την εικόνα - πρότυπο και από την πιθανή διαφορετική κλίμακα μεταξύ των δύο, και στη μετέπειτα ταύτιση των ομόλογων σημείων ενδιαφέροντος, λόγω της κατάλληλης περιγραφής τους. Εξάλλου, οι περιπτώσεις περιστροφής του στιγμιότυπου σε σχέση με την εικόνα - πρότυπο και διαφορετικής κλίμακας μεταξύ των δύο συναντώνται πολύ συχνά σε μία εφαρμογή επαυξημένης πραγματικότητας. Άλλο επιθυμητό χαρακτηριστικό του SURF είναι η ανεξαρτησία του από μεταβολές φωτισμού και αντίθεσης, οι οποίες ενδέχεται να υφίστανται μεταξύ της εικόνας - πρότυπο και των στιγμιότυπων της κάμερας. Καθώς αυτή η διαδικασία εκτελείται για κάθε στιγμιότυπο, το κριτήριο της ταχύτητας για την επιλογή του αλγορίθμου αυτού καθίσταται πολύ σημαντικό και αποτελεί το βασικό λόγο για τον οποίο δεν χρησιμοποιήθηκε ο υπολογιστικά απαιτητικότερος αλγόριθμος SIFT, μολονότι πληροί τις υπόλοιπες προϋποθέσεις για την επιλογή του ως αλγορίθμου ανίχνευσης και περιγραφής σημείων ενδιαφέροντος.

Μετά από την ανίχνευση και περιγραφή των σημείων ενδιαφέροντος του εκάστοτε στιγμιότυπου της κάμερας μπορεί να πραγματοποιηθεί η ψηφιακή συνταύτιση αυτού και της εικόνας - πρότυπο, μέσω της ταύτισης των διανυσμάτων περιγραφικών χαρακτηριστικών των εντοπισμένων σημείων. Ωστόσο, από την αυτόματη αυτή διαδικασία συχνά προκύπτουν – εκτός από τις σωστές – και λανθασμένες ομολογίες σημείων, οι οποίες πρέπει να εξαλειφθούν. Για το σκοπό αυτό, σε μία εκ των εφαρμογών εφαρμόζεται ο έλεγχος του λόγου των αποστάσεων των εγγύτερων γειτόνων (ratio test), ενώ στην άλλη γίνεται εφαρμογή του διασταυρωμένου ελέγχου (cross-check test), όπως παρουσιάστηκαν στις παραγράφους 2.3.3.1 και 2.3.3.2 αντίστοιχα. Επίσης, για την εξάλειψη των άκυρων ομολογιών εφαρμόζεται ένας επιπρόσθετος έλεγχος, κατά τον οποίο τα σημεία θεωρούνται ομόλογα εφόσον η νόρμα του διανύσματος της διαφοράς των περιγραφικών χαρακτηριστικών των δύο αντίστοιχων σημείων είναι μικρότερη από ένα καθορισμένο κατώφλι.

Μετά από την εύρεση κοινών σημείων στην εικόνα - πρότυπο και στο στιγμιότυπο της κάμερας, μπορεί να υπολογιστεί η σχέση που συνδέει την εικόνα - πρότυπο και το τμήμα του στιγμιότυπου της κάμερας που εικονίζει το πρότυπο αντικείμενο. Η σχέση αυτή είναι μία σχέση ομογραφίας στο δισδιάστατο προβολικό χώρο. Στην υποενότητα 3.4.2 αποδεικνύεται ότι η σχέση του δισδιάστατου προβολικού μετασχηματισμού συνδέει τις εικόνες ενός επίπεδου αντικειμένου. Μολονότι υφίστανται και άλλοι δισδιάστατοι μετασχηματισμοί, η ομογραφία είναι η γενικότερη όλων και έχει εφαρμογή σε όλες τις περιπτώσεις μετασχηματισμού εικονοσυντεταγμένων μεταξύ εικόνων του ίδιου δισδιάστατου αντικειμένου, χωρίς την ανάγκη θεώρησης κάποιας παραδοχής. Λόγω της φύσης των εφαρμογών επαυξημένης πραγματικότητας, δεν είναι δυνατή η θεώρηση κάποιας παραδοχής, όπως

είναι η ύπαρξη μικρών προοπτικών παραμορφώσεων μεταξύ της εικόνας - πρότυπο και του στιγμιότυπου, καθώς το τελευταίο μπορεί να έχει ληφθεί υπό αρκετά διαφορετικές γωνίες ω, φ, κ σε σχέση με την εικόνα - πρότυπο και από διαφορετική θέση – γεγονός που αποτελεί και τη συνήθη περίπτωση – και επιπλέον μπορεί να έχει ληφθεί με ευρυγώνιο φακό και το πρότυπο αντικείμενο να βρίσκεται κοντά στη φωτογραφική μηχανή. Συνεπώς, ο αμέσως προηγούμενος στην ιεραρχία αφινικός μετασχηματισμός δεν θα είχε εφαρμογή (βλ. υποενότητα 3.4.2) και φυσικά ούτε οι υπόλοιποι.

Ο υπολογισμός των στοιχείων του προβολικού αυτού μετασχηματισμού μπορεί να γίνει εφόσον από την προηγούμενη διαδικασία προκύψουν τουλάχιστον τέσσερις ομολογίες σημείων στην εικόνα - πρότυπο και στο στιγμιότυπο της κάμερας. Ωστόσο, αν οι ομολογίες είναι λιγότερες από τέσσερις, δεν είναι δυνατός ο υπολογισμός της ομογραφίας μεταξύ των δύο επιπέδων και η μετέπειτα επαύξηση της σκηνής. Στην περίπτωση των εφαρμογών που αναπτύχθηκαν για την παρούσα εργασία, η εύρεση του προβολικού μετασχηματισμού συνδυάζεται με τον αλγόριθμο RANSAC (βλ. παράγραφο 2.3.3.3) για την ταυτόχρονη απόρριψη των άκυρων ομολογιών, οι οποίες δεν επαληθεύουν τη σχέση ομογραφίας που συνδέει τις δύο εικόνες. Μέσω του αλγορίθμου RANSAC υπολογίζεται ο δισδιάστατος προβολικός μετασχηματισμός από το καλύτερο δείγμα των τεσσάρων ομολογιών αλλά και οι έγκυρες ομολογίες (inliers) για τη συγκεκριμένη λύση. Στη συνέχεια, μέσω του αλγορίθμου βελτιστοποίησης Levenberg-Marquardt υπολογίζεται η τελική καλύτερη εκτίμηση του πίνακα της ομογραφίας, με βάση όλες τις έγκυρες ομολογίες που εντοπίστηκαν από το RANSAC.

Εφόσον ήταν δυνατός ο υπολογισμός των συντελεστών του δισδιάστατου προβολικού μετασχηματισμού, μπορούν να μετασχηματιστούν τα σημεία από το σύστημα συντεταγμένων της εικόνας - προτύπο στο σύστημα του εκάστοτε στιγμιότυπου της κάμερας, αλλά και αντίστροφα. Έτσι, με δεδομένους τους συντελεστές της ομογραφίας, μετασχηματίζονται οι συντεταγμένες pixel των τεσσάρων κορυφών της εικόνας - πρότυπο από το σύστημα συντεταγμένων αυτής στο σύστημα συντεταγμένων του στιγμιότυπου της κάμερας. Με βάση τη διαδικασία αυτή εντοπίζονται οι τέσσερις κορυφές της εικόνας - πρότυπο στο εκάστοτε στιγμιότυπο.

Σύμφωνα με όσα έχουν περιγραφεί έως τώρα, είναι γνωστά τα στοιχεία του εσωτερικού προσανατολισμού της κάμερας και οι συντεταγμένες pixel των τεσσάρων κορυφών του επίπεδου πρότυπου αντικειμένου στο εκάστοτε στιγμιότυπο. Για τον υπολογισμό του εξωτερικού προσανατολισμού του συγκεκριμένου στιγμιότυπου απαιτείται επιπλέον η γνώση των συντεταγμένων των κορυφών του πρότυπου αντικειμένου στο επίγειο σύστημα αναφοράς. Επομένως, ορίζεται ένα επίγειο σύστημα αναφοράς και καθορίζονται οι επίγειες συντεταγμένες των κορυφών του πρότυπου αντικειμένου στο σύστημα αυτό. Έτσι, καθίσταται δυνατός ο υπολογισμός του εξωτερικού προσανατολισμού του στιγμιότυπου, με βάση τα στοιχεία του εσωτερικού προσανατολισμού και τις τέσσερις κορυφές του πρότυπου

επίπεδου αντικειμένου στο επίγειο σύστημα συντεταγμένων και στο σύστημα συντεταγμένων του στιγμιότυπου της κάμερας, σε pixels. Αυτά που περιγράφηκαν έως τώρα γίνονται με χρήση της βιβλιοθήκης OpenCV. Η τελευταία χρησιμοποιείται για να εξάγει τις παραπάνω πληροφορίες από τα στιγμιότυπα της κάμερας, για τον υπολογισμό του εξωτερικού προσανατολισμού κάθε στιγμιότυπου, ενώ η OpenGL χρησιμοποιείται για το σχεδιασμό της επαυξημένης σκηνής.

Μέσω της OpenGL, γίνεται η σχεδίαση του στιγμιότυπου της κάμερας ως φόντου σε παράθυρο που δημιουργείται από τη βιβλιοθήκη freeglut για την επαύξηση της πραγματικότητας. Η διαδικασία αυτή πραγματοποιείται μέσω μίας τεχνικής απόδοσης υφής (texture mapping), η οποία περιγράφεται στην παράγραφο 5.4.3.3 του παρόντος κεφαλαίου. Τελευταίο και κρισιμότερο στοιχείο της εφαρμογής αποτελεί η σχεδίαση του τρισδιάστατου μοντέλου στο εν λόγω παράθυρο στην κατάλληλη θέση, κλίμακα και προοπτική και με το σωστό προσανατολισμό. Προκειμένου να επιτευχθεί αυτό, ακολουθείται η διαδικασία που περιγράφεται στη συνέχεια.

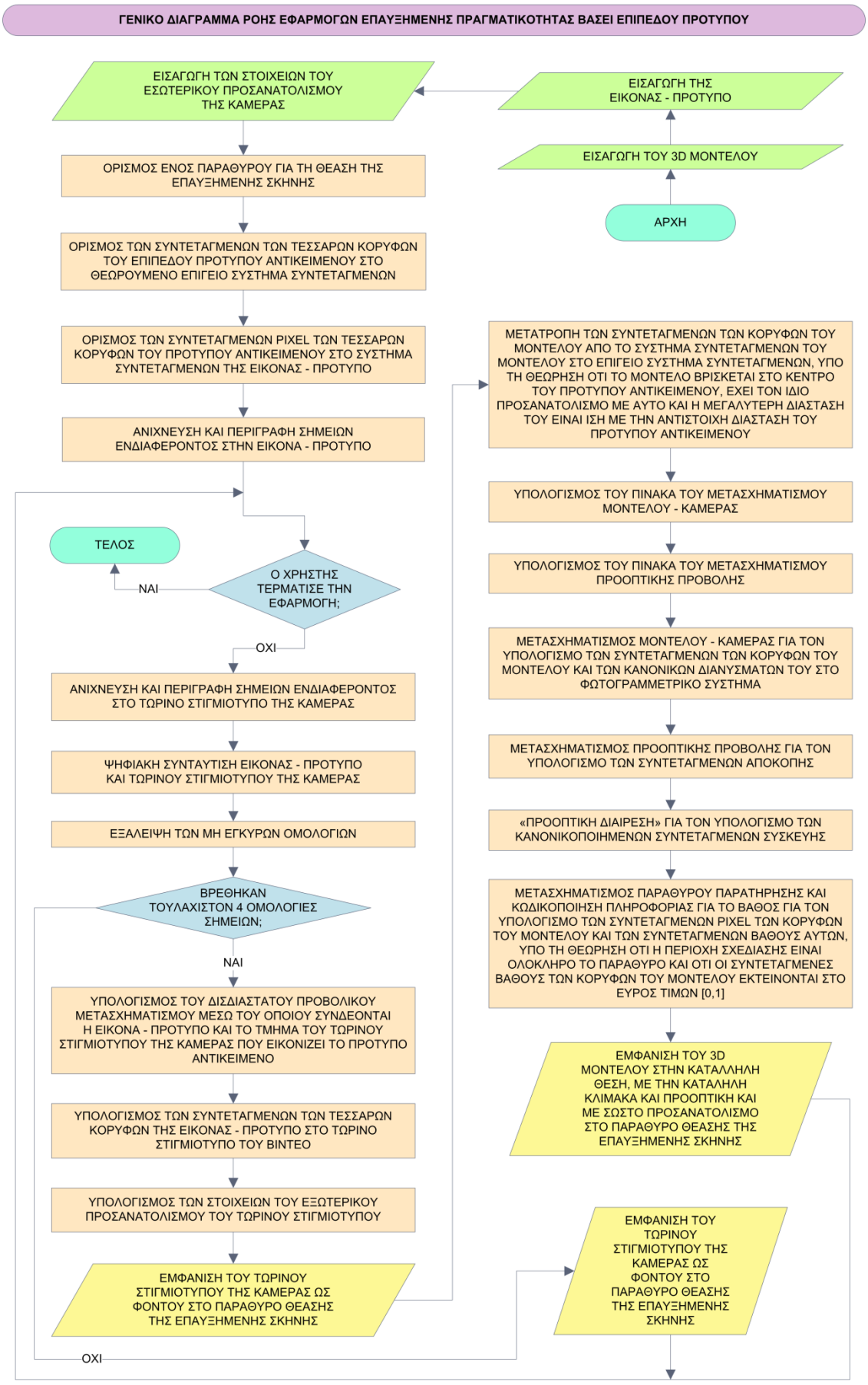
Τα στοιχεία του εξωτερικού προσανατολισμού ορίζουν το μετασχηματισμό κάμερας (viewing transformation), ο οποίος μετασχηματίζει τις συντεταγμένες του μοντέλου από το επίγειο στο φωτογραμμετρικό σύστημα συντεταγμένων, όπως το τελευταίο ορίζεται στην OpenGL (βλ. υποενότητα 4.4.1). Ο πίνακας μοντέλου - κάμερας, που πρέπει να οριστεί στην OpenGL, μετασχηματίζει τις συντεταγμένες από το σύστημα του μοντέλου στο φωτογραμμετρικό σύστημα συντεταγμένων. Συνεπώς, για την κατασκευή του πίνακα του μετασχηματισμού μοντέλου - κάμερας (modelview matrix) από τα στοιχεία του εξωτερικού προσανατολισμού του στιγμιότυπου, απαραίτητη προϋπόθεση είναι η ταύτιση του συστήματος συντεταγμένων του μοντέλου με το επίγειο σύστημα συντεταγμένων. Ωστόσο, οι συντεταγμένες μοντέλου μπορεί να αναφέρονται σε ένα οποιοδήποτε σύστημα συντεταγμένων. Για το λόγο αυτό, γίνεται η μετατροπή των συντεταγμένων των κορυφών του μοντέλου από το σύστημα του μοντέλου στο επίγειο σύστημα συντεταγμένων, θεωρώντας ότι το κέντρο του μοντέλου ταυτίζεται με το κέντρο του επίγειου συστήματος συντεταγμένων, ότι ο προσανατολισμός του είναι ίδιος με τον προσανατολισμό του πρότυπου αντικειμένου και ότι η μεγαλύτερη διάσταση του μοντέλου είναι ίση με τη μεγαλύτερη διάσταση του πρότυπου αντικειμένου. Στις συντεταγμένες αυτές εφαρμόζεται στη συνέχεια ο μετασχηματισμός μοντέλου - κάμερας. Ωστόσο, αν δεν είναι επιθυμητή η τοποθέτηση του μοντέλου στο κέντρο του πρότυπου αντικειμένου, η ταύτιση του προσανατολισμού του με τον προσανατολισμό του πρότυπου αντικειμένου και η ισότητα της μεγαλύτερης διάστασής του με τη μεγαλύτερη διάσταση του πρότυπου αντικειμένου, ο πίνακας του μετασχηματισμού μοντέλου - κάμερας είναι το γινόμενο του πίνακα που δημιουργείται από τα στοιχεία του εξωτερικού προσανατολισμού του στιγμιότυπου και του κατάλληλου πίνακα μετάθεσης - στροφής - αλλαγής κλίμακας.

Επίσης, με βάση τα στοιχεία του εσωτερικού προσανατολισμού της κάμερας και τις διαστάσεις της εικόνας και με επιλογή του κοντινού και του μακρινού επιπέδου αποκοπής, υπολογίζεται ο πίνακας του μετασχηματισμού προοπτικής προβολής της OpenGL (βλ. παράγραφο 4.4.2.1). Η προοπτική προβολή επιλέγεται αντί της ορθογραφικής λόγω της ρεαλιστικής απεικόνισης του τρισδιάστατου αντικειμένου στη σκηνή μέσω αυτής.

Από τους δύο αυτούς πίνακες, δηλαδή τον πίνακα του μετασχηματισμού μοντέλου - κάμερας και τον πίνακα του μετασχηματισμού προοπτικής προβολής, είναι δυνατός ο υπολογισμός των συντεταγμένων pixel των κορυφών του τρισδιάστατου μοντέλου στο στιγμιότυπο – με βάση τις μετασχηματισμένες του συντεταγμένες στο επίγειο σύστημα αναφοράς – αλλά και η αποθήκευση πληροφορίας για το βάθος κάθε κορυφής, για την επισήμανση των κορυφών οι οποίες θα φαίνονται στο τελικό παράθυρο θέασης της επαυξημένης σκηνής. Η διαδικασία αυτή γίνεται αυτόματα από την OpenGL. Συγκεκριμένα, μέσω του πίνακα του μετασχηματισμού μοντέλου - κάμερας υπολογίζονται οι συντεταγμένες των κορυφών του τρισδιάστατου μοντέλου στο φωτογραμμετρικό σύστημα, όπως ορίζεται στην OpenGL, με βάση τις συντεταγμένες αυτών στο σύστημα του μοντέλου. Με δεδομένο τον πίνακα του μετασχηματισμού προοπτικής προβολής υπολογίζονται οι ομογενείς συντεταγμένες αποκοπής, ενώ μέσω «προοπτικής διαίρεσης», δηλαδή διαίρεσης των τριών πρώτων με την τέταρτη ομογενή συντεταγμένη, υπολογίζονται οι κανονικοποιημένες συντεταγμένες συσκευής. Τέλος, επειδή δεν ορίζεται κάποιος μετασχηματισμός παραθύρου παρατήρησης ούτε ένα εύρος για τις τιμές του βάθους, ο μετασχηματισμός παραθύρου παρατήρησης και η κωδικοποίηση της πληροφορίας για το βάθος γίνονται από την OpenGL υπό τη θεώρηση ότι το μοντέλο μπορεί να σχεδιαστεί σε ολόκληρο το παράθυρο και ότι οι συντεταγμένες βάθους έχουν εύρος τιμών [0,1].

Τελικά, με βάση τα παραπάνω, το τρισδιάστατο μοντέλο σχεδιάζεται με το σωστό τρόπο στο παράθυρο επαύξησης της σκηνής, το φόντο του οποίου είναι το εκάστοτε στιγμιότυπο της κάμερας. Το διάγραμμα ροής της διαδικασίας που περιγράφηκε παραπάνω εικονίζεται στη Διαγραμματική παρουσίαση 5-1.

Οι εφαρμογές επαυξημένης πραγματικότητας βάσει επίπεδου προτύπου που αναπτύχθηκαν σύμφωνα με τη μεθοδολογία που αναφέρθηκε προγραμματίστηκαν ώστε να αποτελούν εφαρμογές κονσόλας Win32 (Win32 console applications), δηλαδή προγράμματα τα οποία αλληλεπιδρούν με το χρήστη μέσω ενός περιβάλλοντος χρήστη που χρησιμοποιεί μόνο κείμενο (text user interface) και προορίζονται για το λειτουργικό σύστημα των Windows.



Διαγραμματική παρουσίαση 5-1. Γενικό διάγραμμα ροής των εφαρμογών επαυξημένης πραγματικότητας βάσει επίπεδου προτύπου.

5.2 Βαθμονόμηση

Όπως ήδη αναφέρθηκε, τα στοιχεία του εσωτερικού προσανατολισμού της κάμερας πρέπει να εισαχθούν στο πρόγραμμα προκειμένου να επιτευχθεί η σωστή επαύξηση των σκηνών. Για το λόγο αυτό, πρέπει να προηγηθεί η διαδικασία της βαθμονόμησης της κάμερας που λαμβάνει τα στιγμιότυπα του πραγματικού κόσμου. Στο πλαίσιο αυτό, πραγματοποιείται βαθμονόμηση της κάμερας ενός κινητού τηλεφώνου Samsung Galaxy Y Pro young GT-B5510, με λειτουργικό σύστημα Android, αλλά και της κάμερας ενός φορητού υπολογιστή TURBO-X μοντέλου M765TU με χρήση ελεύθερου κώδικα που διανέμεται μαζί με τη βιβλιοθήκη OpenCV και βρίσκεται στο αρχείο calibration.cpp.

Το αρχείο πηγαίου κώδικα calibration.cpp υπέστη τη διαδικασία που περιγράφεται στην υποενότητα 4.2.1 ώστε να μετατραπεί σε εκτελέσιμο αρχείο (.exe) μέσω του περιβάλλοντος ανάπτυξης Microsoft Visual C++ Express 2010. Με χρήση του εκτελέσιμου αρχείου πραγματοποιήθηκε η διαδικασία της βαθμονόμησης. Για τη διεξαγωγή της τελευταίας απαιτήθηκε το άνοιγμα του παραθύρου εντολών (command window) στο φάκελο στον οποίο αποθηκεύτηκε η εφαρμογή calibraition.exe και η πληκτρολόγηση της κατάλληλης εντολής για την πραγματοποίηση της βαθμονόμησης, όπως θα παρουσιαστεί στη συνέχεια.

Εικόνα 5-1. Εικόνες σκακιέρας από το κινητό τηλέφωνο Samsung Galaxy Y Pro young GT-B5510 που χρησιμοποιήθηκαν για τη βαθμονόμηση της κάμεράς του.

Η διαδικασία της βαθμονόμησης με χρήση επίπεδης σκακιέρας, που πραγματοποιείται μέσω του εκτελέσιμου αρχείου calibration.exe, είναι εντελώς αυτόματη και οι παρεμβάσεις του χρήστη περιορίζονται στη λήψη εικόνων ή βίντεο της σκακιέρας και στον καθορισμό ορισμένων παραμέτρων. Η βαθμονόμηση της κάμερας του κινητού τηλεφώνου έγινε με λήψη εικόνων ενώ η κάμερα του υπολογιστή βαθμονομήθηκε μέσω λήψης βίντεο.

Συγκεκριμένα, για τη βαθμονόμηση μέσω λήψης εικόνων της επίπεδης σκακιέρας, αρχικό βήμα αποτελεί η φωτογράφησή της, από διαφορετικές αποστάσεις και υπό διαφορετικές οπτικές γωνίες. Για το σκοπό αυτό δημιουργήθηκε, μέσω χρήσης του σχεδιαστικού προγράμματος AutoCAD, ένα σχέδιο σκακιέρας, το οποίο τυπώθηκε σε χαρτί A4 και φωτογραφήθηκε από την κάμερα του κινητού. Οι φωτογραφίες οι οποίες λήφθηκαν παρουσιάζονται στην Εικόνα 5-1. Κατά τη λήψη των εικόνων δόθηκε προσοχή στην ταύτιση της ανάλυσής τους με την ανάλυση του βίντεο που εγγράφεται από το κινητό. Καθώς ο αριθμός των εικονοστοιχείων είναι 320 x 240 στην περίπτωση λήψης βίντεο από το συγκεκριμένο κινητό, μεταβλήθηκαν και οι ρυθμίσεις λήψης των φωτογραφιών ώστε ο αριθμός των εικονοστοιχείων τους να είναι ομοίως 320 x 240. Αυτό έγινε διότι η μεταβολή της ανάλυσης των εικόνων, μολονότι δεν επηρεάζει τους συντελεστές της διαστροφής, επηρεάζει τη σταθερά της μηχανής και τις συντεταγμένες του πρωτεύοντος σημείου.

```
<?xml version="1.0"?>
<opencv_storage><images>
GTB5510_images/img1.jpg
GTB5510_images/img2.jpg
GTB5510_images/img3.jpg
GTB5510_images/img4.jpg
GTB5510_images/img5.jpg
GTB5510_images/img6.jpg
GTB5510_images/img7.jpg
GTB5510_images/img8.jpg
GTB5510_images/img9.jpg
GTB5510_images/img10.jpg
GTB5510_images/img11.jpg
GTB5510_images/img12.jpg
GTB5510_images/img13.jpg
GTB5510_images/img14.jpg
</images></opencv_storage>
```

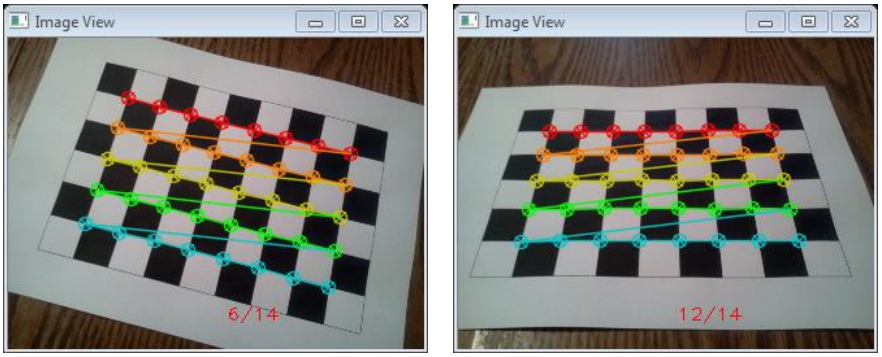
Εικόνα 5-2. Περιεχόμενο αρχείου Imagelist_GTB5510.xml.

Οι εικόνες με ονομασίες img1, img2, …, img14 αποθηκεύτηκαν σε ένα φάκελο «GTB5510_images», ο οποίος τοποθετήθηκε στον ίδιο φάκελο στον οποίο βρίσκεται και η εφαρμογή calibration.exe. Επιπλέον, δημιουργήθηκε ένα XML αρχείο με την ονομασία

Imagelist_GTB5510.xml, που περιλαμβάνει πληροφορίες για τις εικόνες της σκακιέρας που θα χρησιμοποιηθούν στη βαθμονόμηση. Το περιεχόμενο του αρχείου αυτού φαίνεται στην Εικόνα 5-2.

Το αρχείο αυτό πρέπει να εισαχθεί στην εφαρμογή βαθμονόμησης προκειμένου αυτή να εντοπίσει τις παραπάνω εικόνες. Η εντολή η οποία πληκτρολογείται στο παράθυρο εντολών, το οποίο έχει ανοίξει στο φάκελο που βρίσκεται η εφαρμογή, είναι η εξής: «calibration.exe Imagelist_GTB5510_creator Imagelist_GTB5510.xml *.jpg calibration -w 8 -h 5 -s 0.025 -o calibration_output_GTB5510.yml -oe -su Imagelist_GTB5510.xml». Μέσω της εντολής αυτής, το πρόγραμμα μπορεί να εντοπίσει τις εικόνες που τοποθετήθηκαν στο φάκελο GTB5510_images, μέσω των πληροφοριών του αρχείου Imagelist_GTB5510xml, ενώ ενημερώνεται ότι οι εσωτερικές γωνίες της σκακιέρας είναι 8 κατά τη μία διεύθυνση και 5 κατά την άλλη. Επίσης, το μέγεθος του κάθε τετραγώνου της σκακιέρας θεωρείται 0.025m, γεγονός το οποίο επηρεάζει μόνο τις παραμέτρους του εξωτερικού προσανατολισμού και όχι τα στοιχεία του εσωτερικού προσανατολισμού. Μέσω των προθεμάτων -oe, και -su επισημαίνεται ότι στο αρχείο με τα αποτελέσματα της βαθμονόμησης θα συμπεριληφθούν και τα στοιχεία του εξωτερικού προσανατολισμού κάθε λήψης και ότι μετά τη διεξαγωγή της βαθμονόμησης θα παρουσιαστούν και οι εικόνες που διορθώθηκαν από την επίδραση της διαστροφής. Επίσης, μέσω της παραπάνω εντολής δηλώνεται το όνομα του αρχείου που θα περιλαμβάνει το αποτέλεσμα της βαθμονόμησης («calibration_output_GTB5510.yml»).

Κατά τη διεξαγωγή της διαδικασίας της βαθμονόμησης, όπως παρουσιάστηκε στην υποενότητα 3.5.2, ανιχνεύονται οι εσωτερικές γωνίες της σκακιέρας. Στην Εικόνα 5-3 απεικονίζονται ενδεικτικά δύο εικόνες των οποίων έχουν εντοπιστεί οι εσωτερικές γωνίες μέσω της εφαρμογής της βαθμονόμησης που υλοποιήθηκε με χρήση της βιβλιοθήκης OpenCV.



Εικόνα 5-3. Εύρεση των εσωτερικών γωνιών της σκακιέρας κατά τη διαδικασία της βαθμονόμησης.

Το αποτελέσματα που προκύπτουν από την παραπάνω διαδικασία είναι ο πίνακας της μηχανής, οι τρεις συντελεστές του πολυωνύμου της ακτινικής διαστροφής k_1, k_2 και k_3, οι δύο συντελεστές του πολυωνύμου της ασύμμετρης διαστροφής εκκεντρότητας p_1 και p_2, το σφάλμα επαναπροβολής για κάθε εικόνα (συνολικά 14 σφάλματα επαναπροβολής), το μέσο

σφάλμα επαναπροβολής και ο εξωτερικός προσανατολισμός κάθε εικόνας. Αναφορικά με τον τελευταίο, το πρόγραμμα βαθμονόμησης υπολογίζει τα τρία στοιχεία του διανύσματος μετάθεσης και τα τρία στοιχεία του διανύσματος στροφής που αντιστοιχούν στη μετάθεση και στροφή του θεωρούμενου επίγειου συστήματος συντεταγμένων για να ταυτιστεί με το φωτογραμμετρικό, όπως ορίζεται στη βιβλιοθήκη OpenCV. Τα κυριότερα αποτελέσματα που προέκυψαν από τη διαδικασία της βαθμονόμησης της κάμερας του κινητού τηλεφώνου παρουσιάζονται στον Πίνακα 5-1.

Αποτελέσματα βαθμονόμησης της κάμερας του κινητού τηλεφώνου				
Σταθερά μηχανής		**Διαστροφή**		
c_x (pixels)	**301.3**	**Συμμετρική ακτινική διαστροφή**	**k_1**	**0.1585658**
c_y (pixels)	**302.8**		**k_2**	**-0.7155924**
Πρωτεύον σημείο			**k_3**	**1.1343407**
x_0 (pixels)	**154.8**	**Ασύμμετρη διαστροφή εκκεντρότητας**	**p_1**	**-0.0013607**
y_0 (pixels)	**119.2**		**p_2**	**-0.0002139**
Σφάλμα επαναπροβολής (pixels)		**0.15**		

Πίνακας 5-1. Αποτελέσματα βαθμονόμησης της κάμερας του κινητού τηλεφώνου.

Εκτός από την κάμερα του κινητού, βαθμονομήθηκε και η κάμερα ενός φορητού υπολογιστή με το ίδιο πρόγραμμα, μέσω λήψης βίντεο της σκακιέρας. Η εντολή που πληκτρολογήθηκε στο παράθυρο εντολών είναι η ακόλουθη: «calibration.exe -w 8 -h 5 -d 2000 -s 0.025 -o laptop_calibration.yml -oe». Η παραπάνω εντολή δηλώνει βαθμονόμηση μέσω στιγμιότυπων της κάμερας του υπολογιστή σε πραγματικό χρόνο. Μέσω αυτής υποδηλώνεται ότι ο αριθμός των εσωτερικών γωνιών της σκακιέρας είναι 8 κατά τη μία διάσταση και 5 κατά την άλλη και ότι ο ελάχιστος χρόνος που πρέπει να μεσολαβήσει μεταξύ δύο διαδοχικών προσπαθειών του προγράμματος να εντοπίσει τη σκακιέρα είναι δύο δευτερόλεπτα. Επίσης, μέσω της ίδιας εντολής επισημαίνεται ότι το μέγεθος των τετραγώνων της σκακιέρας είναι 0.025 σε μονάδες επίγειου συστήματος αναφοράς – κάτι το οποίο δεν επηρεάζει τα στοιχεία του εσωτερικού προσανατολισμού – ότι το όνομα του αρχείου εξόδου με τα αποτελέσματα της βαθμονόμησης είναι «laptop_calibration.yml» και ότι στο αρχείο αυτό είναι επιθυμητή η παράθεση και των στοιχείων του εξωτερικού προσανατολισμού κάθε λήψης. Για την περάτωση της βαθμονόμησης λήφθηκαν δέκα στιγμιότυπα του βίντεο. Στην Εικόνα 5-4 παρατίθεται ένα στιγμιότυπο από την αναγνώριση της σκακιέρας (αριστερά) και ένα στιγμιότυπο που χρησιμοποιήθηκε για τη βαθμονόμηση (δεξιά).

Εικόνα 5-4. Στιγμιότυπα από τη διαδικασία της βαθμονόμησης της κάμερας του φορητού υπολογιστή. Αναγνώριση της σκακιέρας [αριστερά] και λήψη του στιγμιότυπου για χρήση στη διαδικασία βαθμονόμησης [δεξιά].

Τα κυριότερα αποτελέσματα που προέκυψαν από τη διαδικασία της βαθμονόμησης της κάμερας του φορητού υπολογιστή TURBO-X παρουσιάζονται στον Πίνακα 5-2, ενώ στο αρχείο με τα αποτελέσματα της βαθμονόμησης αποθηκεύονται τα ίδια στοιχεία που αναγράφηκαν και προηγουμένως για τη βαθμονόμηση της κάμερας του κινητού τηλεφώνου.

Αποτελέσματα βαθμονόμησης της κάμερας του φορητού υπολογιστή				
Σταθερά μηχανής		**Διαστροφή**		
c_x (pixels)	**651.5**	**Συμμετρική ακτινική διαστροφή**	**k_1**	**0.1453706**
c_y (pixels)	**654.3**		**k_2**	**-0.4381366**
Πρωτεύον σημείο			**k_3**	**0.2061427**
x_0 (pixels)	**315.4**	**Ασύμμετρη διαστροφή εκκεντρότητας**	**p_1**	**-0.0192786**
y_0 (pixels)	**225.6**		**p_2**	**-0.0012197**
Σφάλμα επαναπροβολής (pixels)		**0.31**		

Πίνακας 5-2. Αποτελέσματα βαθμονόμησης της κάμερας του φορητού ηλεκτρονικού υπολογιστή.

5.3 Μορφή Αρχείων OBJ

Το OBJ είναι μία μορφή αρχείου που ορίζει τη γεωμετρία ενός αντικειμένου και αναπτύχθηκε από την εταιρία γραφικών υπολογιστή Wavefront Technologies για το πακέτο κίνησης τρισδιάστατων γραφικών Advanced Visualizer. Τα αρχεία OBJ συναντώνται σε μορφή ASCII με κατάληξη .obj και μπορούν να τροποποιηθούν με έναν απλό επεξεργαστή κειμένου, αλλά και σε δυαδική μορφή υπό την κατάληξη .mod. Για το σκοπό των εφαρμογών χρησιμοποιήθηκαν αρχεία κειμένου κατάληξης .obj, τα οποία περιγράφονται στην ενότητα

αυτή. Τα τελευταία αναπαριστούν την τρισδιάστατη γεωμετρία ενός αντικειμένου ενσωματώνοντας πληροφορίες για τη θέση κάθε κορυφής, αλλά και κάθε κορυφής υφής (texture coordinate vertex), για τα κανονικά διανύσματα (normals) και για τα πρόσωπα (faces), τα οποία ορίζουν κάθε πολύγωνο ως μία λίστα κορυφών, συντεταγμένων υφής και κανονικών διανυσμάτων. Στον Πίνακα 5-3 παρουσιάζονται τα κυριότερα στοιχεία της γεωμετρίας που ορίζεται σε ένα αρχείο OBJ και η συνήθης σύνταξή τους. Τα περισσότερα από τα στοιχεία αυτά συναντώνται και στα τρία αρχεία που διατέθηκαν για τις ανάγκες των εφαρμογών.

Μία γεωμετρική κορυφή σε ένα αρχείο OBJ συναντάται με τη σύνταξη «**v** x y z» όπου τα x, y, z είναι πραγματικοί αριθμοί που δηλώνουν τις τρισδιάστατες συντεταγμένες της κορυφής. Για ρητές καμπύλες (rational curves) και επιφάνειες προστίθεται και μία τέταρτη ομογενής συντεταγμένη w. Μία κορυφή υφής δηλώνεται με τη σύνταξη «**vt** u v» όπου u είναι η τιμή για την οριζόντια διεύθυνση της υφής και v η τιμή της κατακόρυφης διεύθυνσης της υφής, η οποία έχει νόημα στην περίπτωση δισδιάστατης υφής. Στις περιπτώσεις τρισδιάστατης υφής προστίθεται μία τρίτη συντεταγμένη w που δηλώνει το βάθος της υφής. Ένα κανονικό διάνυσμα ορίζεται στο αρχείο με τη μορφή «**vn** i j k» όπου i, j, k είναι οι συντεταγμένες των κανονικών διανυσμάτων κατά τους τρεις άξονες.

Τα πρόσωπα μπορεί να οριστούν μέσω των γεωμετρικών κορυφών και δηλώνονται στο αρχείο OBJ με τη μορφή «**f** v_1 v_2 … v_n», όπου τα v_1, v_2, …, v_n είναι αριθμοί εντός του εύρους τιμών [1, k] και k είναι ο αριθμός των γεωμετρικών κορυφών του μοντέλου. Ο αριθμός 1 αντιστοιχεί στην πρώτη γεωμετρική κορυφή που παρουσιάστηκε στο αρχείο, ο 2 στη δεύτερη κ.ο.κ. Επίσης, ο ορισμός των προσώπων μπορεί να περιλαμβάνει και συντεταγμένες υφής. Στην περίπτωση αυτή, τα πρόσωπα συναντώνται στο αρχείο με τη σύνταξη «**f** v_1/t_1 v_2/t_2 … v_n/t_n», όπου ο αριθμός t_i δηλώνει τη συντεταγμένη υφής που αντιστοιχεί στην κορυφή v_i και λαμβάνει τιμές στο εύρος [1, m], όπου m ο αριθμός των κορυφών υφής. Ο ορισμός του προσώπου αν γίνει μέσω χρήσης των κανονικών διανυσμάτων λαμβάνει τη μορφή «**f** $v_1//n_1$ $v_2//n_2$ … $v_n//n_n$», όπου n_i οι αριθμοί που αντιστοιχούν στο κανονικό διάνυσμα της γεωμετρικής κορυφής, με την ίδια λογική όπως παραπάνω. Τέλος, ένα πρόσωπο μπορεί να περιλαμβάνει όλες τις ανωτέρω πληροφορίες, οπότε ορίζεται ως εξής: «**f** $v_1/t_1/n_1$ $v_2/t_2/n_2$ … $v_n/t_n/n_n$». Οι γεωμετρικές κορυφές, οι κορυφές υφής και τα κανονικά διανύσματα αντί να ορίζονται μέσω αριθμών που δηλώνουν την απόλυτη θέση τους στη συγκεκριμένη ομάδα του αρχείου, μπορούν να οριστούν και μέσω αρνητικών αριθμών που δείχνουν τη σχετική θέση σε σχέση με το συγκεκριμένο πρόσωπο. Έτσι, ο αριθμός -1 για παράδειγμα δηλώνει την αμέσως προηγούμενη γεωμετρική κορυφή ή κορυφή υφής ή το αμέσως προηγούμενο κανονικό διάνυσμα του συγκεκριμένου προσώπου. Επίσης, το αρχείο OBJ μπορεί να δηλώσει σημεία («**p** v_1») τα οποία ορίζονται από μία γεωμετρική κορυφή, και γραμμές («**l** v_1 v_2 … v_n»), που ορίζονται από n γεωμετρικές κορυφές οι οποίες ενώνονται μεταξύ τους.

Αναπαράσταση της τρισδιάστατης γεωμετρίας μέσω αρχείου OBJ	
Δεδομένα κορυφών	
Γεωμετρική κορυφή	**v** x y z
Κορυφή υφής	**vt** u v
Κανονικό διάνυσμα κορυφής	**vn** i j k
Στοιχεία	
Σημείο	**p** v_1
Γραμμή	**l** v_1 v_2 … v_n
Πρόσωπο	**f** v_1 v_2 … v_n
Πρόσωπο με συντεταγμένες υφής	**f** v_1/t_1 v_2/t_2 … v_n/t_n
Πρόσωπο με κανονικά διανύσματα	**f** $v_1//n_1$ $v_2//n_2$ … $v_n//n_n$
Πρόσωπο με συν/νες υφής και κανονικά διανύσματα	**f** $v_1/t_1/n_1$ $v_2/t_2/n_2$ … $v_n/t_n/n_n$
Ομαδοποίηση	
Όνομα ομάδας	**g** group_name
Χαρακτηριστικά	
Όνομα υλικού	**usemtl** metarial_name
Όνομα αρχείου υλικού MTL	**mtlib** mtl_name.mtl

Πίνακας 5-3. Αναπαράσταση της τρισδιάστατης γεωμετρίας μέσω αρχείου OBJ.

Το αρχείο OBJ υποστηρίζει τη δυνατότητα ομαδοποίησης των στοιχείων. Έτσι, η σύνταξη «**g** group_name» δηλώνει το όνομα της ομάδας των στοιχείων που ακολουθούν, για την καλύτερη οργάνωση των στοιχείων του αρχείου και τη διευκόλυνση του χειρισμού των στοιχείων που αποτελούν το μοντέλο. Εκτός της παραπάνω ομαδοποίησης, υποστηρίζονται οι ομάδες εξομάλυνσης, οι ομάδες συγχώνευσης και η ανάθεση ονόματος σε ένα αντικείμενο του αρχείου.

Επίσης, το αρχείο OBJ υποστηρίζει και υλικά παραπέμποντας σε ένα ή και περισσότερα εξωτερικά αρχεία υλικού, που έχουν την ονομασία MTL (Material Template Library), μέσω της εντολής «**mtlib** mtl_name.mtl». Στο αρχείο OBJ μέσω της εντολής «**usemtl** material_name» ορίζεται το όνομα του υλικού για το στοιχείο που ακολουθεί. Το όνομα του υλικού δηλώνεται στο αρχείο MTL. Το τελευταίο μπορεί να ορίσει διάφορα υλικά μέσω της λέξης - κλειδί «**newmtl**» ακολουθούμενης από το όνομα του υλικού. Μετά από τη δήλωση του ονόματος του υλικού ακολουθεί η περιγραφή του στο αρχείο αυτό. Επίσης, το αρχείο MTL μπορεί να περιέχει αναφορά σε εξωτερικό υλικό υφής (textured material), δηλαδή σε μία εικόνα, μέσω της εντολής **map_kd** ακολουθούμενης από το όνομα του αρχείου της εικόνας [188], [189], [190], [191].

Τα τρισδιάστατα μοντέλα επαύξησης των εφαρμογών παρουσιάζονται στην Εικόνα 5-5. Το ψηφιακό μοντέλο εδάφους (ΨΜΕ) που χρησιμοποιείται στην εφαρμογή επαυξημένης

πραγματικότητας εσωτερικού χώρου (Εφαρμογή 1) ορίζεται σε ένα αρχείο OBJ που συνοδεύεται από ένα αρχείο MTL, το οποίο περιέχει πληροφορία για μία εικόνα που αποτελεί την «υφή» του. Το ίδιο συμβαίνει και με το τρισδιάστατο μοντέλο του αγγείου που χρησιμοποιείται στην εφαρμογή επαυξημένης πραγματικότητας εξωτερικού χώρου (Εφαρμογή 2). Συνεπώς, και τα δύο μοντέλα περιλαμβάνουν συντεταγμένες υφής που αναφέρονται σε σημεία των δύο εικόνων που τα συνοδεύουν. Αντίθετα, το μοντέλο του αγάλματος που χρησιμοποιείται στην Εφαρμογή 2 ορίζεται σε ένα μόνο αρχείο OBJ και δε συνοδεύεται από κάποιο αρχείο MTL, μη περιέχοντας έτσι πληροφορίες για κάποιο υλικό ή για εξωτερική εικόνα και μη διαθέτοντας συντεταγμένες υφής.

Εικόνα 5-5. Τρισδιάστατα μοντέλα που χρησιμοποιήθηκαν στις εφαρμογές. ΨΜΕ [αριστερά], αγγείο [κέντρο] και άγαλμα [δεξιά].

5.4 Εφαρμογή 1 - Επαυξημένη Πραγματικότητα σε Εσωτερικό Χώρο

Στην εφαρμογή αυτή, το τρισδιάστατο μοντέλο που επαυξάνει την πραγματική σκηνή είναι ένα ψηφιακό μοντέλο εδάφους (ΨΜΕ) της περιοχής διαστάσεων περίπου 18km x 12km εκατέρωθεν της τεχνητής λίμνης του ποταμού Λάδωνα στο νομό Αρκαδίας της Πελοποννήσου και διατέθηκε από το Εργαστήριο Φωτογραμμετρίας της Σχολής Αγρονόμων και Τοπογράφων Μηχανικών του Εθνικού Μετσόβιου Πολυτεχνείου για την εκπόνηση της παρούσας εργασίας. Το επίπεδο πρότυπο αντικείμενο βάσει του οποίου γίνεται η επαύξηση της σκηνής είναι μία ορθοφωτογραφία της ίδιας περιοχής. Το ΨΜΕ επικάθεται στην ορθοφωτογραφία και κάθε φορά που αυτή βρίσκεται στο οπτικό πεδίο της κάμερας, ο χρήστης έχει τη δυνατότητα να δει αντί αυτής, το ΨΜΕ της στη θέση που βρίσκεται αυτή και με τον ίδιο προσανατολισμό και μέγεθος. Έτσι, δίνεται στο χρήστη η εντύπωση ότι δεν υπάρχει η ορθοφωτογραφία στην πραγματική σκηνή, αλλά το ΨΜΕ αυτής και του δίνεται η δυνατότητα να εξετάσει το ανάγλυφο της περιοχής. Συνεπώς, η πληροφορία η οποία

προστίθεται είναι αυτή της τρίτης διάστασης. Παράλληλα, δεν αφαιρείται κάποια πληροφορία, καθώς το ΨΜΕ έχει ως υφή την ορθοφωτογραφία, η οποία επεξεργάστηκε ώστε να αυξηθεί η αντίθεσή της, για την επίτευξη καλύτερου οπτικού αποτελέσματος.

Ο χρήστης της εφαρμογής μπορεί να χρησιμοποιήσει την κάμερα του υπολογιστή του για τη θέαση της επαυξημένης σκηνής σε πραγματικό χρόνο, να εισαγάγει κάποιο αρχείο βίντεο, το περιεχόμενο του οποίου θα επαυξηθεί μέσω της εφαρμογής, ή να παρακολουθήσει μία προεπισκόπηση της εφαρμογής με την επαύξηση των σκηνών ενός προεπιλεγμένου βίντεο. Επίσης, του δίνεται η δυνατότητα να δει εφόσον το επιλέξει – εκτός από την επαυξημένη σκηνή – τη βασική διαδικασία που πραγματοποιείται με χρήση της βιβλιοθήκης OpenCV προκειμένου να επιτευχθεί το τελικό αποτέλεσμα της εφαρμογής. Ακόμη, έχει την επιλογή να εισαγάγει τα στοιχεία του εσωτερικού προσανατολισμού της κάμερας του υπολογιστή του ή της κάμερας με την οποία πραγματοποιήθηκε η λήψη του βίντεο ή να χρησιμοποιήσει τα προεπιλεγμένα στοιχεία του εσωτερικού προσανατολισμού που προέκυψαν από τη βαθμονόμηση της κάμερας του φορητού υπολογιστή TURBO-X μοντέλου M765TU (στην περίπτωση χρήσης της κάμερας του υπολογιστή) ή του κινητού τηλεφώνου Samsung Galaxy Y Pro young GT-B5510 με λειτουργικό σύστημα Android (στην περίπτωση επαύξησης των σκηνών ενός αρχείου βίντεο). Γενικά, αν ο χρήστης δεν πληκτρολογήσει τα στοιχεία του εσωτερικού προσανατολισμού της κάμερας και χρησιμοποιήσει διαφορετική κάμερα από τις παραπάνω για τις οποίες έγινε η βαθμονόμηση, το αποτέλεσμα δεν θα είναι το επιθυμητό, καθώς θα έχει γίνει λάθος εκτίμηση των στοιχείων του εξωτερικού προσανατολισμού του στιγμιότυπου και επιπλέον το τρισδιάστατο μοντέλο θα φαίνεται με λανθασμένη προοπτική, λόγω της μη σωστής κατασκευής του πίνακα του προοπτικού μετασχηματισμού. Στην περίπτωση αυτή, το αποτέλεσμα της εφαρμογής θα είναι η μη ρεαλιστική επαύξηση της πραγματικότητας.

Μία εφαρμογή επαυξημένης πραγματικότητας, σύμφωνα με τον ορισμό που έδωσε ο Azuma (βλ. ενότητα 1.1) πρέπει να συνδυάζει το πραγματικό αντικείμενο με το εικονικό σε πραγματικό περιβάλλον, να επιτρέπει την αλληλεπίδραση σε πραγματικό χρόνο και να εγγράφει τα εικονικά αντικείμενα στον τρισδιάστατο χώρο. Η εφαρμογή αυτή συνδυάζει και τα τρία αυτά χαρακτηριστικά στην περίπτωση που η επαύξηση της σκηνής γίνει σε πραγματικό χρόνο μέσω της κάμερας του υπολογιστή, με αποτέλεσμα να μπορεί χωρίς αμφιβολία να χαρακτηριστεί ως μία πραγματική εφαρμογή επαυξημένης πραγματικότητας. Ωστόσο, στην περίπτωση εισαγωγής κάποιου αρχείου βίντεο δεν τηρείται το δεύτερο χαρακτηριστικό, καθώς το βίντεο λήφθηκε σε προγενέστερη χρονική στιγμή, με αποτέλεσμα ο συνδυασμός του πραγματικού με το εικονικό αντικείμενο να μη γίνεται σε πραγματικό χρόνο. Στην περίπτωση επαύξησης των σκηνών ενός αρχείου βίντεο, η τοποθέτηση του τρισδιάστατου μοντέλου στη σωστή θέση, με τη σωστή κλίμακα και προοπτική και με τον κατάλληλο προσανατολισμό δεν οφείλεται σε κάποιου είδους προεπεξεργασία, αλλά αντίθετα ακολουθείται η ίδια διαδικασία που ακολουθείται και στην περίπτωση της

επαύξησης της πραγματικότητας σε πραγματικό χρόνο μέσω της κάμερας του υπολογιστή. Ωστόσο, σύμφωνα με τον αυστηρό αυτό ορισμό η επαύξηση των σκηνών του βίντεο δεν χαρακτηρίζεται ως εφαρμογή επαυξημένης πραγματικότητας καθώς δεν ισχύει το δεύτερο χαρακτηριστικό. Στο πλαίσιο της παρούσας εργασίας το είδος αυτό της επαύξησης θα χαρακτηρίζεται ως επαυξημένη πραγματικότητα, καθώς συγκαταλέγεται στη γενική περίπτωση τοποθέτησης επιπρόσθετων εικονικών πληροφοριών στον πραγματικό κόσμο χωρίς κάποιου είδους χειροκίνητη προεπεξεργασία. Αντίθετα, όλη η διαδικασία είναι πλήρως αυτοματοποιημένη.

Η ονομασία «επαυξημένη πραγματικότητα σε εσωτερικό χώρο» ή «indoor augmented reality» επιλέχτηκε για να διαφοροποιηθεί η εφαρμογή αυτή από τη δεύτερη, η οποία απαιτεί αρχεία βίντεο και τιμές GPS που λήφθηκαν σε εξωτερικό χώρο. Ωστόσο, ο χαρακτηρισμός αυτός δεν αποκλείει τη χρήση της Εφαρμογής 1 και σε εξωτερικό χώρο, στην περίπτωση που η ορθοφωτογραφία βρίσκεται σε εξωτερικό χώρο. Στην ενότητα αυτή περιγράφονται τα βήματα που ακολουθήθηκαν κατά τον προγραμματισμό της εφαρμογής και παρουσιάζονται τα αποτελέσματά της.

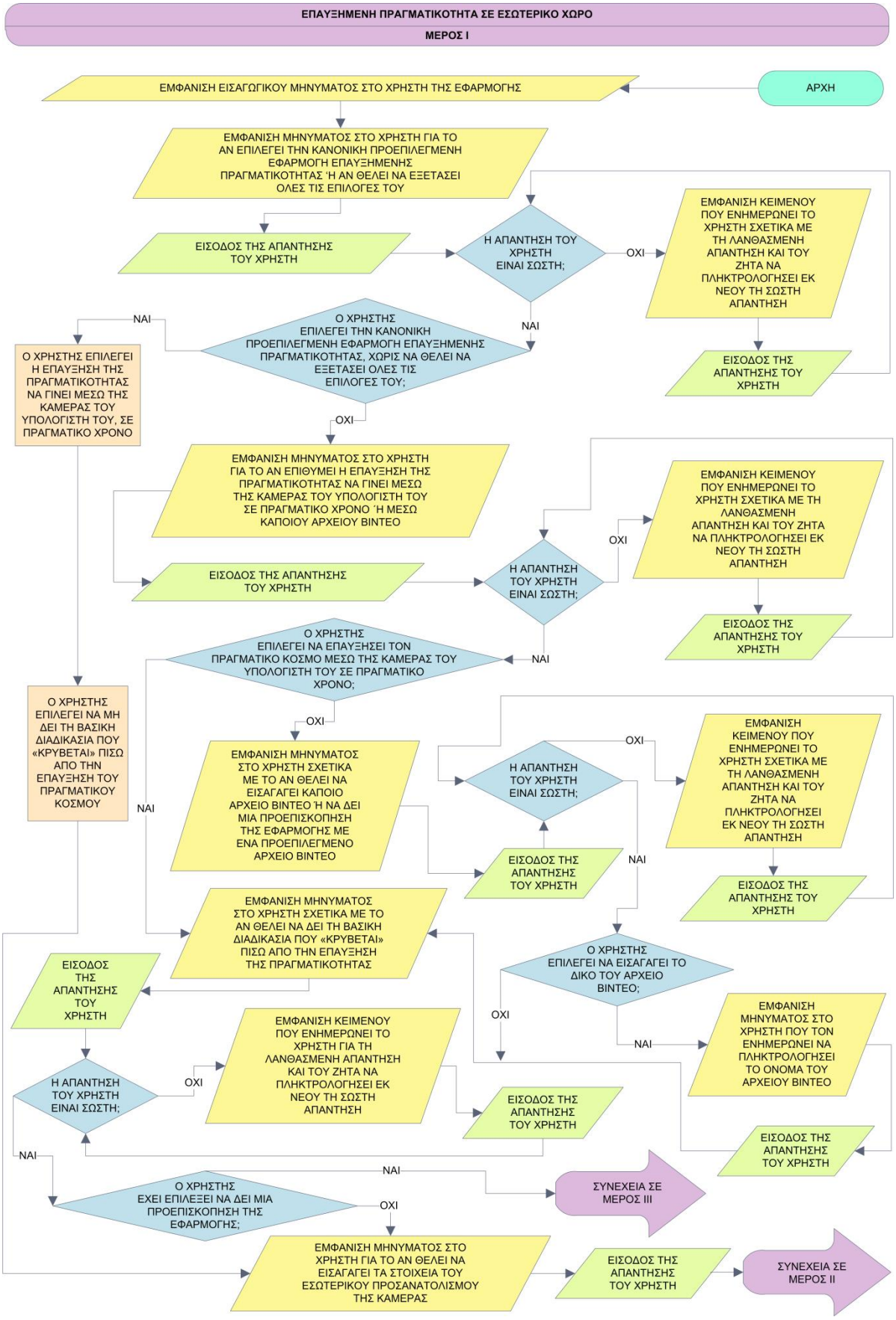
5.4.1 Αρχικά βήματα για τον καθορισμό των επιλογών του χρήστη

Αρχικό βήμα της εφαρμογής αποτελεί ο καθορισμός των επιλογών του χρήστη. Για το σκοπό αυτό, στο παράθυρο της εφαρμογής κονσόλας Win32 εμφανίζονται διαδοχικά μηνύματα και ο χρήστης καλείται να δηλώσει τις επιλογές του μέσω πληκτρολόγησης των κατάλληλων χαρακτήρων. Παρακάτω παρουσιάζονται όλοι οι πιθανοί συνδυασμοί επιλογών του χρήστη.

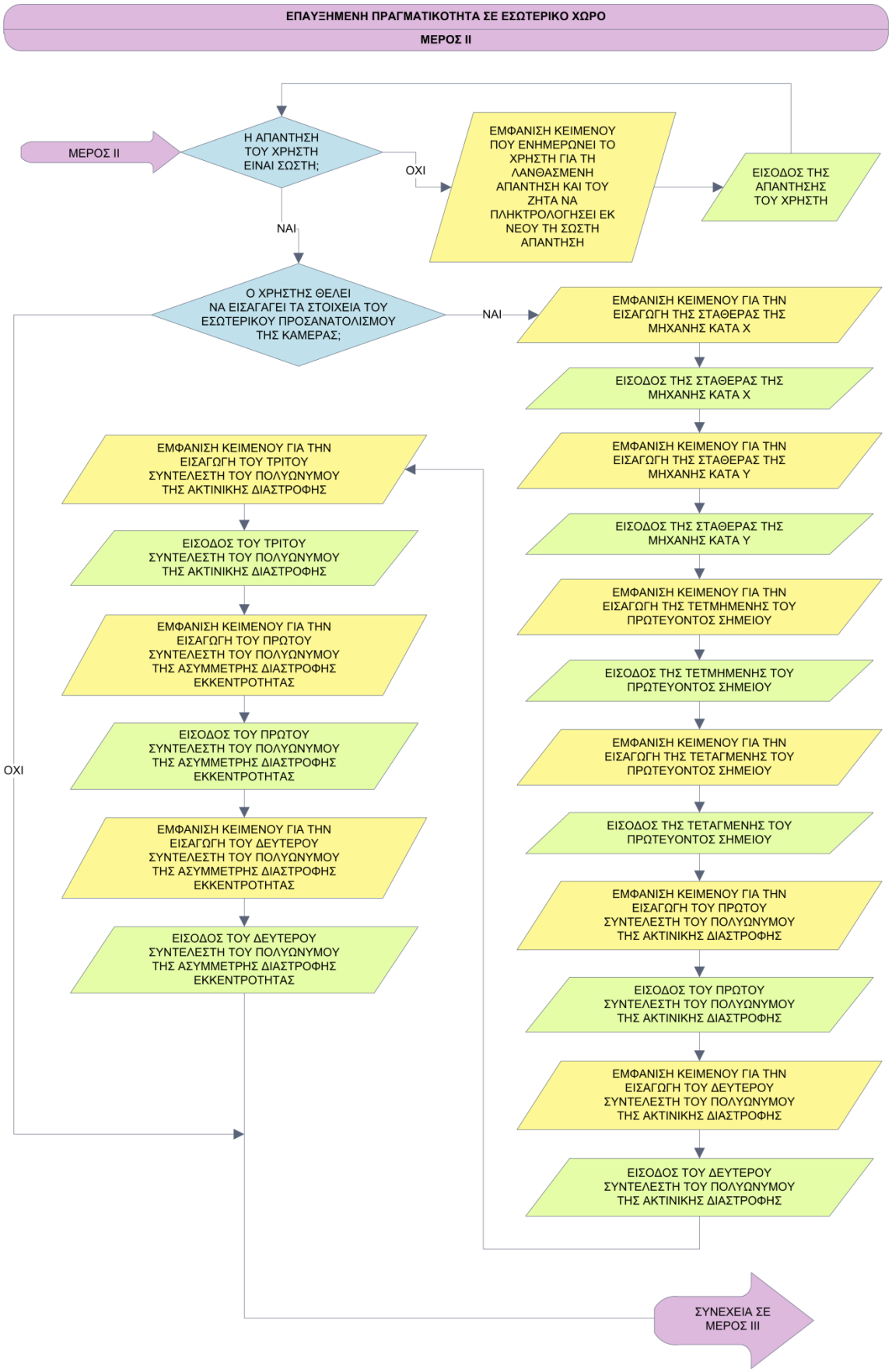
- Επαύξηση της πραγματικότητας σε πραγματικό χρόνο μέσω λήψης στιγμιότυπων από την κάμερα του υπολογιστή, εμφάνιση μόνο του παραθύρου απεικόνισης των επαυξημένων σκηνών, πληκτρολόγηση των στοιχείων του εσωτερικού προσανατολισμού της κάμερας από το χρήστη.
- Επαύξηση της πραγματικότητας σε πραγματικό χρόνο μέσω λήψης στιγμιότυπων από την κάμερα του υπολογιστή, εμφάνιση μόνο του παραθύρου απεικόνισης των επαυξημένων σκηνών, μη εισαγωγή των στοιχείων του εσωτερικού προσανατολισμού της κάμερας από το χρήστη.
- Επαύξηση της πραγματικότητας σε πραγματικό χρόνο μέσω λήψης στιγμιότυπων από την κάμερα του υπολογιστή, εμφάνιση όχι μόνο του παραθύρου επαύξησης της πραγματικότητας αλλά και παραθύρων που παρουσιάζουν τη βασική διαδικασία που διεξάγεται μέσω της βιβλιοθήκης OpenCV, πληκτρολόγηση των στοιχείων του εσωτερικού προσανατολισμού της κάμερας από το χρήστη.

- Επαύξηση της πραγματικότητας σε πραγματικό χρόνο μέσω λήψης στιγμιότυπων από την κάμερα του υπολογιστή, εμφάνιση όχι μόνο του παραθύρου επαύξησης της πραγματικότητας αλλά και παραθύρων που παρουσιάζουν τη βασική διαδικασία που διεξάγεται μέσω της βιβλιοθήκης OpenCV, μη εισαγωγή των στοιχείων του εσωτερικού προσανατολισμού της κάμερας από το χρήστη.
- Επαύξηση των σκηνών ενός αρχείου βίντεο που εισάγεται από το χρήστη, εμφάνιση μόνο του παραθύρου απεικόνισης των επαυξημένων σκηνών, πληκτρολόγηση των στοιχείων του εσωτερικού προσανατολισμού της κάμερας με την οποία έγινε η λήψη του βίντεο από το χρήστη.
- Επαύξηση των σκηνών ενός αρχείου βίντεο που εισάγεται από το χρήστη, εμφάνιση όχι μόνο του παραθύρου επαύξησης της πραγματικότητας αλλά και παραθύρων που παρουσιάζουν τη βασική διαδικασία που διεξάγεται μέσω της βιβλιοθήκης OpenCV, πληκτρολόγηση των στοιχείων του εσωτερικού προσανατολισμού της κάμερας με την οποία έγινε η λήψη του βίντεο από το χρήστη.
- Επαύξηση των σκηνών ενός αρχείου βίντεο που εισάγεται από το χρήστη, εμφάνιση μόνο του παραθύρου απεικόνισης των επαυξημένων σκηνών, μη εισαγωγή των στοιχείων του εσωτερικού προσανατολισμού της κάμερας με την οποία έγινε η λήψη του βίντεο από το χρήστη.
- Επαύξηση των σκηνών ενός αρχείου βίντεο που εισάγεται από το χρήστη, εμφάνιση όχι μόνο του παραθύρου επαύξησης της πραγματικότητας αλλά και παραθύρων που παρουσιάζουν τη βασική διαδικασία που διεξάγεται μέσω της βιβλιοθήκης OpenCV, μη εισαγωγή των στοιχείων του εσωτερικού προσανατολισμού της κάμερας με την οποία έγινε η λήψη του βίντεο από το χρήστη.
- Επαύξηση των σκηνών ενός προεπιλεγμένου αρχείου βίντεο, εμφάνιση μόνο του παραθύρου απεικόνισης των επαυξημένων σκηνών.
- Επαύξηση των σκηνών ενός προεπιλεγμένου αρχείου βίντεο, εμφάνιση όχι μόνο του παραθύρου επαύξησης της πραγματικότητας αλλά και παραθύρων που παρουσιάζουν τη βασική διαδικασία που διεξάγεται μέσω της βιβλιοθήκης OpenCV.

Οι επιλογές του χρήστη ελέγχονται για την ορθότητά τους και σε περίπτωση λανθασμένης επιλογής ο χρήστης καλείται εκ νέου να πληκτρολογήσει το σωστό χαρακτήρα. Στις Διαγραμματικές παρουσιάσεις 5-2 και 5-3 φαίνεται η διαδικασία καθορισμού των επιλογών του χρήστη με τη μορφή διαγραμμάτων ροής (βλ. Παράρτημα Α για τα χρησιμοποιούμενα σύμβολα). Στο Παράρτημα Γ, το οποίο αποτελεί εγχειρίδιο χρήσης της Εφαρμογής 1, εικονίζονται τα μηνύματα που εμφανίζονται κατά την εκτέλεση της εφαρμογής, αναφέρονται οδηγίες για τη σωστή εισαγωγή των επιλογών του χρήστη, ώστε να γίνει κατανοητός ο τρόπος αλληλεπίδρασης του χρήστη με την εφαρμογή, και παρουσιάζονται κάποιες λεπτομέρειες σχετικές με τον τρόπο λειτουργίας της.



Διαγραμματική παρουσίαση 5-2. Εφαρμογή 1, μέρος Ι.



Διαγραμματική παρουσίαση 5-3. Εφαρμογή 1, μέρος ΙΙ.

5.4.2 Ενδιάμεσα βήματα πριν την επαναληπτική διαδικασία επαύξησης της σκηνής κάθε στιγμιότυπου

Αφού καθοριστούν οι επιλογές του χρήστη, είναι δυνατή η συνέχιση της διαδικασίας για τη διεκπεραίωση των βημάτων που πρέπει να διεξαχθούν πριν ξεκινήσει η επαναληπτική διαδικασία για την επαύξηση της σκηνής κάθε στιγμιότυπου.

Επόμενο βήμα μετά τον καθορισμό των επιλογών του χρήστη αποτελεί η δημιουργία του 3x3 πίνακα της μηχανής, με τα στοιχεία του εσωτερικού προσανατολισμού της κάμερας, και του 5x1 διανύσματος με τους συντελεστές του πολυωνύμου της ακτινικής διαστροφής (k_1, k_2, k_3) και της ασύμμετρης διαστροφής εκκεντρότητας (p_1, p_2). Ο πίνακας της μηχανής δίνεται από την εξίσωση (5.1), θεωρώντας ότι οι πλευρές των εικονοστοιχείων είναι κάθετες μεταξύ τους (s=0), ενώ το διάνυσμα των συντελεστών της διαστροφής δίνεται από τη σχέση (5.2). Η σειρά με την οποία αποθηκεύονται οι συντελεστές στο διάνυσμα είναι η αναγραφόμενη στην εν λόγω σχέση, λόγω της σειράς με την οποία τους δέχονται οι συναρτήσεις της βιβλιοθήκης OpenCV προκειμένου να χρησιμοποιηθούν για τον υπολογισμό άλλων στοιχείων, όπως των παραμέτρων του εξωτερικού προσανατολισμού.

$$\mathbf{K} = \begin{bmatrix} c_x & 0 & x_0 \\ 0 & c_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{5.1}$$

$$\mathbf{v}_{\text{διαστρ.}} = \begin{bmatrix} k_1 & k_2 & p_1 & p_2 & k_3 \end{bmatrix}^T \tag{5.2}$$

Ο πίνακας της μηχανής και το διάνυσμα με τους συντελεστές της διαστροφής δημιουργούνται από:

- τα στοιχεία που δίνει ο χρήστης, εφόσον επέλεξε να πληκτρολογήσει τα στοιχεία του εσωτερικού προσανατολισμού της κάμερας, ή
- τα στοιχεία του εσωτερικού προσανατολισμού της κάμερας του φορητού υπολογιστή TURBO-X, όπως προέκυψαν από τη βαθμονόμηση που πραγματοποιήθηκε, εφόσον δεν έγινε εισαγωγή των στοιχείων του εσωτερικού προσανατολισμού από το χρήστη και ο τελευταίος επέλεξε τη ζωντανή επαύξηση της πραγματικότητας μέσω της κάμερας του υπολογιστή, ή
- τα στοιχεία του εσωτερικού προσανατολισμού της κάμερας του κινητού τηλεφώνου Samsung, όπως προέκυψαν από τη βαθμονόμηση που πραγματοποιήθηκε, εφόσον δεν έγινε εισαγωγή των στοιχείων του εσωτερικού προσανατολισμού από το χρήστη και ο τελευταίος επέλεξε την επαύξηση των σκηνών ενός αρχείου βίντεο.

Ακολουθεί η αποθήκευση των διαστάσεων του βίντεο ή των στιγμιότυπων της κάμερας του υπολογιστή, για τον καθορισμό του μεγέθους του παραθύρου επαύξησης της πραγματικότητας. Για το λόγο αυτό, διαβάζεται το πρώτο στιγμιότυπο του βίντεο ή της

κάμερας του υπολογιστή. Ανάλογα με τις προηγούμενες επιλογές του χρήστη, ανοίγει η κάμερα του υπολογιστή ή το αρχείο βίντεο που εισήχθη από το χρήστη ή το προεπιλεγμένο αρχείο βίντεο. Στην περίπτωση που δεν άνοιξε επιτυχώς, εμφανίζεται κατάλληλο προειδοποιητικό μήνυμα για τον τερματισμό της εφαρμογής (βλ. Παράρτημα Γ).

Αν η κάμερα του υπολογιστή ή το αρχείο βίντεο άνοιξε επιτυχώς, ακολουθούν τα βήματα που απαιτούνται για την αρχικοποίηση και τη δημιουργία ενός παραθύρου για την επαύξηση των σκηνών μέσω της βιβλιοθήκης freeglut. Πρώτο βήμα αποτελεί η αρχικοποίηση της βιβλιοθήκης αυτής και ακολουθεί ο ορισμός διαφόρων παραμέτρων για την απεικόνιση της επαυξημένης σκηνής στο παράθυρο που θα δημιουργηθεί. Οι παράμετροι που ορίζονται είναι η χρήση του χρωματικού μοντέλου RGB (κόκκινο - πράσινο - μπλε) για την απεικόνιση της σκηνής και η χρήση διπλού «buffer» («double buffering»). Σύμφωνα με το χρωματικό μοντέλο RGB, κάθε χρώμα αποτελεί γραμμικό συνδυασμό των τριών βασικών χρωμάτων (κόκκινου, πράσινου και μπλε), ενώ η χρήση του διπλού «buffer» ορίζει τη χρήση δύο «buffers», κατά την οποία ο ένας απεικονίζεται στο χρήστη όσο η σκηνή σχεδιάζεται στο δεύτερο. Με τη χρήση αυτής της τεχνικής εξασφαλίζεται ότι δεν θα απεικονιστεί μία ημιτελώς σχεδιασμένη σκηνή, αλλά θα εμφανιστεί μόνο όταν είναι πλήρως σχεδιασμένη, με αποτέλεσμα να μην παρατηρείται «τρεμόπαιγμα» της σκηνής σε περίπτωση κίνησης [60], [192]. Στη συνέχεια, καθορίζεται το μέγεθος του παραθύρου επαύξησης της πραγματικότητας (σε pixels), το οποίο τίθεται ίσο με το μέγεθος του πρώτου στιγμιότυπου της κάμερας του υπολογιστή ή του αρχείου βίντεο. Επίσης, πριν τη δημιουργία του παραθύρου καθορίζεται η θέση του στην οθόνη. Για τον καθορισμό της τελευταίας, ορίζεται η απόσταση σε pixels της επάνω αριστερής κορυφής του παραθύρου από την επάνω αριστερή κορυφή της οθόνης, κατά τις δύο διευθύνσεις. Τέλος, δημιουργείται το παράθυρο, με τίτλο «AR Window», το οποίο θα είναι υπεύθυνο για την απεικόνιση της επαυξημένης σκηνής.

Επόμενο βήμα αποτελεί η αρχικοποίηση της βιβλιοθήκης GLEW. Στην περίπτωση που αυτή δεν είναι επιτυχής, εμφανίζεται κατάλληλο μήνυμα που ειδοποιεί για τον επερχόμενο τερματισμό της εφαρμογής (βλ. Παράρτημα Γ).

Στη συνέχεια, ανάλογα με την προηγούμενη επιλογή του χρήστη για επαύξηση της πραγματικότητας μέσω της κάμερας του υπολογιστή ή για θέαση των επαυξημένων σκηνών ενός αρχείου βίντεο, καθορίζεται η συνάρτηση η οποία θα καλείται από τη βιβλιοθήκη freeglut κάθε φορά που απαιτείται ο σχεδιασμός της σκηνής και η συνάρτηση η οποία θα καλείται συνεχώς, όταν δεν λαμβάνονται «γεγονότα» (events) προς επεξεργασία. Ο σχεδιασμός της σκηνής απαιτείται να γίνει την πρώτη φορά που δημιουργείται το παράθυρο αλλά και κάθε φορά που αυτό «κρύβεται» από κάποιο άλλο και επανέρχεται στο προσκήνιο. Τα «γεγονότα» μπορεί να είναι το πάτημα πλήκτρων του πληκτρολογίου, το πάτημα ή η απελευθέρωση των κουμπιών του ποντικιού, η κίνησή του όταν κάποιο από τα κουμπιά του

είναι πατημένο, η μεταβολή του μεγέθους του παραθύρου και γενικά οι ενέργειες αλληλεπίδρασης με τις μονάδες εισόδου [193]. Στο πλαίσιο της εφαρμογής, καθορίζεται η ίδια συνάρτηση η οποία θα καλείται όταν απαιτείται επανασχεδιασμός της σκηνής και συνεχώς όταν δεν λαμβάνονται «γεγονότα» προς επεξεργασία και είναι η συνάρτηση η οποία επαυξάνει την πραγματικότητα. Ωστόσο, η τελευταία διαφέρει ανάλογα με το αν ο χρήστης επέλεξε επαύξηση των σκηνών του βίντεο ή των ζωντανών στιγμιότυπων της κάμερας. Με τον τρόπο αυτό, η διαδικασία που ακολουθείται για την επαύξηση της πραγματικότητας θα εκτελείται συνεχώς, για κάθε νέο στιγμιότυπο του βίντεο ή της ζωντανής μετάδοσης της κάμερας του υπολογιστή. Οι συναρτήσεις αυτές παρουσιάζονται αναλυτικά στην υποενότητα 5.4.3.

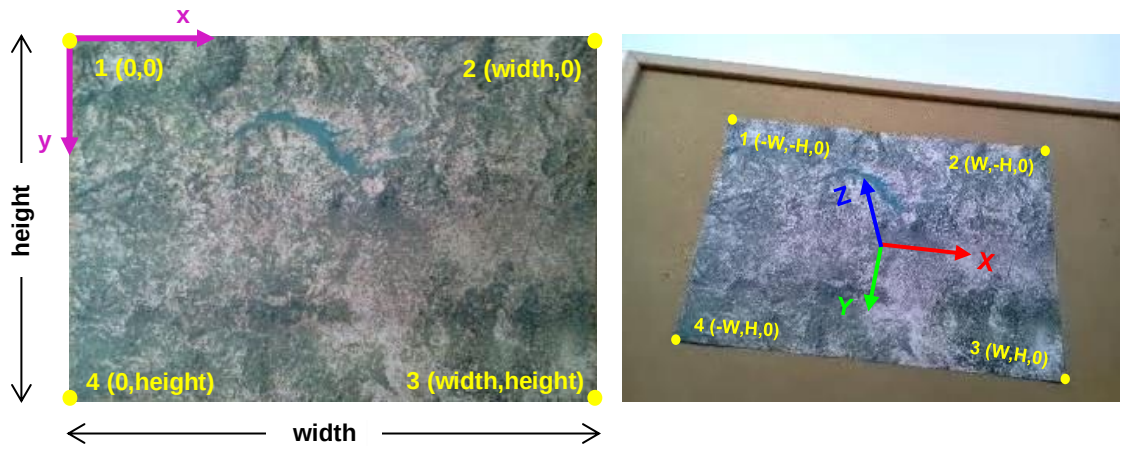
Επίσης, καθορίζεται η συνάρτηση η οποία θα καλείται κάθε φορά που ο χρήστης πατά κάποιο πλήκτρο του πληκτρολογίου. Σύμφωνα με τη συνάρτηση αυτή, επιλέγεται ο τερματισμός της εφαρμογής αν ο χρήστης πατήσει το πλήκτρο «Escape». Επιπλέον, καθορίζεται η συνάρτηση που θα καλείται κάθε φορά που ο χρήστης μεταβάλλει το μέγεθος του παραθύρου «AR Window». Η συνάρτηση αυτή διατηρεί το λόγο των πλευρών του στιγμιότυπου σταθερό και ίσο με τον αρχικό λόγο αυτών, ώστε να μην παραμορφώνεται η σκηνή για να καταλάβει όλο το διαθέσιμο εύρος του παραθύρου του οποίου μεταβλήθηκε το μέγεθος. Αν δεν οριστεί μία συνάρτηση για το σκοπό αυτό, κάθε φορά που ο χρήστης μεταβάλλει το μέγεθος του παραθύρου, η σκηνή παραμορφώνεται ώστε να καταλάβει όλο το διαθέσιμο εύρος του παραθύρου, καθώς πραγματοποιείται ο προεπιλεγμένος μετασχηματισμός παραθύρου παρατήρησης. Σύμφωνα με αυτόν, οι παράμετροι width και height των σχέσεων (4.56) και (4.57) τίθενται ίσοι με τις νέες διαστάσεις του παραθύρου, ενώ θεωρείται ότι $x_1=y_1=0$. Αυτό σημαίνει ότι οι διαστάσεις της σκηνής είναι ίσες με τις διαστάσεις του παραθύρου και ότι το κάτω αριστερά pixel της σκηνής σχεδιάζεται στην κάτω αριστερή γωνία του παραθύρου. Αντίθετα, είναι επιθυμητή η εύρεση των διαστάσεων της σκηνής ώστε αυτή να μην καταλαμβάνει όλο το διαθέσιμο παράθυρο και να μην παραμορφώνεται. Για το λόγο αυτό πραγματοποιείται η διαδικασία που περιγράφεται στη συνέχεια. Αρχικά υπολογίζεται ο λόγος του νέου πλάτους του παραθύρου προς το αρχικό πλάτος του στιγμιότυπου (έστω AW) και ακολούθως ο λόγος του νέου ύψους του παραθύρου προς το αρχικό ύψος του στιγμιότυπου (έστω AH). Οι διαστάσεις του στιγμιότυπου υπολογίζονται με πολλαπλασιασμό των αρχικών διαστάσεών του με το μικρότερο εκ των δύο παραπάνω λόγων. Έτσι, αν μικρότερος είναι ο λόγος AW, οι νέες διαστάσεις του στιγμιότυπου γίνονται AW·width και AW·height, ενώ στην αντίθετη περίπτωση οι νέες διαστάσεις του είναι AH·width και AH·height, όπου ως width και height συμβολίζονται οι αρχικές διαστάσεις του στιγμιότυπου σε pixels (πλάτος και ύψος αντίστοιχα).

Επόμενο βήμα αποτελεί η είσοδος της εικόνας - πρότυπο στο πρόγραμμα, η οποία είναι μία φωτογραφία της τυπωμένης ορθοφωτογραφίας, ενώ στη συνέχεια υπολογίζονται οι

διαστάσεις της, δηλαδή το πλάτος και το ύψος της σε pixels. Κατόπιν, ορίζονται οι συντεταγμένες του πρότυπου επίπεδου αντικειμένου στο σύστημα συντεταγμένων της εικόνας - πρότυπο, δηλαδή οι συντεταγμένες pixel των τεσσάρων κορυφών αυτής, με αρχή το επάνω αριστερά εικονοστοιχείο, όπως φαίνεται στην Εικόνα 5-6, αριστερά. Ακολούθως, ορίζεται το επίγειο σύστημα συντεταγμένων με αρχή το κέντρο του πρότυπου αντικειμένου, άξονες X και Y στο επίπεδο του αντικειμένου και άξονα Z κάθετο σε αυτό με φορά προς την κάμερα, όπως φαίνεται στην Εικόνα 5-6, δεξιά. Στο σύστημα αυτό ορίζονται οι συντεταγμένες των τεσσάρων κορυφών του επίπεδου αντικειμένου. Η τρίτη συντεταγμένη Z είναι μηδενική, καθώς το αντικείμενο είναι επίπεδο και το επίγειο σύστημα συντεταγμένων ορίστηκε πάνω στο επίπεδο του αντικειμένου. Οι οριζοντιογραφικές επίγειες συντεταγμένες των τεσσάρων κορυφών του πρότυπου αντικειμένου ορίζονται με βάση τη θεώρηση ότι κυμαίνονται στο εύρος τιμών [-1,1] και αναγράφονται στην Εικόνα 5-6, δεξιά. Στην ίδια εικόνα ως W και H συμβολίζονται το κανονικοποιημένο πλάτος και ύψος της εικόνας - πρότυπο, όπως ορίζονται στη σχέση (5.3), στην οποία width και height είναι οι διαστάσεις της εικόνας - πρότυπο.

$$W = \frac{\text{width}}{\max(\text{width},\text{height})} \quad \text{και} \quad H = \frac{\text{height}}{\max(\text{width},\text{height})} \tag{5.3}$$

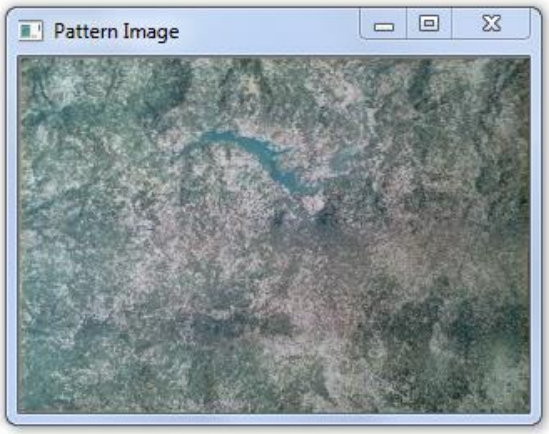
Έτσι, στην περίπτωση της συγκεκριμένης εικόνας - πρότυπο, της οποίας το πλάτος είναι μεγαλύτερο από το ύψος, δηλαδή ισχύει max(width,height)=width από τη σχέση (5.3) προκύπτει : W=1 και H<1.



Εικόνα 5-6. Σύστημα συντεταγμένων εικόνας [αριστερά], επίγειο σύστημα συντεταγμένων [δεξιά] και συντεταγμένες των τεσσάρων κορυφών του επίπεδου αντικειμένου στα δύο συστήματα.

Μετά τον ορισμό των συντεταγμένων των τεσσάρων αντίστοιχων σημείων της εικόνας και του χώρου, δημιουργείται ένα παράθυρο με χρήση της OpenCV με τίτλο «Pattern Image» και εμφανίζεται σε αυτό η εικόνα - πρότυπο, στην περίπτωση που ο χρήστης είχε επιλέξει να παρακολουθήσει τη βασική διαδικασία που ακολουθείται, μέσω της OpenCV, προκειμένου να επιτευχθεί η επαύξηση της πραγματικότητας. Το παράθυρο με την εικόνα - πρότυπο φαίνεται στην Εικόνα 5-7. Όταν ο χρήστης πατήσει οποιοδήποτε πλήκτρο, το παράθυρο

καταστρέφεται (κλείνει) και το πρόγραμμα συνεχίζει με το επόμενο βήμα, που είναι η ανίχνευση σημείων ενδιαφέροντος με χρήση του αλγορίθμου SURF.



Εικόνα 5-7. Παράθυρο «Pattern Image» στο οποίο εμφανίζεται η εικόνα - πρότυπο της Εφαρμογής 1.

Είτε εμφανίστηκε η εικόνα - πρότυπο είτε όχι, ανάλογα με την επιθυμία του χρήστη που δηλώθηκε στην αρχή της εφαρμογής, επόμενο βήμα αποτελεί η εξαγωγή των σημείων ενδιαφέροντος με τον ανιχνευτή του αλγορίθμου SURF, μέσω συνάρτησης της βιβλιοθήκης OpenCV. Όπως αναφέρθηκε στην υποενότητα 2.2.2, στην οποία περιγράφεται ο αλγόριθμος αυτός, ο SURF δεν χρησιμοποιεί πληροφορία χρώματος. Συνεπώς, πρώτο βήμα είναι η μετατροπή της εικόνας σε ασπρόμαυρη. Αυτό γίνεται με υλοποιημένη συνάρτηση από τη βιβλιοθήκη OpenCV, η οποία δέχεται την έγχρωμη εικόνα και τη μετατρέπει σε ασπρόμαυρη. Η αρχική εικόνα έχει τρία κανάλια (κόκκινο, πράσινο και μπλε). Η ψηφιακή τιμή του αντίστοιχου pixel της εικόνας προκύπτει ως ο σταθμισμένος μέσος όρος των ψηφιακών τιμών των τριών χρωμάτων, ως εξής:

$$D = 0.299 \cdot R + 0.587 \cdot G + 0.114 \cdot B \tag{5.4}$$

όπου με το χαρακτήρα D έχει συμβολιστεί η ψηφιακή τιμή του εικονοστοιχείου της ασπρόμαυρης εικόνας εξόδου, ενώ R, G και B είναι οι τιμές του κόκκινου, πράσινου και μπλε καναλιού αντίστοιχα της έγχρωμης εικόνας για το εικονοστοιχείο αυτό.

Μετά τη μετατροπή της εικόνας από έγχρωμη σε ασπρόμαυρη, πραγματοποιείται η ανίχνευση των σημείων ενδιαφέροντος και στη συνέχεια η περιγραφή αυτών, με χρήση του SURF (βλ. υποενότητα 2.2.2), ο οποίος έχει υλοποιηθεί στη βιβλιοθήκη OpenCV. Αρχικά καθορίζονται οι παράμετροι που δέχεται ο αλγόριθμος, μία εκ των οποίων είναι ένα κατώφλι που αντιστοιχεί στην ελάχιστη τιμή της ορίζουσας της παραλλαγής του Εσσιανού Πίνακα, την οποία πρέπει να διαθέτει ένα εικονοστοιχείο προκειμένου να μπορεί να χαρακτηριστεί ως σημείο ενδιαφέροντος. Η τιμή αυτή εξαρτάται από την εκάστοτε εφαρμογή και βρίσκεται μέσω δοκιμών, ενώ όσο μικρότερη είναι, τόσο περισσότερα σημεία ενδιαφέροντος μπορούν να εντοπιστούν. Η τιμή του κατωφλίου η οποία κρίθηκε ικανοποιητική για τη συγκεκριμένη εφαρμογή και επιλέχτηκε μέσω διαφορετικών δοκιμών είναι ο αριθμός 200. Επίσης,

καθορίζονται ο αριθμός των οκτάβων και των επιπέδων κλίμακας κάθε οκτάβας, το μέγεθος του διανύσματος περιγραφικών χαρακτηριστικών κάθε σημείου ενδιαφέροντος (64 ή 128) και ο υπολογισμός ή μη του κυρίαρχου προσανατολισμού του σημείου ενδιαφέροντος. Συγκεκριμένα, επιλέχτηκε ο χώρος κλίμακας να χωριστεί σε τέσσερις οκτάβες και η κάθε μία να περιλαμβάνει τέσσερα επίπεδα κλίμακας, επιλογές οι οποίες προτείνονται και από τους συγγραφείς του αλγορίθμου SURF [85]. Επιπλέον, καθορίζεται το διάνυσμα των περιγραφικών χαρακτηριστικών κάθε σημείου ενδιαφέροντος να έχει διάσταση 64, και δεν γίνεται χρήση της περισσότερο απαιτητικής σε χρόνο παραλλαγής SURF-128 του αλγορίθμου SURF (βλ. παράγραφο 2.2.2.7), καθώς η κανονική εκδοχή του SURF που χρησιμοποιεί περιγραφέα 64 διαστάσεων δίνει εξίσου ικανοποιητικά αποτελέσματα. Τέλος, μολονότι δίνεται η δυνατότητα χρήσης της παραλλαγής U-SURF του αλγορίθμου SURF (βλ. παράγραφο 2.2.2.6), η οποία παραλείπει το στάδιο ανάθεσης προσανατολισμού, καθορίζεται η διεξαγωγή του υπολογισμού του κυρίαρχου προσανατολισμού κάθε σημείου ενδιαφέροντος, ώστε κατά την κυρίως διαδικασία της συνταύτισης να μπορούν να ταυτιστούν τα ομόλογα σημεία, ανεξάρτητα αν η εικόνα - πρότυπο και το εκάστοτε στιγμιότυπο διαφέρουν κατά μία στροφή. Ενώ το κατώφλι αλλάζει ανάλογα με την εκάστοτε περιοχή ενδιαφέροντος, οι υπόλοιπες παράμετροι για τη ανίχνευση και περιγραφή των σημείων ενδιαφέροντος κάθε φορά που γίνεται χρήση του αλγορίθμου SURF, τόσο στην Εφαρμογή 1 όσο και στην Εφαρμογή 2, είναι οι ίδιες.

Μετά τον καθορισμό των παραπάνω παραμέτρων πραγματοποιείται η ανίχνευση των σημείων ενδιαφέροντος στην ασπρόμαυρη εικόνα - πρότυπο. Αποτέλεσμα της διαδικασίας αυτής είναι η δομή δεδομένων «KeyPoint», η οποία περιλαμβάνει τις ιδιότητες για κάθε σημείο ενδιαφέροντος το οποίο εντοπίστηκε μέσω του ανιχνευτή του αλγορίθμου SURF. Οι ιδιότητες αυτές είναι οι εξής [194], [195]:

- η θέση του σημείου στην εικόνα σε pixels
- η διάμετρος της περιοχής ενδιαφέροντος του σημείου αυτού, η οποία εξαρτάται από την κλίμακα στην οποία εντοπίστηκε
- ο κυρίαρχος προσανατολισμός του σημείου ενδιαφέροντος, ο οποίος μετράται σε μοίρες δεξιόστροφα σε σχέση με το σύστημα συντεταγμένων εικόνας της OpenCV
- η απόκριση του συγκεκριμένου σημείου, που αντιστοιχεί στην ορίζουσα της παραλλαγής του Εσσιανού Πίνακα
- ο αριθμός της οκτάβας στην οποία εντοπίστηκε
- ένας αριθμός (1 ή -1) ανάλογα με το ίχνος του Εσσιανού Πίνακα που δείχνει αν πρόκειται για ανοικτή blob σε σκούρο περιβάλλον ή σκούρα blob σε ανοικτό περιβάλλον.

Η δομή αυτή μπορεί να γίνει κατανοητή ως ένα διάνυσμα σε έναν πολυδιάστατο χώρο. Το μέγεθος του διανύσματος είναι ίσο με τον αριθμό των σημείων ενδιαφέροντος που

ανιχνεύτηκαν, ενώ κάθε στοιχείο του διανύσματος είναι οι προαναφερθείσες ιδιότητες κάθε σημείου.

Στην περίπτωση που ο χρήστης είχε επιλέξει να παρακολουθήσει τη βασική διαδικασία που διεξάγεται μέσω της βιβλιοθήκης OpenCV, δημιουργείται ένα παράθυρο, με τίτλο «SURF Features» στο οποίο εμφανίζεται η έγχρωμη εικόνα - πρότυπο με τα σημεία ενδιαφέροντος τα οποία εντοπίστηκαν. Η OpenCV έχει τη δυνατότητα σχεδίασης των σημείων ενδιαφέροντος με τέτοιο τρόπο ώστε να φαίνεται ο κυρίαρχος προσανατολισμός κάθε σημείου και το μέγεθος της περιοχής ενδιαφέροντός του, ανάλογα με την κλίμακα στην οποία αυτό εντοπίστηκε. Ωστόσο, λόγω του πλήθους των σημείων ενδιαφέροντος και του μικρού μεγέθους της εικόνας, ο παραπάνω τρόπος απεικόνισης των σημείων δεν δημιουργεί ένα ευκρινές αποτέλεσμα, παρόλο που επιλέχτηκε η χρήση τυχαίων διαφορετικών χρωμάτων για κάθε σημείο (βλ. Εικόνα 5-8, αριστερά). Επομένως, στην εφαρμογή επιλέγεται η απεικόνιση των σημείων ενδιαφέροντος στην εικόνα - πρότυπο με τον τρόπο που φαίνεται στην Εικόνα 5-8, δεξιά, στην οποία έχει σχεδιαστεί ένας κύκλος με κέντρο το κάθε σημείο ενδιαφέροντος, ίδιου χρώματος και μεγέθους για όλα τα σημεία. Το παράθυρο αυτό παραμένει ανοικτό όταν ο χρήστης δεν πατά κάποιο πλήκτρο, ενώ κλείνει και συνεχίζει η διαδικασία επαύξησης της πραγματικότητας όταν αυτός πατήσει οποιοδήποτε πλήκτρο.

Εικόνα 5-8. Απεικόνιση των σημείων ενδιαφέροντος στην εικόνα - πρότυπο της Εφαρμογής 1 με ταυτόχρονη εμφάνιση οπτικών πληροφοριών για την κλίμακα και τον προσανατολισμό κάθε σημείου [αριστερά] και απεικόνιση των σημείων ενδιαφέροντος μέσω απλών κύκλων χωρίς διαφοροποίηση του μεγέθους αυτών [δεξιά], η οποία τελικά χρησιμοποιήθηκε.

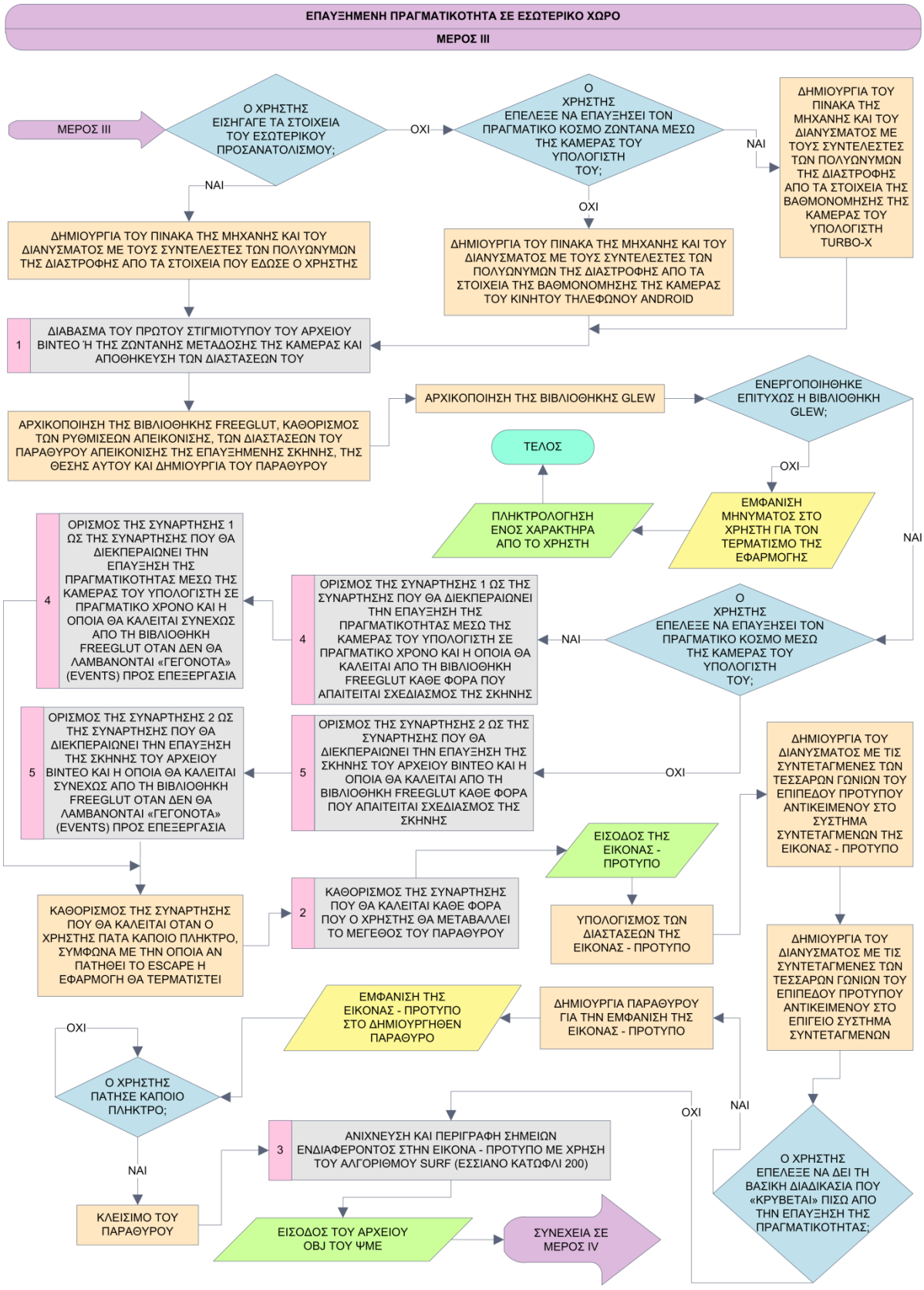
Επόμενο βήμα αποτελεί η εξαγωγή των διανυσμάτων περιγραφικών χαρακτηριστικών των σημείων ενδιαφέροντος, η οποία πραγματοποιείται με δεδομένη τη δομή δεδομένων «KeyPoint». Αποτέλεσμα της διαδικασίας αυτής είναι ένας πίνακας n γραμμών και 64 στηλών, όπου n ο αριθμός των σημείων ενδιαφέροντος που εντοπίστηκαν και 64 το επιλεγμένο μέγεθος του διανύσματος περιγραφικών χαρακτηριστικών κάθε σημείου ενδιαφέροντος. Σε κάθε γραμμή του πίνακα αυτού αποθηκεύεται το διάνυσμα περιγραφικών χαρακτηριστικών του αντίστοιχου σημείου της δομής «Keypoint». Ωστόσο, αν για κάποιο σημείο δεν είναι δυνατός ο υπολογισμός του διανύσματος περιγραφικών χαρακτηριστικών,

το σημείο αυτό, το οποίο αναπαρίσταται από τις ιδιότητές του στη δομή «KeyPoint», αφαιρείται από αυτή [196].

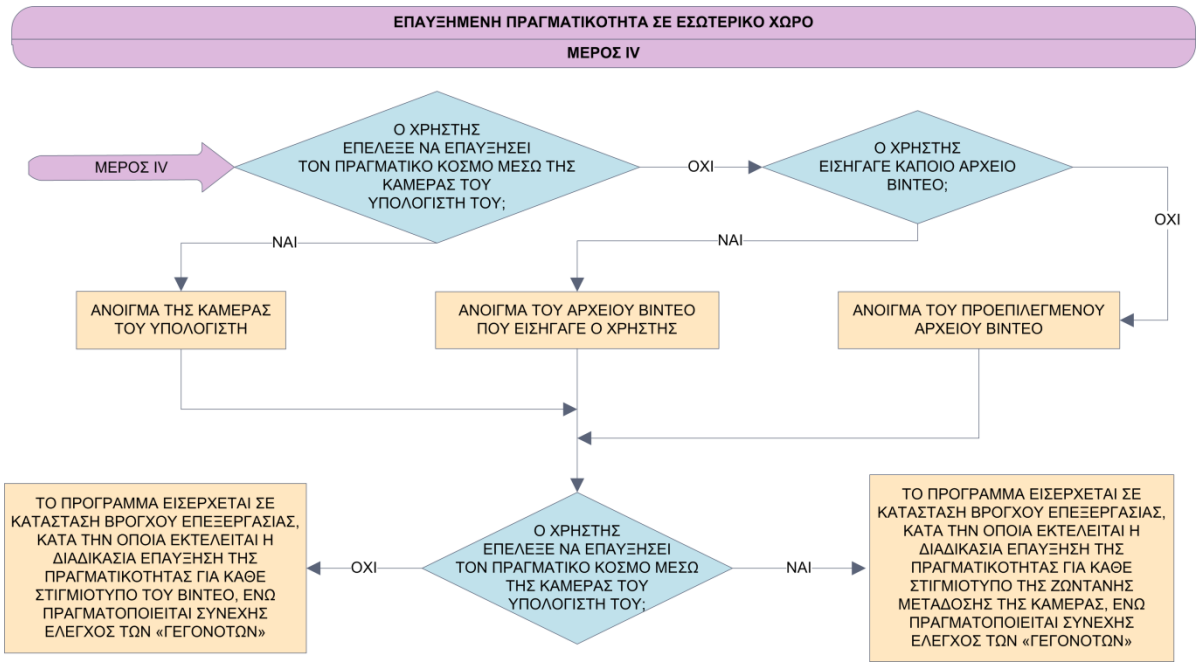
Στη συνέχεια, εισάγεται στο πρόγραμμα το ΨΜΕ (βλ. Εικόνα 5-5, αριστερά), με χρήση της βιβλιοθήκης GLM. Το ΨΜΕ διατέθηκε σε μορφότυπο OBJ υπό τη συνοδεία ενός αρχείου MTL και μίας εικόνας. Ωστόσο, στο πλαίσιο της εφαρμογής δεν χρησιμοποιήθηκε το εξωτερικό αρχείο υλικού MTL (σβήνοντας τις αντίστοιχες εντολές από το αρχείο OBJ), με αποτέλεσμα να διαβάζεται μόνο το περιεχόμενο του αρχείου OBJ. Έτσι, στο βήμα σχεδιασμού του ΨΜΕ στη σκηνή, η εικόνα που συνοδεύει το αρχείο OBJ εισάγεται στο πρόγραμμα και μέσω της διαδικασίας απόδοσης υφής της OpenGL το τρισδιάστατο μοντέλο σχεδιάζεται ορθά με την εικόνα που το συνοδεύει (βλ. παράγραφο 5.4.3.4).

Ακολούθως, αν ο χρήστης είχε επιλέξει να επαυξήσει τον πραγματικό κόσμο μέσω της κάμερας του υπολογιστή του, ανοίγει η κάμερα του υπολογιστή. Διαφορετικά, σε περίπτωση που εισήγαγε κάποιο αρχείο βίντεο, ανοίγει το τελευταίο, ή το προεπιλεγμένο αρχείο βίντεο από την εφαρμογή εφόσον ο χρήστης είχε επιλέξει να παρακολουθήσει μία προεπισκόπηση αυτής. Στη συνέχεια, εκτελείται η διαδικασία επαύξησης της πραγματικότητας για κάθε στιγμιότυπο του βίντεο, μέχρι αυτό να σταματήσει ή μέχρι ο χρήστης να πατήσει το πλήκτρο «Escape», ή για κάθε στιγμιότυπο της ζωντανής μετάδοσης της κάμερας του υπολογιστή, μέχρι ο χρήστης να πατήσει το πλήκτρο «Escape». Αυτό γίνεται μέσω της ενεργοποίησης του βρόγχου επεξεργασίας ο οποίος ταυτόχρονα ελέγχει όλα τα «γεγονότα». Αυτό σημαίνει ότι κάθε φορά που απαιτείται σχεδιασμός της σκηνής αλλά και συνεχώς, όταν δεν λαμβάνονται «γεγονότα» προς επεξεργασία, εκτελείται η διαδικασία η οποία θα πραγματοποιεί την επαύξηση της πραγματικότητας κάθε στιγμιότυπου, μέσω κλήσης της κατάλληλης συνάρτησης. Παράλληλα, εξετάζονται και οι αλληλεπιδράσεις του χρήστη με τις μονάδες εισόδου και όταν απαιτείται πραγματοποιούνται οι κατάλληλες ενέργειες, όπως έχουν οριστεί. Συγκεκριμένα, αν πατηθεί το πλήκτρο «Escape» η εφαρμογή σταματά, ενώ αν ο χρήστης μεταβάλλει το μέγεθος του παραθύρου στο οποίο απεικονίζεται η επαυξημένη σκηνή, το μέγεθος της τελευταίας μεταβάλλεται κατάλληλα, ώστε να διατηρείται σταθερός ο λόγος των μηκών των πλευρών της. Ενώ οι δύο τελευταίες συναρτήσεις περιγράφηκαν στην παράγραφο αυτή, στην παράγραφο 5.4.3 θα παρουσιαστεί αναλυτικά η επαναληπτική διαδικασία επαύξησης της πραγματικότητας.

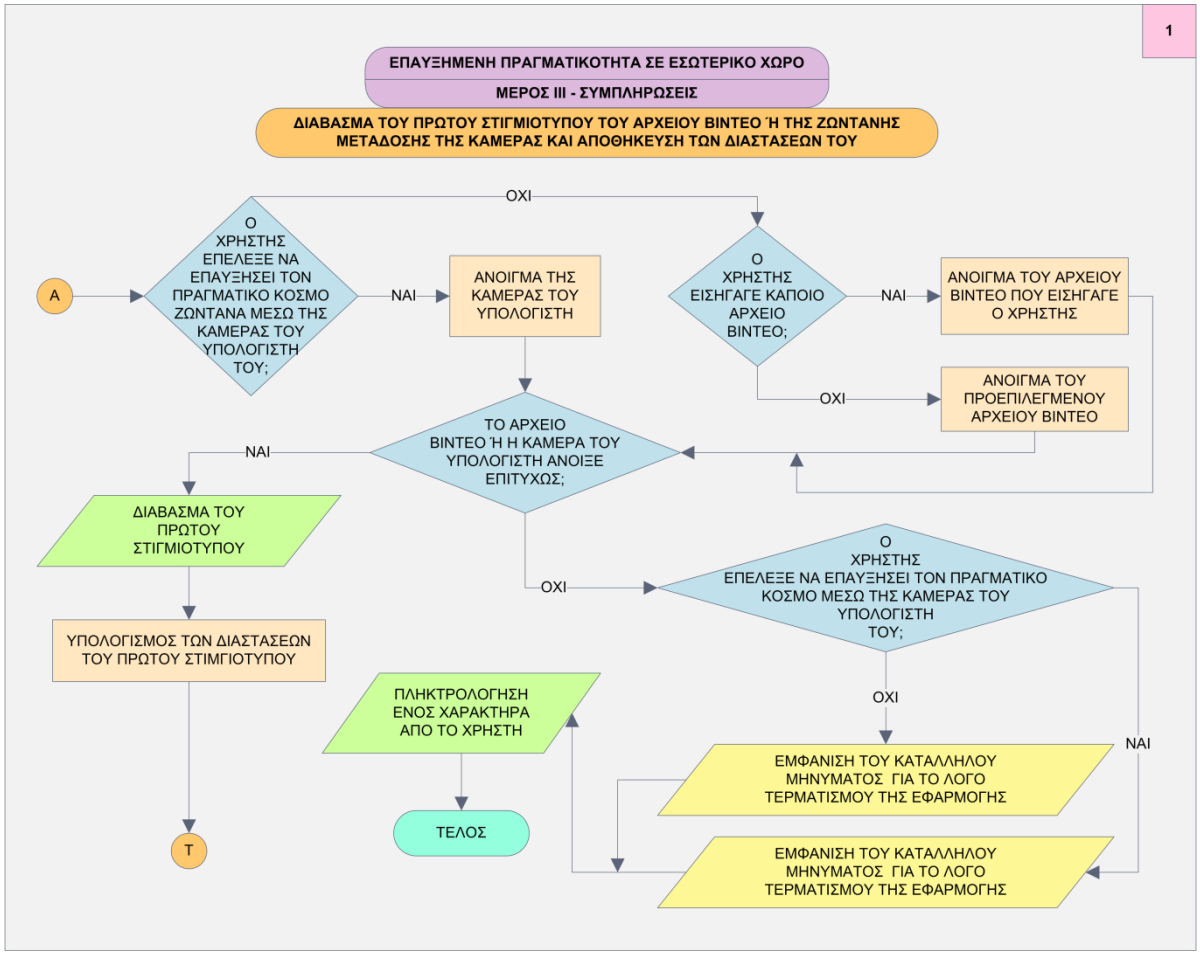
Στις Διαγραμματικές παρουσιάσεις 5-4 και 5-5 φαίνεται το διάγραμμα ροής της διαδικασίας που περιγράφηκε στην παρούσα υποενότητα, το οποίο αποτελεί συνέχεια της Διαγραμματικής παρουσίασης 5-3. Στις Διαγραμματικές παρουσιάσεις 5-6, 5-7 και 5-8 απεικονίζονται τα πλαίσια συμπληρώσεων της Διαγραμματικής Παρουσίασης 5-3.



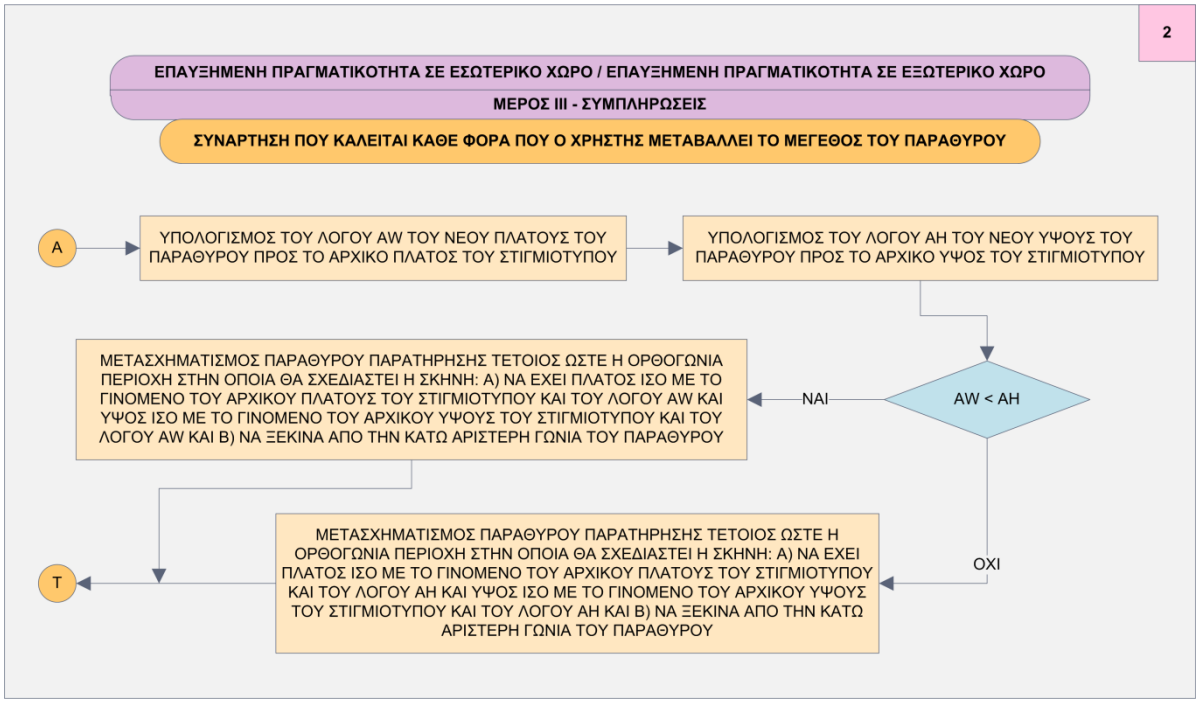
Διαγραμματική παρουσίαση 5-4. Εφαρμογή 1, μέρος III.



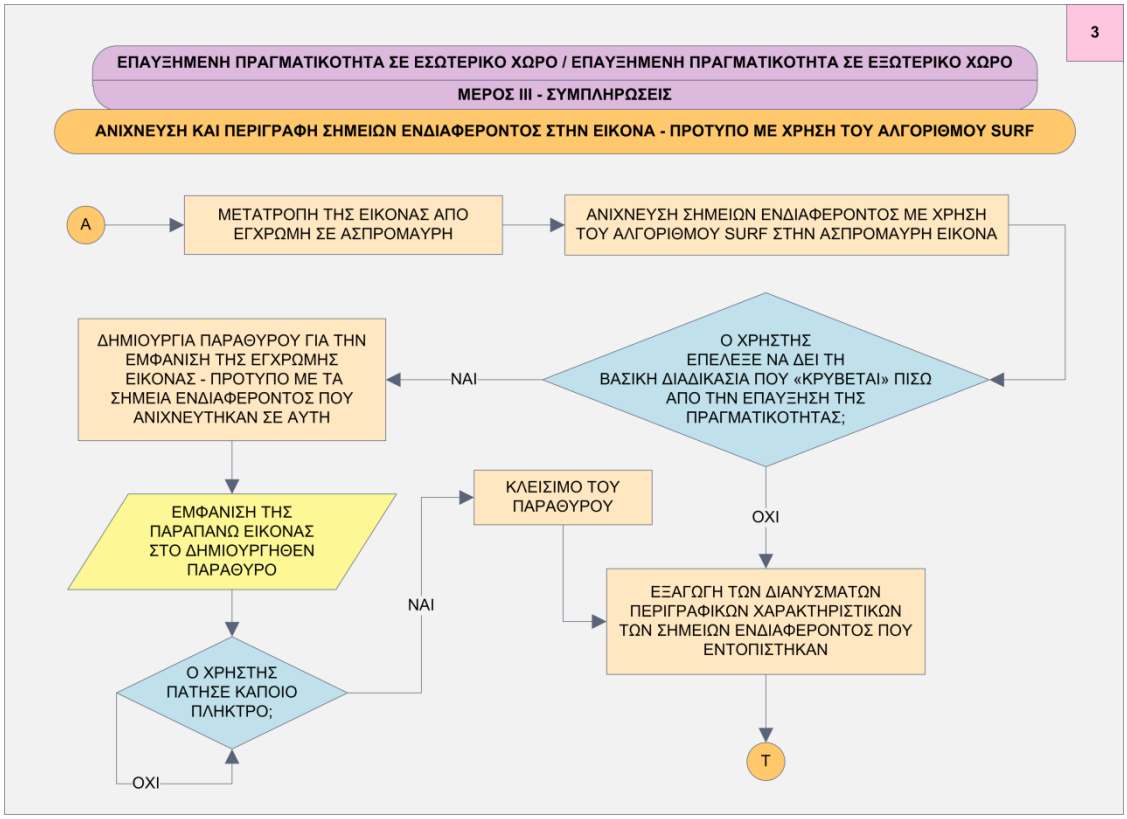
Διαγραμματική παρουσίαση 5-5. Εφαρμογή 1, μέρος IV.



Διαγραμματική παρουσίαση 5-6. Εφαρμογή 1, πλαίσιο συμπληρώσεων 1.



Διαγραμματική παρουσίαση 5-7. Εφαρμογή 1 - Εφαρμογή 2, πλαίσιο συμπληρώσεων 2.



Διαγραμματική παρουσίαση 5-8. Εφαρμογή 1 - Εφαρμογή 2, πλαίσιο συμπληρώσεων 3.

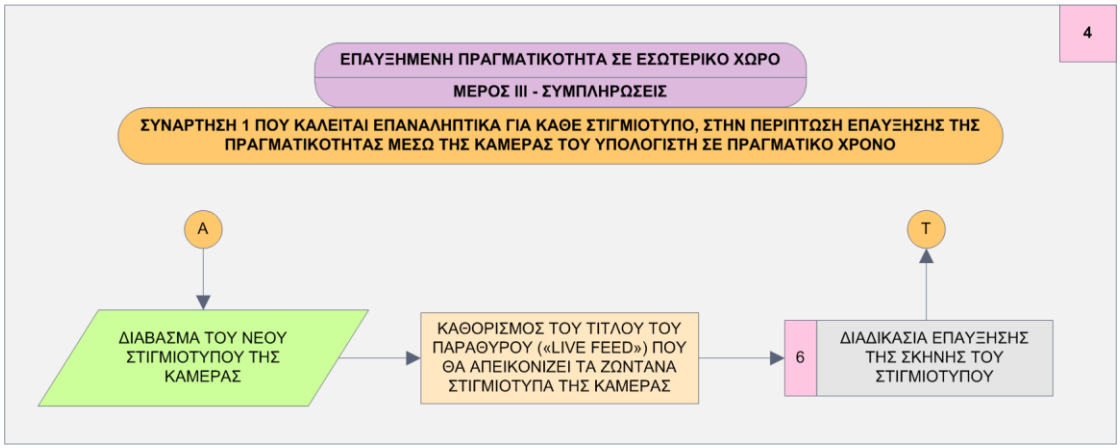
5.4.3 Επαναληπτική διαδικασία επαύξησης της σκηνής κάθε στιγμιότυπου

Στην παρούσα υποενότητα παρουσιάζεται η επαναληπτική διαδικασία η οποία εκτελείται για κάθε στιγμιότυπο του βίντεο ή της ζωντανής μετάδοσης της κάμερας. Πρόκειται για τη συνάρτηση η οποία θα καλείται συνεχώς μετά από την είσοδο του προγράμματος στο βρόγχο επεξεργασίας, ο οποίος ταυτόχρονα ελέγχει όλα τα «γεγονότα». Οποιαδήποτε εντολή γραφεί μετά την ενεργοποίηση του βρόγχου επεξεργασίας δεν εκτελείται. Για το λόγο αυτό, το τέλος της εφαρμογής πρέπει να προγραμματιστεί να γίνει είτε μέσω της αλληλεπίδρασης του χρήστη με τις συσκευές εισόδου, είτε μέσω της συνάρτησης που καλείται συνεχώς μετά την είσοδο του προγράμματος στο βρόγχο επεξεργασίας. Για το λόγο αυτό, στο διάγραμμα ροής της Διαγραμματικής παρουσίασης 5-5 μετά την ενεργοποίηση της ατέρμονα βρόγχου δεν τοποθετείται η χαρακτηριστική λέξη «ΤΕΛΟΣ».

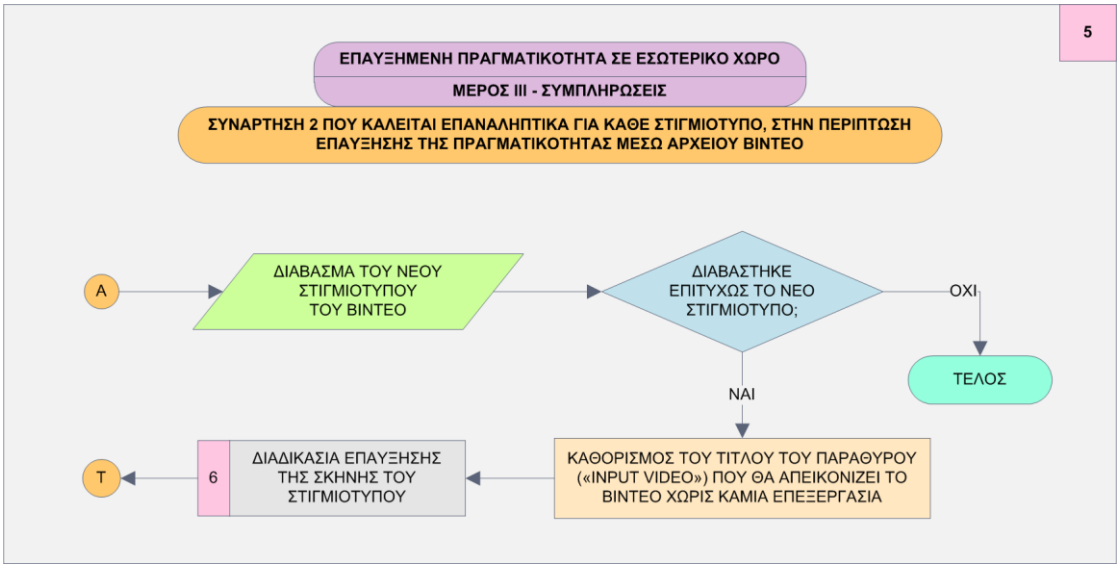
Όπως αναφέρθηκε στην υποενότητα 5.4.2, ανάλογα με το αν ο χρήστης επέλεξε επαύξηση των σκηνών του βίντεο ή των ζωντανών στιγμιότυπων της κάμερας, ορίζεται διαφορετική συνάρτηση επαύξησης της πραγματικότητας.

- Στην περίπτωση επαύξησης της πραγματικότητας σε πραγματικό χρόνο μέσω λήψης στιγμιότυπων από την κάμερα του υπολογιστή (βλ. Διαγραμματική παρουσίαση 5-9), αρχικά διαβάζεται το νέο στιγμιότυπο. Στη συνέχεια, ορίζεται ο τίτλος του παραθύρου στο οποίο θα παρουσιάζονται τα στιγμιότυπα της κάμερας («Live Feed») – στην περίπτωση που ο χρήστης επέλεξε να παρακολουθήσει τη βασική διαδικασία που εκτελείται μέσω της OpenCV – και ακολουθεί η επεξεργασία του στιγμιότυπου για την επαύξηση της σκηνής.
- Στην περίπτωση επαύξησης των σκηνών ενός αρχείου βίντεο (βλ. Διαγραμματική παρουσίαση 5-10), γίνεται προσπάθεια να διαβαστεί το εκάστοτε στιγμιότυπο του βίντεο. Αν αυτό δεν διαβάστηκε, γεγονός το οποίο οφείλεται πιθανότατα στο τέλος του αρχείου βίντεο, δηλαδή στη μη ύπαρξη άλλων στιγμιότυπων, η εφαρμογή τερματίζεται. Διαφορετικά, καθορίζεται ο τίτλος του παραθύρου που θα απεικονίζει το βίντεο («Input Video») – εφόσον ο χρήστης επέλεξε να παρακολουθήσει τη βασική διαδικασία που εκτελείται μέσω της βιβλιοθήκης OpenCV – και ακολουθεί η διαδικασία επαύξησης της πραγματικότητας για το συγκεκριμένο στιγμιότυπο.

Η διαδικασία επαύξησης της πραγματικότητας για κάθε στιγμιότυπο, που είναι κοινή και για τις δύο περιπτώσεις, αναλύεται στις παραγράφους που ακολουθούν.



Διαγραμματική παρουσίαση 5-9. Εφαρμογή 1, πλαίσιο συμπληρώσεων 4.



Διαγραμματική παρουσίαση 5-10. Εφαρμογή 1, πλαίσιο συμπληρώσεων 5.

5.4.3.1 Υπολογισμός του εξωτερικού προσανατολισμού του στιγμιότυπου

Αρχικό βήμα της διαδικασίας επαύξησης της σκηνής του στιγμιότυπου είναι η μετατροπή του σε ασπρόμαυρο, προκειμένου να πραγματοποιηθεί η ανίχνευση και περιγραφή των σημείων ενδιαφέροντος από τον αλγόριθμο SURF. Το κατώφλι που ορίζεται είναι το ίδιο με αυτό που ορίστηκε για την ανίχνευση των σημείων ενδιαφέροντος στην εικόνα - πρότυπο (200).

Επόμενο βήμα αποτελεί η ψηφιακή συνταύτιση της εικόνας - πρότυπο και του στιγμιότυπου, με δεδομένους τους πίνακες με τα διανύσματα περιγραφικών χαρακτηριστικών των σημείων ενδιαφέροντος της εικόνας - πρότυπο και του στιγμιότυπου. Στην εφαρμογή αυτή επιλέγεται να διεξαχθεί η συνταύτιση με χρήση του διασταυρωμένου ελέγχου (cross-check test) ως μεθόδου απαλοιφής των μη έγκυρων ομολογιών (βλ. παράγραφο 2.3.3.2) ενώ κριτήριο για την ταύτιση των διανυσμάτων περιγραφικών χαρακτηριστικών θεωρείται η ευκλείδεια απόσταση (βλ. υποενότητα 2.3.1). Με βάση τη διαδικασία αυτή, τα εικονοσημεία i και j

θεωρούνται ομόλογα εφόσον για το εικονοσημείο i της εικόνας - πρότυπο το εικονοσημείο j του στιγμιότυπου είναι το κοντινότερο σε αυτό, με βάση την ευκλείδεια απόσταση των διανυσμάτων περιγραφικών χαρακτηριστικών τους, και επιπρόσθετα εφόσον το κοντινότερο του εικονοσημείου j του στιγμιότυπου είναι το εικονοσημείο i της εικόνας - πρότυπο, με βάση το ίδιο κριτήριο. Η διαδικασία αυτή διεξάγεται μέσω υλοποιημένης συνάρτησης της βιβλιοθήκης OpenCV. Αποτέλεσμα είναι μία δομή «DMatch» που περιλαμβάνει για κάθε εντοπισμένο ζεύγος ομολογιών:

- ένα δείκτη που φανερώνει τη θέση του σημείου j στη δομή «Keypoint» του στιγμιότυπου
- ένα δείκτη για τη θέση του σημείου i στη δομή «Keypoint» της εικόνας - πρότυπο
- ένα δείκτη που χαρακτηρίζει την εικόνα πρότυπο (ο οποίος δεν είναι χρήσιμος για την εφαρμογή)
- την ευκλείδεια απόσταση των διανυσμάτων περιγραφικών χαρακτηριστικών των σημείων i και j.

Όπως και η δομή «Keypoint», η δομή «DMatch» που προκύπτει από τη συνταύτιση μπορεί να γίνει κατανοητή ως ένα διάνυσμα σε έναν πολυδιάστατο χώρο, κάθε στοιχείο του οποίου είναι οι τέσσερις προαναφερθείσες ιδιότητες του ζεύγους ομόλογων σημείων.

Οι ομολογίες που προκύπτουν από την παραπάνω διαδικασία περιέχουν αρκετά χονδροειδή λάθη, τα οποία οφείλονται στο γεγονός ότι δεν έχει οριστεί κάποιο κατώφλι για την αποδοχή ή την απόρριψή τους και ο μόνος έλεγχος ο οποίος γίνεται είναι το cross-check test. Για το λόγο αυτό, θεωρείται ότι το ζεύγος σημείων είναι ομόλογο εφόσον η ευκλείδεια απόσταση των διανυσμάτων περιγραφικών χαρακτηριστικών τους είναι μικρότερη της τιμής 0.22, η οποία βρέθηκε μετά από δοκιμές. Με βάση αυτή τη θεώρηση απορρίπτονται αρκετές μη έγκυρες ομολογίες.

Επόμενο βήμα αποτελεί ο υπολογισμός του δισδιάστατου προβολικού μετασχηματισμού μεταξύ της εικόνας - πρότυπο και του στιγμιότυπου, με δεδομένες τις ευρεθείσες ομολογίες σημείων. Η εύρεση μίας αρχικής εκτίμησης του πίνακα του μετασχηματισμού αυτού γίνεται μέσω του αλγορίθμου RANSAC (βλ. παράγραφο 2.3.3.3), ο οποίος ταυτόχρονα απορρίπτει τις μη έγκυρες ομολογίες που είναι πιθανό να εξακολουθούν να υπάρχουν, ενώ η λύση αυτή βελτιώνεται μέσω ενός αλγορίθμου βελτιστοποίησης, με χρήση όλων των έγκυρων ομολογιών που εντοπίστηκαν από το RANSAC. Η διαδικασία αυτή πραγματοποιείται στην περίπτωση που οι διατιθέμενες ομολογίες σημείων είναι τουλάχιστον πέντε. Παρόλο που απαιτούνται κατ' ελάχιστο τέσσερις ομολογίες για τον υπολογισμό της ομογραφίας, στο πλαίσιο της εφαρμογής θεωρείται ότι αν οι ομολογίες είναι λιγότερες από πέντε το πρότυπο αντικείμενο δεν έχει εντοπιστεί στην εικόνα, οπότε δεν έχει νόημα ο υπολογισμός του πίνακα του δισδιάστατου προβολικού μετασχηματισμού. Εξάλλου, αν οι ομολογίες ήταν τέσσερις

(γεγονός το οποίο σημαίνει ότι η χρήση του αλγορίθμου RANSAC δεν θα είχε νόημα) και μία ή περισσότερες εξ' αυτών ήταν λανθασμένη/ες, θα ήταν λανθασμένος και ο πίνακας της ομογραφίας, με αποτέλεσμα τη μη σωστή τοποθέτηση του τρισδιάστατου μοντέλου στο στιγμιότυπο. Επομένως, εφόσον οι ομολογίες είναι τουλάχιστον πέντε, υπολογίζεται ο 3x3 πίνακας της ομογραφίας, ο οποίος συνδέει τις συντεταγμένες pixel της εικόνας - πρότυπο με τις συντεταγμένες pixel του πρότυπου επίπεδου αντικειμένου που εντοπίστηκε στο στιγμιότυπο. Η σχέση αυτή εφαρμόζεται σε ομογενείς συντεταγμένες και είναι η ακόλουθη:

$$s \cdot \begin{bmatrix} x_{frame} \\ y_{frame} \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \cdot \begin{bmatrix} x_{pattern} \\ y_{pattern} \\ 1 \end{bmatrix} \tag{5.5}$$

όπου h_{ij} τα στοιχεία του πίνακα **H** της ομογραφίας, $x_{pattern}$, $y_{pattern}$ οι συντεταγμένες pixel ενός σημείου της εικόνας - πρότυπο, x_{frame}, y_{frame} οι συντεταγμένες pixel του ίδιου σημείου στο στιγμιότυπο και s οποιαδήποτε μη μηδενική σταθερά. Ο μετασχηματισμός αυτός δεν μεταβάλλεται με τον πολλαπλασιασμό του πίνακα **H** με μη μηδενικό συντελεστή και έτσι ο **H** «κανονικοποιείται», ώστε:

$$h_{33} = 1 \tag{5.6}$$

Η διαδικασία εύρεσης του προβολικού μετασχηματισμού από το σύνολο των έγκυρων ομολογιών γίνεται μέσω συνάρτησης της OpenCV και μπορεί να συνοψισθεί ως εξής:

- Επιλογή ενός τυχαίου δείγματος τεσσάρων ομολογιών από το σύνολο των ομολογιών.
- Εκτίμηση του πίνακα του δισδιάστατου προβολικού μετασχηματισμού με βάση το παραπάνω δείγμα.
- Υπολογισμός της «ποιότητας» της υπολογισμένης ομογραφίας, η οποία δίνεται από τον αριθμό των έγκυρων ομολογιών (inliers) για τη συγκεκριμένη λύση. Οι έγκυρες ομολογίες εντοπίζονται μέσω του υπολογισμού της ευκλείδειας απόστασης μεταξύ της υπολογισμένης θέσης του σημείου στο στιγμιότυπο του βίντεο, με βάση τον προβολικό μετασχηματισμό, και της πραγματικής θέσης του, όπως προέκυψε από τη συνταύτιση. Αν η απόσταση αυτή είναι μικρότερη από ένα κατώφλι, το οποίο στην περίπτωση της εφαρμογής ορίστηκε ίσο με 5 pixels, η ομολογία θεωρείται έγκυρη για τη συγκεκριμένη λύση του προβολικού μετασχηματισμού (inlier), αλλιώς θεωρείται άκυρη (outlier), όπως γίνεται φανερό από την ακόλουθη σχέση:

$$\begin{aligned} &\text{Αν} \quad \left\| point_{frame} - computed_point_{frame} \right\| \le 5 \quad \rightarrow \quad \text{inlier} \\ &\text{όπου} \quad point_{frame} = \begin{bmatrix} x_{frame} & y_{frame} \end{bmatrix}^T \quad \text{και} \\ &computed_point_{frame} = \left[\frac{h_{11} \cdot x_{pattern} + h_{12} \cdot y_{pattern} + h_{13}}{h_{31} \cdot x_{pattern} + h_{32} \cdot y_{pattern} + 1} \quad \frac{h_{21} \cdot x_{pattern} + h_{22} \cdot y_{pattern} + h_{23}}{h_{31} \cdot x_{pattern} + h_{32} \cdot y_{pattern} + 1} \right]^T \end{aligned} \tag{5.7}$$

- Επανάληψη των παραπάνω βημάτων μέχρι να εντοπιστεί η λύση με την επιθυμητή ποιότητα. Τα βήματα αυτά επαναλαμβάνονται το πολύ N=2000 φορές, όπου N ο μέγιστος αριθμός των επαναλήψεων που καθορίζεται στη συνάρτηση της OpenCV.
- Επιλογή της λύσης που υπολογίστηκε από τον αλγόριθμο RANSAC και οδηγεί στο μεγαλύτερο αριθμό των έγκυρων ομολογιών ως μίας αρχικής εκτίμησης του πίνακα της ομογραφίας. Εκτός από τους συντελεστές του προβολικού μετασχηματισμού υπολογίζεται και ένα διάνυσμα που περιλαμβάνει πληροφορίες για τις θέσεις των έγκυρων ομολογιών στη δομή «DMatch». Συγκεκριμένα, το διάνυσμα αυτό έχει διάσταση ίδια με τον αριθμό των ομολογιών – έγκυρων και μη – της δομής «DMatch» και στην περίπτωση που πρόκειται για μη έγκυρη αντιστοίχιση (outlier), το αντίστοιχο στοιχείο του διανύσματος είναι κενό, ενώ αν πρόκειται για έγκυρη ομολογία (inlier), το αντίστοιχο στοιχείο λαμβάνει μία καθορισμένη τιμή.
- Βελτίωση της αρχικής εκτίμησης του πίνακα της ομογραφίας, που υπολογίστηκε από το καλύτερο δείγμα των τεσσάρων ομολογιών, μέσω του αλγορίθμου Levenberg-Marquardt, με χρήση όλων των έγκυρων ομολογιών που εντοπίστηκαν από τον αλγόριθμο RANSAC, για την περαιτέρω ελαχιστοποίηση του σφάλματος επαναπροβολής.

Εφόσον ο πίνακας του προβολικού μετασχηματισμού υπολογίστηκε με χρήση τουλάχιστον πέντε ομολογιών, δηλαδή αν τα inliers που προέκυψαν από τον αλγόριθμο RANSAC είναι τουλάχιστον πέντε, θεωρείται ότι το πρότυπο έχει εντοπιστεί στο στιγμιότυπο και μπορεί να πραγματοποιηθεί η διαδικασία επαύξησής του. Η επιλογή του αριθμού πέντε ως ορίου για την επαύξηση ή μη της σκηνής, υπαγορεύεται από το γεγονός ότι εφόσον οι έγκυρες ομολογίες με βάση τον αλγόριθμο RANSAC είναι τέσσερις, αυτές συνιστούν το τυχαίο δείγμα το οποίο χρησιμοποιήθηκε για τον υπολογισμό του πίνακα και επομένως υπάρχει πιθανότητα να μην είναι σωστό.

Αν οι έγκυρες ομολογίες είναι τουλάχιστον πέντε, υπολογίζονται οι συντεταγμένες των τεσσάρων κορυφών της εικόνας - πρότυπο στο στιγμιότυπο με βάση τον πίνακα της ομογραφίας, ο οποίος μετασχηματίζει τις συντεταγμένες της εικόνας - πρότυπο σε συντεταγμένες στιγμιότυπου. Η παραπάνω διαδικασία γίνεται με χρήση συνάρτησης της βιβλιοθήκης OpenCV, η οποία μετασχηματίζει τις συντεταγμένες των τεσσάρων κορυφών της εικόνας - πρότυπο σε ομογενείς συντεταγμένες (θέτοντας την τρίτη συντεταγμένη w ίση με τη μονάδα), τις πολλαπλασιάζει με τον 3x3 πίνακα του προβολικού μετασχηματισμού και υπολογίζει τις ομογενείς συντεταγμένες των κορυφών του πρότυπου αντικειμένου στο στιγμιότυπο. Οι συντεταγμένες pixel αυτών προκύπτουν με διαίρεση με την τρίτη ομογενή συντεταγμένη τους (βλ. σχέσεις (5.8) και (5.9)).

$$\begin{bmatrix} x'_{frame} \\ y'_{frame} \\ w'_{frame} \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{bmatrix} \cdot \begin{bmatrix} x_{pattern} \\ y_{pattern} \\ 1 \end{bmatrix} \tag{5.8}$$

$$\begin{aligned} x_{frame} &= x'_{frame} / w'_{frame} \\ y_{frame} &= y'_{frame} / w'_{frame} \end{aligned} \tag{5.9}$$

Στη συνέχεια, υπολογίζεται ο εξωτερικός προσανατολισμός του συγκεκριμένου στιγμιότυπου με δεδομένα τις συντεταγμένες pixel των τεσσάρων κορυφών του πρότυπου επίπεδου αντικειμένου στο στιγμιότυπο, τις συντεταγμένες αυτών στο επίγειο σύστημα αναφοράς και τα στοιχεία του εσωτερικού προσανατολισμού, τα οποία δίνονται από τον 3x3 πίνακα της μηχανής και από το πέντε διαστάσεων διάνυσμα με τους συντελεστές της διαστροφής. Ο υπολογισμός των στοιχείων του εξωτερικού προσανατολισμού πραγματοποιείται μέσω συνάρτησης της βιβλιοθήκης OpenCV, η οποία λύνει το πρόβλημα «Perspective-n-Point» («PnP»). Η διαδικασία επίλυσης αυτού του προβλήματος παρουσιάστηκε στην ενότητα 3.6. Πρόκειται για υπολογισμό του εξωτερικού προσανατολισμού στιγμιότυπου από ομολογίες σημείων τα οποία ορίζουν ένα δισδιάστατο σημειοσύνολο, αφού και οι τέσσερις κορυφές του πρότυπου αντικειμένου βρίσκονται στο ίδιο επίπεδο. Αποτέλεσμα της επίλυσης του προβλήματος «PnP» είναι το διάνυσμα μετάθεσης **t** και το διάνυσμα στροφής **r**, τα οποία ορίζουν τη μετάθεση και τη στροφή που πρέπει να υποστεί το επίγειο σύστημα συντεταγμένων για να ταυτιστεί με το φωτογραμμετρικό. Η αρχή του επίγειου συστήματος συντεταγμένων ορίστηκε στο κέντρο του επίπεδου πρότυπου αντικειμένου, με άξονες X και Y παράλληλους στους άξονες της εικόνας - πρότυπο και άξονα Z κάθετο στο επίπεδο του αντικειμένου (βλ. Εικόνα 5-6, δεξιά). Το φωτογραμμετρικό σύστημα είναι εκείνο το οποίο ορίζεται από την OpenCV (βλ. Εικόνα 4-2, δεξιά).

Όπως αναφέρθηκε στην υποενότητα 4.3.3, το θεωρούμενο φωτογραμμετρικό σύστημα συντεταγμένων της OpenCV είναι αριστερόστροφο, ενώ το φωτογραμμετρικό σύστημα που θεωρείται από την OpenGL είναι δεξιόστροφο. Συγκεκριμένα, οι άξονες Y και Z των δύο συστημάτων έχουν αντίθετη φορά, όπως φαίνεται στην Εικόνα 4-2. Συνεπώς, προκειμένου τα αποτελέσματα αυτά να χρησιμοποιηθούν για την κατασκευή του πίνακα του μετασχηματισμού μοντέλου - κάμερας της OpenGL, πρέπει να μεταβληθούν τα πρόσημα των μεταθέσεων κατά Y και Z αλλά και των στροφών κατά Y και Z του επίγειου συστήματος ώστε να ταυτιστεί με το φωτογραμμετρικό. Συνεπώς, το διάνυσμα στροφής προκειμένου να αναφέρεται στον ορισμό του φωτογραμμετρικού συστήματος της OpenGL μετασχηματίζεται σύμφωνα με τη σχέση (5.10)[15]. Στη συνέχεια, μετατρέπεται σε 3x3 πίνακα στροφής, σύμφωνα με τις σχέσεις (3.111) και (3.112).

[15] Η αλλαγή των προσήμων των δύο τελευταίων στοιχείων του διανύσματος στροφής ισοδυναμεί με αλλαγή των προσήμων των στροφών κατά Y και Z του επίγειου συστήματος για να ταυτιστεί με το φωτογραμμετρικό.

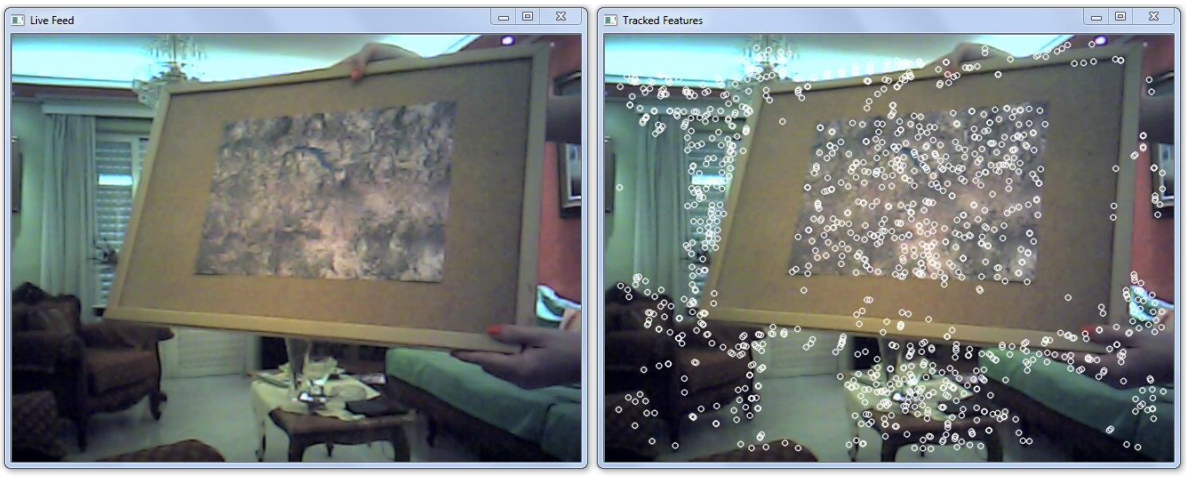
$$\begin{bmatrix} r_{OpenGL_1} \\ r_{OpenGL_2} \\ r_{OpenGL_3} \end{bmatrix} = \begin{bmatrix} r_{OpenCV_1} \\ -r_{OpenCV_2} \\ -r_{OpenCV_3} \end{bmatrix} \quad (5.10)$$

Ανάλογα, το διάνυσμα μετάθεσης προκειμένου να αναφέρεται στη μετάθεση από το επίγειο σύστημα στο φωτογραμμετρικό όπως ορίζεται στην OpenGL, μετασχηματίζεται ως εξής:

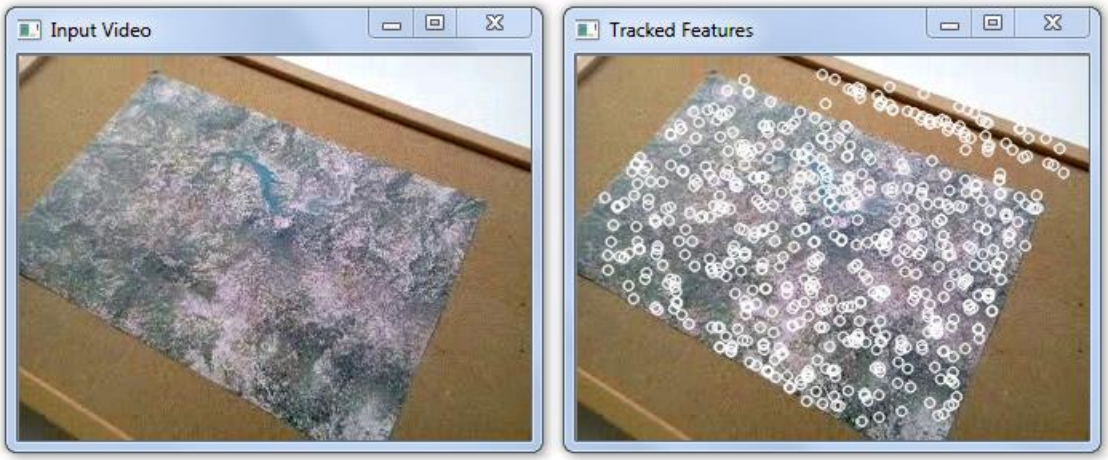
$$\begin{bmatrix} t_{OpenGL_1} \\ t_{OpenGL_2} \\ t_{OpenGL_3} \end{bmatrix} = \begin{bmatrix} t_{OpenCV_1} \\ -t_{OpenCV_2} \\ -t_{OpenCV_3} \end{bmatrix} \quad (5.11)$$

5.4.3.2 Εμφάνιση των κατάλληλων παραθύρων μέσω της OpenCV

Ο υπολογισμός του εξωτερικού προσανατολισμού του στιγμιότυπου είναι το τελικό αποτέλεσμα το οποίο προσδοκάται να επιτευχθεί μέσω της βιβλιοθήκης OpenCV. Ωστόσο, προτού προχωρήσει η διαδικασία επαύξησης της πραγματικότητας με τη βοήθεια της OpenGL, προγραμματίζεται η εμφάνιση κάποιων παραθύρων μέσω της OpenCV, τα οποία δείχνουν τα κύρια αποτελέσματα της διαδικασίας που ακολουθήθηκε μέσω της βιβλιοθήκης αυτής, εφόσον ο χρήστης το επέλεξε στην αρχή της εφαρμογής. Συγκεκριμένα, στην περίπτωση κατά την οποία οι ομολογίες που προέκυψαν είναι περισσότερες των πέντε, δημιουργείται ένα παράθυρο με τίτλο που ήδη είχε καθοριστεί ανάλογα με το αν πρόκειται για επαύξηση αρχείου βίντεο («Input Video») ή ζωντανής μετάδοσης της κάμερας του υπολογιστή («Live Feed») και σε αυτό εμφανίζεται το αντίστοιχο στιγμιότυπο. Επίσης, δημιουργείται ένα παράθυρο με τίτλο «Tracked Features» για την απεικόνιση του στιγμιότυπου στο οποίο έχουν σχεδιαστεί τα σημεία ενδιαφέροντος που εντοπίστηκαν από τον αλγόριθμο SURF (βλ. Εικόνες 5-9 και 5-10).



Εικόνα 5-9. Παράθυρο «Live Feed» [αριστερά], στο οποίο απεικονίζονται τα ζωντανά στιγμιότυπα της κάμερας, και παράθυρο «Tracked Features» [δεξιά], στο οποίο απεικονίζονται τα στιγμιότυπα με τα σημεία ενδιαφέροντος που εντοπίστηκαν σε αυτά, για μία τυχαία σκηνή της ζωντανής μετάδοσης της κάμερας για την Εφαρμογή 1.



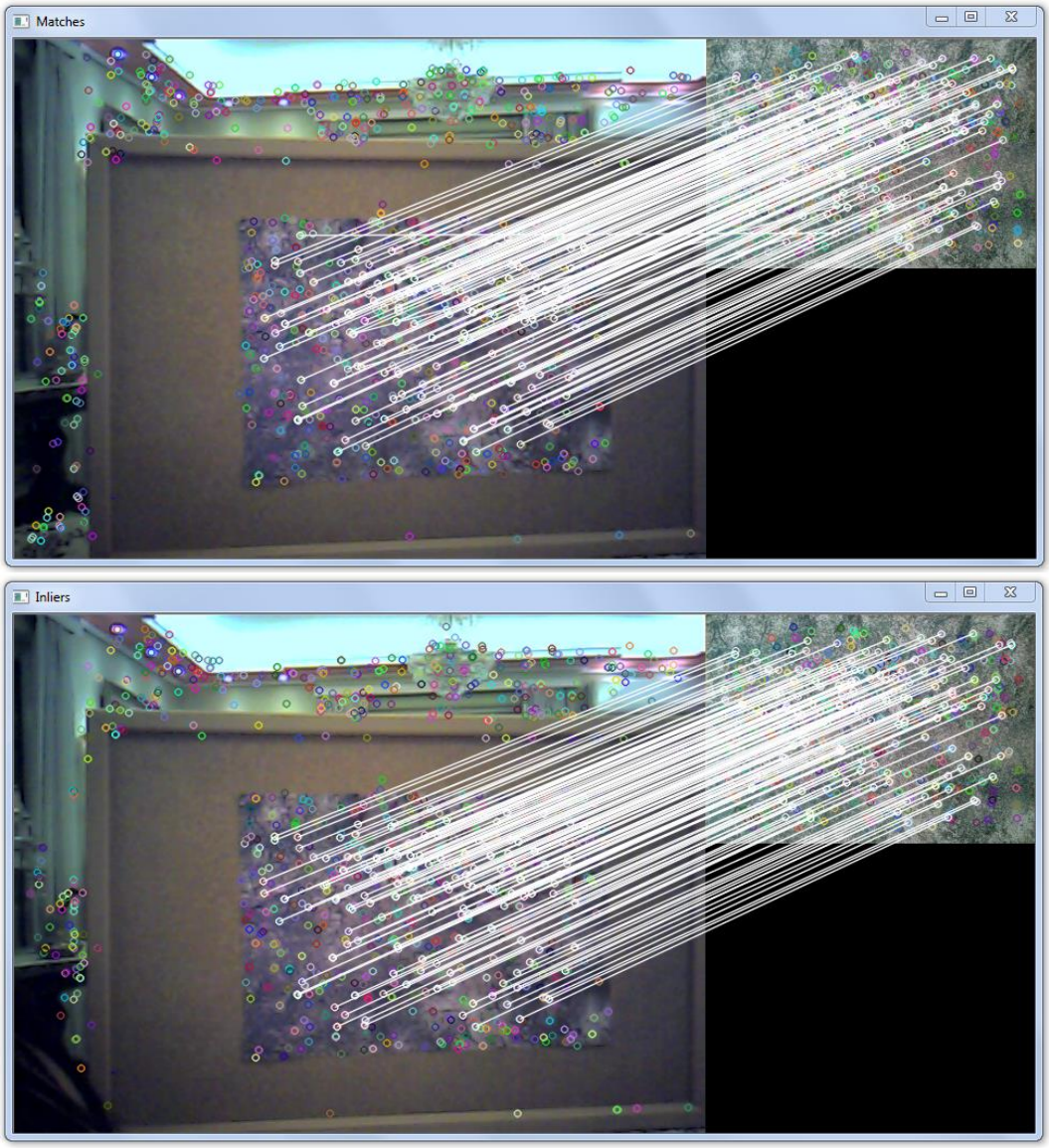
Εικόνα 5-10. Παράθυρο «Input Video» [αριστερά], στο οποίο απεικονίζονται τα στιγμιότυπα του αρχείου βίντεο, και παράθυρο «Tracked Features» [δεξιά], στο οποίο απεικονίζονται τα στιγμιότυπα του βίντεο με τα σημεία ενδιαφέροντος που εντοπίστηκαν σε αυτά, για μία τυχαία σκηνή ενός αρχείου βίντεο της Εφαρμογής 1.

Επίσης, εμφανίζονται οι ομολογίες των σημείων που προέκυψαν από τη διαδικασία της συνταύτισης μετά το διασταυρωμένο έλεγχο και την αποδοχή εκείνων για τις οποίες η απόσταση των διανυσμάτων περιγραφικών χαρακτηριστικών των αντίστοιχων σημείων είναι μικρότερη του ορισθέντος κατωφλίου (0.22). Το παράθυρο που δημιουργείται για το σκοπό αυτό έχει τίτλο «Matches». Ωστόσο, η ομογραφία δεν υπολογίζεται από το σύνολο αυτών των ομολογιών, αλλά από τα inliers, που επαληθεύουν την ευρεθείσα γεωμετρική σχέση μεταξύ της εικόνας - πρότυπο και του στιγμιότυπου. Οι έγκυρες αυτές συνταυτίσεις εμφανίζονται σε παράθυρο με τίτλο «Inliers». Στις Εικόνες 5-11 και 5-12 παρουσιάζονται τα δύο παραπάνω παράθυρα για τυχαίο στιγμιότυπο της κάμερας και του βίντεο αντίστοιχα και καθίσταται φανερή η απόρριψη των άκυρων αντιστοιχίσεων που δεν επαληθεύουν τη γεωμετρική σχέση ομογραφίας μεταξύ του στιγμιότυπου και της εικόνας - πρότυπο.

Προκειμένου να καταστεί εύκολος ο έλεγχος της ορθότητας του πίνακα της ομογραφίας από το χρήστη, πραγματοποιείται αναδόμηση του στιγμιότυπου με βάση τον πίνακα ομογραφίας ώστε αυτό να δείχνει το πρότυπο επίπεδο αντικείμενο. Αν η τελική αυτή αναδομημένη εικόνα εικονίζει το πρότυπο (ή τμήμα αυτού) το οποίο βρίσκεται στο στιγμιότυπο, ο πίνακας της ομογραφίας που υπολογίστηκε είναι σωστός. Η πορεία της επανασύστασης είναι από την τελική στην αρχική εικόνα, δηλαδή από την αναδομημένη εικόνα στο στιγμιότυπο. Συγκεκριμένα, δημιουργείται μία εικόνα με διαστάσεις ίσες με τις διαστάσεις της εικόνας - πρότυπο και αυτή αποτελεί τη ζητούμενη αναδομημένη εικόνα. Για κάθε σημείο (x,y) της εικόνας αυτής (το οποίο αντιστοιχεί στο σημείο (x,y) της εικόνας - πρότυπο) υπολογίζονται οι συντεταγμένες του στιγμιότυπου (x_{frame}, y_{frame}) με βάση το δισδιάστατο προβολικό μετασχηματισμό, ως εξής:

$$x_{frame} = \frac{h_{11} \cdot x + h_{12} \cdot y + h_{13}}{h_{31} \cdot x + h_{32} \cdot y + 1} \tag{5.12}$$

$$y_{frame} = \frac{h_{21} \cdot x + h_{22} \cdot y + h_{23}}{h_{31} \cdot x + h_{32} \cdot y + 1} \tag{5.13}$$



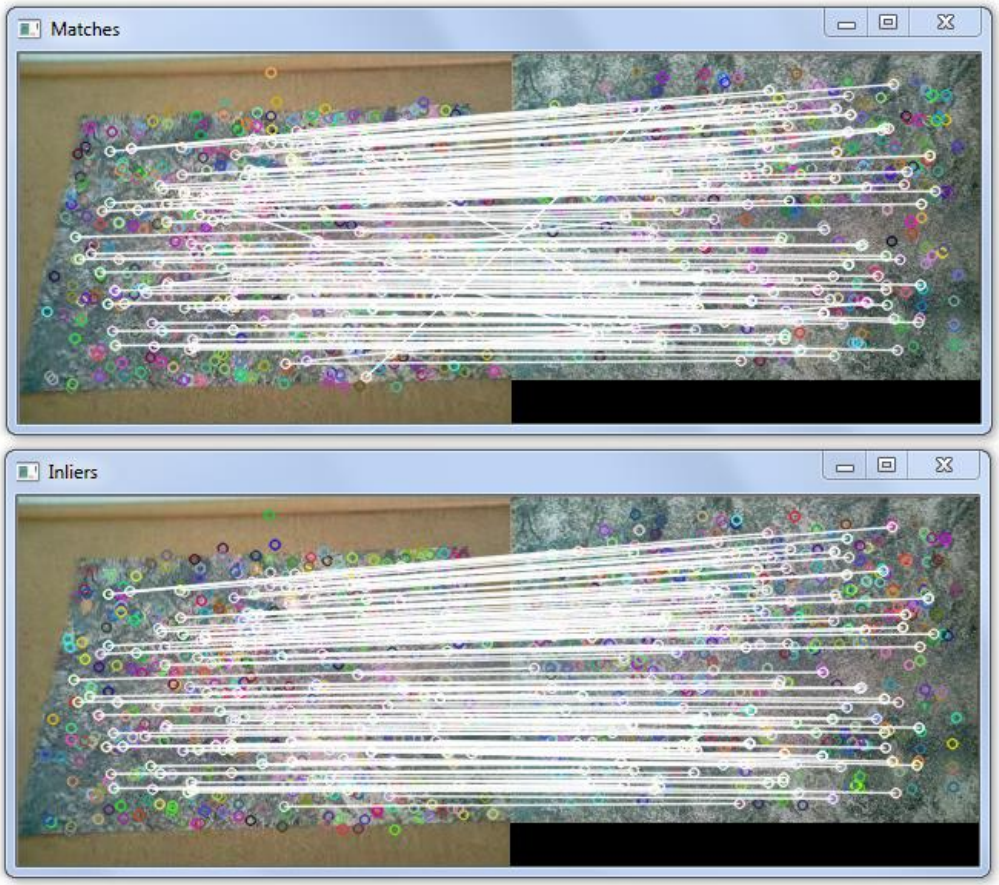
Εικόνα 5-11. Παράθυρο «Matches» με τις ομολογίες που προέκυψαν μετά το διασταυρωμένο έλεγχο και την κατωφλίωση και αντίστοιχο παράθυρο «Inliers» με τις έγκυρες ομολογίες που προέκυψαν από τον αλγόριθμο RANSAC και χρησιμοποιήθηκαν για τον υπολογισμό του πίνακα της ομογραφίας, για ένα τυχαίο στιγμιότυπο της μετάδοσης της κάμερας του υπολογιστή (Εφαρμογή 1).

Επειδή κατά βάση οι συντεταγμένες αυτές δεν είναι ακέραιες, δηλαδή δεν αντιστοιχούν σε κάποιο pixel της εικόνας, δεν είναι δυνατή η άμεση εύρεση της τιμής του χρώματος που πρέπει να ληφθεί από το στιγμιότυπο για να αποτελέσει το χρώμα της νέας αναδομημένης εικόνας. Συνεπώς, καθίσταται αναγκαία μία μέθοδος παρεμβολής για την περάτωση της επανασύστασης του στιγμιότυπου. Στο πλαίσιο της εφαρμογής, επιλέγεται η μέθοδος του εγγύτερου γείτονα ως τεχνική παρεμβολής, σύμφωνα με την οποία η τιμή του χρώματος της αναδομημένης εικόνας είναι η τιμή του χρώματος του εγγύτερου εικονοστοιχείου του στιγμιότυπου στο σημείο με συντεταγμένες (x_{frame}, y_{frame}), που υπολογίστηκε με βάση την ομογραφία. Αν συμβολιστούν ως (x'_{frame}, y'_{frame}) οι ακέραιες κοντινότερες συντεταγμένες

pixel των υπολογισμένων συντεταγμένων με βάση την ομογραφία (x_{frame}, y_{frame}), η αναδόμηση πραγματοποιείται ως εξής:

$$\text{αναδομημένη_εικόνα}\left[x, y\right] = \text{στιγμιότυπο}\left[x'_{frame}, y'_{frame}\right] \quad (5.14)$$

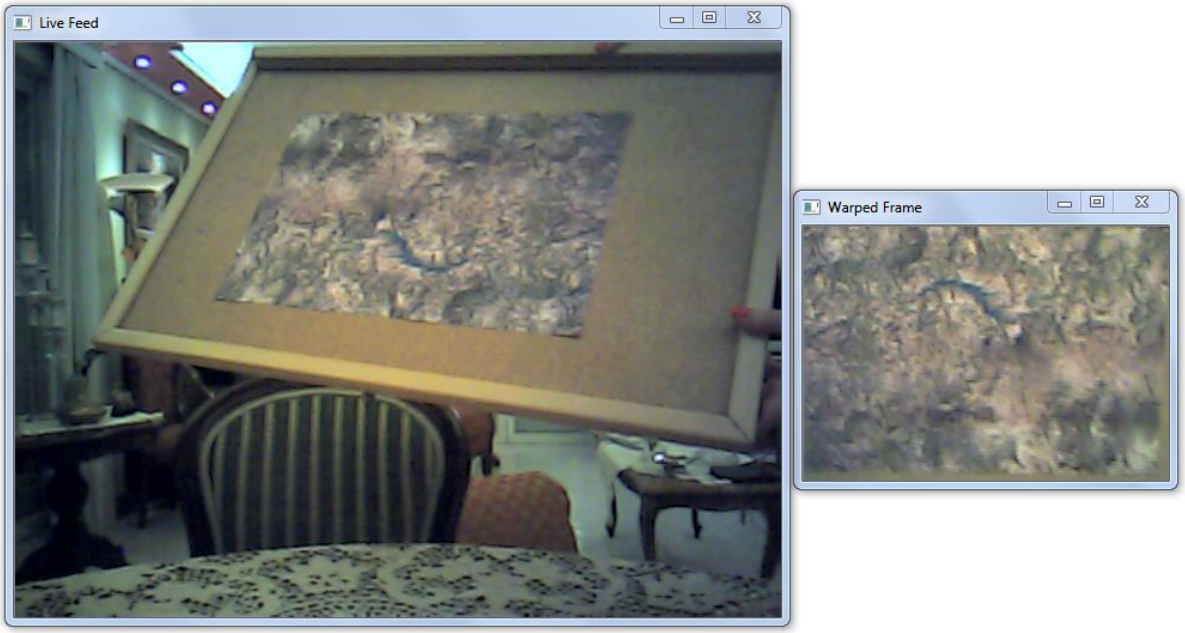
όπου το πρώτο μέλος της ισότητας δηλώνει την τιμή χρώματος της αναδομημένης εικόνας στο pixel της με συντεταγμένες (x,y), ενώ το δεύτερο μέλος της δηλώνει την τιμή χρώματος του στιγμιότυπου στο εικονοσημείο (x'_{frame}, y'_{frame}).



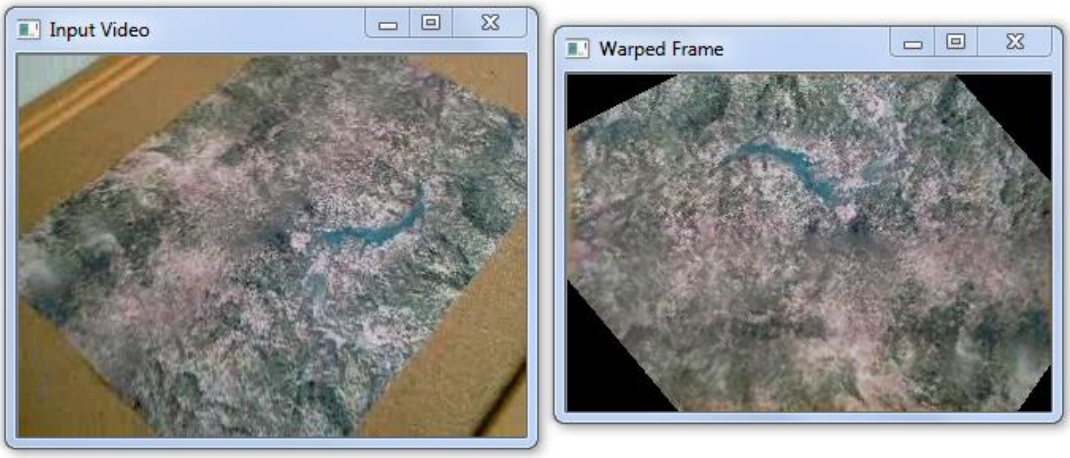
Εικόνα 5-12. Παράθυρο «Matches» με τις ομολογίες που προέκυψαν μετά το διασταυρωμένο έλεγχο και την κατωφλίωση και αντίστοιχο παράθυρο «Inliers» με τις έγκυρες ομολογίες που προέκυψαν από τον αλγόριθμο RANSAC και χρησιμοποιήθηκαν για τον υπολογισμό του πίνακα της ομογραφίας, για μία τυχαία σκηνή ενός αρχείου βίντεο της Εφαρμογής 1.

Η μέθοδος του εγγύτερου γείτονα επιλέχτηκε λόγω των μικρότερων υπολογιστικών της απαιτήσεων και της συνακόλουθης γρηγορότερης ολοκλήρωσης της αναδόμησης μέσω αυτής. Άλλες συχνά χρησιμοποιούμενες μέθοδοι παρεμβολής αποτελούν η διγραμμική και η δικυβική παρεμβολή, στις οποίες η τιμή χρώματος του pixel της αναδομημένης εικόνας υπολογίζεται από τις τιμές χρώματος των τεσσάρων και των δεκαέξι αντίστοιχα εγγύτερων εικονοστοιχείων στο εικονοσημείο (x_{frame}, y_{frame}). Ωστόσο, ενώ αυτές παράγουν καλύτερα οπτικά αποτελέσματα στην ανακατασκευασμένη εικόνα και υπάρχει η δυνατότητα να χρησιμοποιηθούν μέσω της βιβλιοθήκης OpenCV, είναι περισσότερο πολύπλοκες και συνεπώς υπολογιστικά απαιτητικότερες και έτσι δεν επιλέγονται για την απλή εμφάνιση στο

χρήστη της αναδομημένης εικόνας του στιγμιότυπου που δείχνει το πρότυπο αντικείμενο που εντοπίστηκε σε αυτό. Η εν λόγω εικόνα παρουσιάζεται σε παράθυρο με τίτλο «Warped Frame». Στην Εικόνα 5-13, δεξιά παρουσιάζεται η αναδομημένη εικόνα εντός του παραθύρου αυτού, για το αντίστοιχο στιγμιότυπο της ζωντανής μετάδοσης της κάμερας που εικονίζεται στο παράθυρο «Live Feed» στην ίδια εικόνα, αριστερά. Στην Εικόνα 5-14, δεξιά παρουσιάζεται η αναδομημένη εικόνα εντός του παραθύρου «Warped Frame», για το αντίστοιχο στιγμιότυπο του βίντεο που εικονίζεται στο παράθυρο «Input Video» στην ίδια εικόνα, αριστερά.

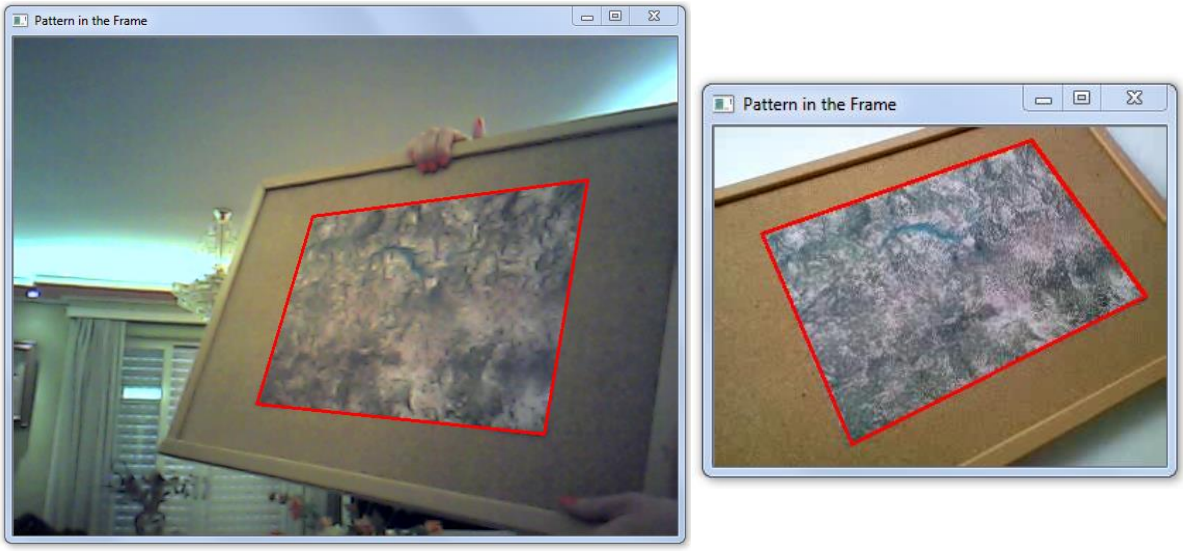


Εικόνα 5-13. Παράθυρο «Live Feed» [αριστερά], που εικονίζει ένα τυχαίο στιγμιότυπο της ζωντανής μετάδοσης της κάμερας του υπολογιστή, και παράθυρο «Warped Frame» [δεξιά], που εικονίζει το αντίστοιχο αναδομημένο στιγμιότυπο που δείχνει το πρότυπο επίπεδο αντικείμενο (Εφαρμογή 1).



Εικόνα 5-14. Παράθυρο «Input Video» [αριστερά], που εικονίζει ένα τυχαίο στιγμιότυπο του βίντεο, και παράθυρο «Warped Frame» [δεξιά], που εικονίζει το αντίστοιχο αναδομημένο στιγμιότυπο που δείχνει το πρότυπο επίπεδο αντικείμενο (Εφαρμογή 1).

Τέλος, δημιουργείται ένα ακόμη παράθυρο, το οποίο παρουσιάζει το κάθε στιγμιότυπο του βίντεο ή της ζωντανής μετάδοσης της κάμερας με μία κλειστή κόκκινη γραμμή σχεδιασμένη σε αυτό, η οποία συνδέει τις τέσσερις κορυφές του πρότυπου αντικειμένου που εντοπίστηκαν στο στιγμιότυπο (βλ. Εικόνα 5-15).



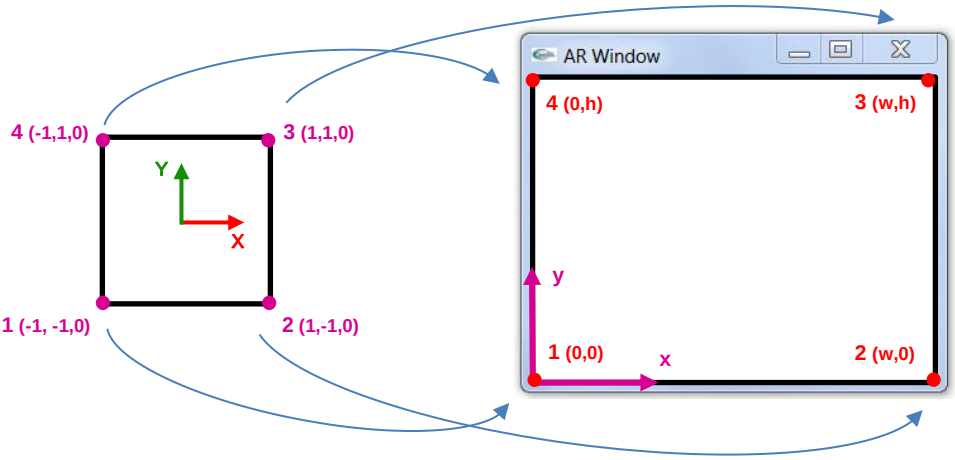
Εικόνα 5-15. Στιγμιότυπα από το παράθυρο «Pattern in the Frame» της Εφαρμογής 1, με μία τυχαία σκηνή της ζωντανής μετάδοσης της κάμερας [αριστερά] και ενός αρχείου βίντεο [δεξιά]. Σε κάθε στιγμιότυπο έχει σχεδιαστεί μία κόκκινη κλειστή γραμμή που ενώνει τις τέσσερις κορυφές του πρότυπου αντικειμένου.

5.4.3.3 Σχεδίαση του στιγμιότυπου ως φόντου της επιφάνειας σχεδίασης του παραθύρου επαύξησης της πραγματικότητας

Επόμενο βήμα είναι η σχεδίαση του στιγμιότυπου του βίντεο ή της κάμερας ως φόντου της επιφάνειας σχεδίασης του παραθύρου «AR Window», που δημιουργήθηκε με τη βιβλιοθήκη freeglut. Αυτό γίνεται μέσω μίας διαδικασίας που ονομάζεται από την OpenGL απόδοση υφής («texture mapping»). Η βασική ιδέα πίσω από τον ορισμό του εκάστοτε στιγμιότυπου ως φόντου στη σχεδιαστική επιφάνεια του παραθύρου αυτού είναι η δημιουργία ενός ορθογωνίου παραλληλεπιπέδου το οποίο καλύπτει όλη τη σχεδιαστική επιφάνεια του παραθύρου «AR Window» και η σχεδίαση πάνω σε αυτό της «υφής» του εκάστοτε στιγμιότυπου. Η διαδικασία αυτή περιγράφεται αναλυτικότερα στη συνέχεια.

Στην ενότητα 4.4 περιγράφηκαν οι μετασχηματισμοί συντεταγμένων που πραγματοποιούνται στην OpenGL ώστε ένα αντικείμενο να σχεδιαστεί στη σκηνή. Στην προκειμένη περίπτωση, το αντικείμενο αυτό είναι ένα τετράγωνο επίπεδο αντικείμενο. Αρχικά δίνονται οι συντεταγμένες των τεσσάρων κορυφών του στο τοπικό σύστημα του μοντέλου, το κέντρο του οποίου θεωρείται στο κέντρο του τετράγωνου επίπεδου αντικειμένου, όπως φαίνεται στην Εικόνα 5-16, αριστερά. Επειδή πρόκειται για επίπεδο αντικείμενο, η τρίτη συντεταγμένη του θεωρείται μηδενική. Οι συντεταγμένες αυτές είναι οι εξής:

$$\begin{aligned} X_{MODEL_1} &= -1 & Y_{MODEL_1} &= -1 & Z_{MODEL_1} &= 0 \\ X_{MODEL_2} &= 1 & Y_{MODEL_2} &= -1 & Z_{MODEL_2} &= 0 \\ X_{MODEL_3} &= 1 & Y_{MODEL_3} &= 1 & Z_{MODEL_3} &= 0 \\ X_{MODEL_4} &= -1 & Y_{MODEL_4} &= 1 & Z_{MODEL_4} &= 0 \end{aligned} \quad (5.15)$$



Εικόνα 5-16. Θεωρούμενο σύστημα συντεταγμένων μοντέλου [αριστερά] και σύστημα συντεταγμένων παραθύρου [δεξιά]. Σχεδίαση του τετραγώνου ως ορθογωνίου που καλύπτει όλη τη διαθέσιμη σχεδιαστική επιφάνεια του παραθύρου.

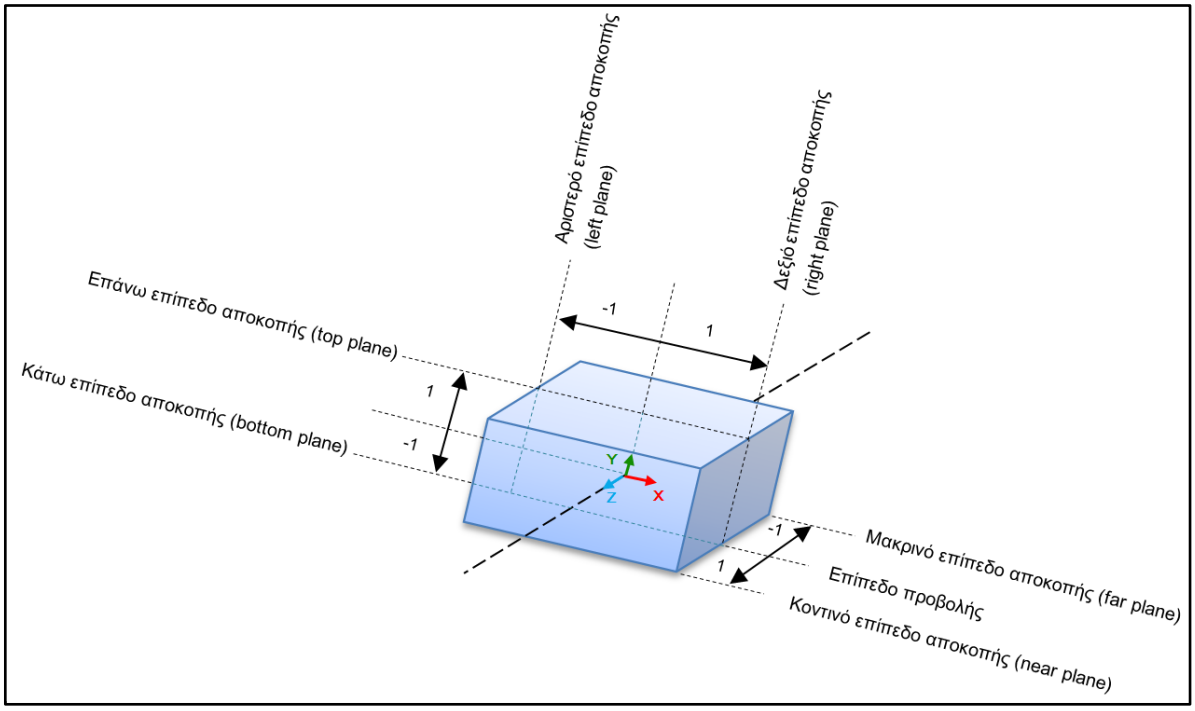
Το σύστημα του μοντέλου θεωρείται ότι ταυτίζεται με το φωτογραμμετρικό σύστημα όπως ορίζεται στην OpenGL. Συνεπώς, ο πίνακας μοντέλου - κάμερας που μετασχηματίζει τις συντεταγμένες από το σύστημα του μοντέλου στο φωτογραμμετρικό σύστημα ορίζεται ως ο μοναδιαίος 4x4 πίνακας. Επομένως, οι συντεταγμένες των τεσσάρων κορυφών του τετράγωνου αντικειμένου στο φωτογραμμετρικό σύστημα είναι οι ίδιες με τις συντεταγμένες τους στο σύστημα του μοντέλου.

Για την εύρεση των συντεταγμένων αποκοπής θεωρείται ορθογραφική προβολή και συμμετρικός όγκος θέασης και οι αποστάσεις right, left, top και bottom θεωρούνται μοναδιαίες. Συνεπώς, προκειμένου να προκύψει η συντεταγμένη αποκοπής Z_{CLIP} μηδενική, όπως γίνεται φανερό από τη σχέση (4.52) που ισχύει για ορθογραφική προβολή και συμμετρικό όγκο θέασης, πρέπει οι αποστάσεις near και far να έχουν ίδια τιμή με αντίθετο πρόσημο. Στο πλαίσιο αυτό, θεωρείται ότι η αρχή του φωτογραμμετρικού συστήματος, που ταυτίζεται με την αρχή του συστήματος του μοντέλου, λαμβάνεται στο κέντρο του συμμετρικού όγκου θέασης, με αποτέλεσμα το κοντινό επίπεδο αποκοπής να βρίσκεται στην αντίθετη πλευρά σε σχέση με εκείνη που θεωρείται συνήθως.[16] Συνεπώς, η σχέση (4.52) για right=1, top=1, near=-1 (με αρνητικό πρόσημο επειδή η σχέση (4.52) προέκυψε θεωρώντας το κοντινό επίπεδο αποκοπής από την αντίθετη πλευρά σε σχέση με το θεωρούμενο στην προκειμένη περίπτωση, όπως φαίνεται στην Εικόνα 5-17) και far=1 γίνεται:

[16] Ο πίνακας προοπτικής προβολής έχει δημιουργηθεί θεωρώντας τις τιμές right, left, top, bottom, near και far θετικές.

$$\begin{bmatrix} X_{CLIP} \\ Y_{CLIP} \\ Z_{CLIP} \\ W_{CLIP} \end{bmatrix} = \begin{bmatrix} \frac{1}{1} & 0 & 0 & 0 \\ 0 & \frac{1}{1} & 0 & 0 \\ 0 & 0 & \frac{2}{-1-1} & \frac{-1+1}{-1-1} \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} X_{EYE} \\ Y_{EYE} \\ Z_{EYE} \\ 1 \end{bmatrix} \rightarrow \begin{bmatrix} X_{CLIP} \\ Y_{CLIP} \\ Z_{CLIP} \\ W_{CLIP} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} X_{EYE} \\ Y_{EYE} \\ Z_{EYE} \\ 1 \end{bmatrix} \rightarrow$$

$$\begin{aligned} X_{CLIP} &= X_{EYE} = X_{MODEL} \\ Y_{CLIP} &= Y_{EYE} = Y_{MODEL} \\ Z_{CLIP} &= -Z_{EYE} = -Z_{MODEL} = 0 \\ W_{CLIP} &= 1 \end{aligned} \tag{5.16}$$



Εικόνα 5-17. Αρχή του φωτογραμμετρικού συστήματος συντεταγμένων που θεωρήθηκε στην περίπτωση της απεικόνισης του τετραγώνου στο παράθυρο μέσω ορθογραφικής προβολής, όγκος θέασης και επίπεδα αποκοπής.

Οι κανονικοποιημένες συντεταγμένες συσκευής στην περίπτωση της ορθογραφικής προβολής είναι ίδιες με τις συντεταγμένες αποκοπής, καθώς προκύπτουν με διαίρεση με την τέταρτη ομογενή συντεταγμένη, η οποία είναι ίση με τη μονάδα:

$$\begin{aligned} X_{NDC_1} &= -1 & Y_{NDC_1} &= -1 & Z_{NDC_1} &= 0 \\ X_{NDC_2} &= 1 & Y_{NDC_2} &= -1 & Z_{NDC_2} &= 0 \\ X_{NDC_3} &= 1 & Y_{NDC_3} &= 1 & Z_{NDC_3} &= 0 \\ X_{NDC_4} &= -1 & Y_{NDC_4} &= 1 & Z_{NDC_4} &= 0 \end{aligned} \tag{5.17}$$

Τέλος, οι εξισώσεις (4.56) και (4.57) του μετασχηματισμού παραθύρου παρατήρησης, θεωρώντας ότι $x_1=y_1=0$ και ότι οι μεταβλητές width και height είναι το πλάτος (w) και ύψος

(h) του παραθύρου αντίστοιχα (τα οποία από τα αρχικά στάδια της εφαρμογής καθορίστηκαν ίσα με το πλάτος και το ύψος των στιγμιότυπων), γίνονται:

$$\begin{aligned} x &= \frac{w}{2} \cdot X_{NDC} + \frac{w}{2} \\ y &= \frac{h}{2} \cdot Y_{NDC} + \frac{h}{2} \end{aligned} \tag{5.18}$$

Συνεπώς, οι συντεταγμένες παραθύρου παρατήρησης των τεσσάρων κορυφών του τετραγώνου είναι οι εξής:

$$\begin{aligned} x_1 &= \frac{w}{2} \cdot (-1) + \frac{w}{2} = 0 \qquad y_1 = \frac{h}{2} \cdot (-1) + \frac{h}{2} = 0 \\ x_2 &= \frac{w}{2} \cdot 1 + \frac{w}{2} = w \qquad y_2 = \frac{h}{2} \cdot (-1) + \frac{h}{2} = 0 \\ x_3 &= \frac{w}{2} \cdot 1 + \frac{w}{2} = w \qquad y_3 = \frac{h}{2} \cdot 1 + \frac{h}{2} = h \\ x_4 &= \frac{w}{2} \cdot (-1) + \frac{w}{2} = 0 \qquad y_4 = \frac{h}{2} \cdot 1 + \frac{h}{2} = h \end{aligned} \tag{5.19}$$

Η αρχή του συστήματος συντεταγμένων του παραθύρου είναι το κάτω αριστερά σημείο του, ο άξονας x κατευθύνεται προς τα δεξιά και ο άξονας y κατευθύνεται προς τα επάνω (βλ. Εικόνα 4-10, δεξιά). Συνεπώς, από τους παραπάνω υπολογισμούς προκύπτει ότι τα σημεία 1, 2, 3 και 4 του τετράγωνου αντικειμένου αντιστοιχούν στις τέσσερις κορυφές του παραθύρου (βλ. Εικόνα 5-16). Αυτό συμβαίνει εφόσον ο χρήστης δεν έχει μεταβάλει το μέγεθος του παραθύρου, καθώς αν αυτό μεταβληθεί, η σχεδιαστική επιφάνεια του παραθύρου δεν είναι – γενικά – ολόκληρη η διαθέσιμη επιφάνεια σχεδίασης αυτού, καθώς επιδιώκεται η διατήρηση του λόγου των πλευρών του στιγμιότυπου, όπως αναφέρθηκε στην υποενότητα 5.4.2. Στην περίπτωση αυτή, τα σημεία 1, 2, 3 και 4 του τετράγωνου αντικειμένου αντιστοιχούν στις τέσσερις κορυφές της σχεδιαστικής επιφάνειας του παραθύρου, σύμφωνα με τη συνάρτηση που καλείται όταν ο χρήστης μεταβάλλει το μέγεθος του παραθύρου.

Έχοντας ορίσει τον τρόπο σχεδιασμού του τετράγωνου αντικειμένου ως ορθογωνίου στο παράθυρο «AR Window», έτσι ώστε να καλύπτει όλη τη διαθέσιμη σχεδιαστική επιφάνεια, απομένει ο ορισμός της υφής του τετράγωνου αντικειμένου, για να δημιουργηθεί το φόντο της σχεδιαστικής επιφάνειας του παραθύρου επαύξησης της πραγματικότητας. Η υφή αυτή είναι το εκάστοτε στιγμιότυπο του βίντεο. Σημαντικό βήμα για την περάτωση της παραπάνω διαδικασίας αποτελεί ο ορισμός των συντεταγμένων υφής των τεσσάρων κορυφών του τετράγωνου αντικειμένου. Η αρχή του συστήματος συντεταγμένων υφής είναι το κάτω αριστερά pixel, ενώ το εύρος τιμών των συντεταγμένων υφής είναι το διάστημα [0,1]. Συνεπώς, προκειμένου ολόκληρο το στιγμιότυπο να αποτελέσει την υφή του τετράγωνου

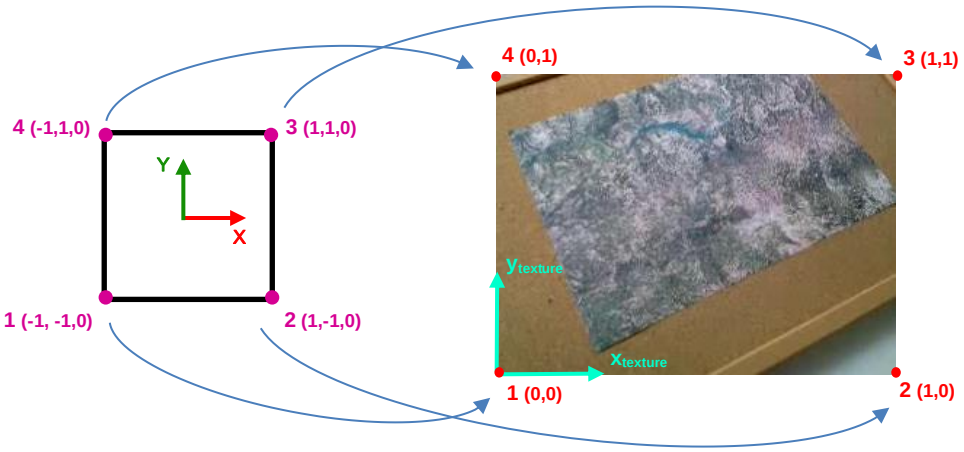
αντικειμένου, οι συντεταγμένες υφής των κορυφών του λαμβάνουν τις τιμές 0 ή 1 και είναι οι ακόλουθες (βλ. Εικόνα 5-18):

$$\begin{aligned} x_{texture_1} &= 0 & y_{texture_1} &= 0 \\ x_{texture_2} &= 1 & y_{texture_2} &= 0 \\ x_{texture_3} &= 1 & y_{texture_3} &= 1 \\ x_{texture_4} &= 0 & y_{texture_4} &= 1 \end{aligned} \tag{5.20}$$

Με τον καθορισμό των συντεταγμένων υφής των κορυφών του τετράγωνου αντικειμένου, οι συντεταγμένες υφής όλων των ενδιάμεσων σημείων προσδιορίζονται αυτόματα μέσω διγραμμικής παρεμβολής από την OpenGL. Η τελευταία για κάθε pixel της επιφάνειας σχεδίασης αναζητά τις συντεταγμένες υφής που του αντιστοιχούν. Συνεπώς, πραγματοποιούνται οι αντίστροφοι μετασχηματισμοί από το σύστημα συντεταγμένων του παραθύρου στο σύστημα συντεταγμένων του μοντέλου με την αντίστροφη σειρά και γίνεται η μετάβαση από τις συντεταγμένες παραθύρου στις κανονικοποιημένες συντεταγμένες συσκευής, εν συνεχεία στις συντεταγμένες αποκοπής, στις συντεταγμένες κάμερας και, τέλος, στις συντεταγμένες μοντέλου. Κατόπιν, από το σύστημα συντεταγμένων του μοντέλου πραγματοποιείται μετάβαση στο σύστημα συντεταγμένων υφής και εντοπίζεται το στοιχείο υφής (texture element ή texel) που θα αποδοθεί στο εκάστοτε pixel. Το γεγονός αυτό συνεπάγεται την πιθανότητα να ανατεθούν σε εικονοστοιχεία του παραθύρου συντεταγμένες υφής οι οποίες δεν συμπίπτουν ακριβώς με τις συντεταγμένες υφής κάποιου texel. Στην περίπτωση αυτή, επιλέγεται να εφαρμοστεί η ταχύτερη μέθοδος του εγγύτερου γείτονα ως μέθοδος παρεμβολής, σύμφωνα με την οποία η τιμή που αποδίδεται στο pixel είναι η τιμή χρώματος του πλησιέστερου στοιχείου υφής. Η μέθοδος του εγγύτερου γείτονα επιλέγεται ως μέθοδος παρεμβολής τόσο σε περιπτώσεις «μεγέθυνσης» της υφής (όταν αυτή μεγεθύνεται κατά την απόδοσή της), όσο και σε περιπτώσεις «σμίκρυνσης» αυτής (όταν πρέπει να σμικρυνθεί για να αποδοθεί) [197], [198].

Σύμφωνα με την παραπάνω ανάλυση, τα βήματα που έγιναν προκειμένου να οριστεί το στιγμιότυπο ως φόντο της επιφάνειας σχεδίασης του παραθύρου επαύξησης της πραγματικότητας παρουσιάζονται στη Διαγραμματική παρουσίαση 5-13 – χωρίς να δίνεται έμφαση στις προγραμματιστικές λεπτομέρειες για την απεικόνιση της υφής μέσω της OpenGL – ενώ στη συνέχεια περιγράφεται συνοπτικά η διαδικασία αυτή. Αρχικά ορίζεται η μέθοδος του εγγύτερου γείτονα ως μέθοδος παρεμβολής. Ακολούθως γίνεται αντικατοπτρισμός του στιγμιότυπου, που έχει «διαβαστεί» από τη βιβλιοθήκη OpenCV, κατά τον άξονα x, λόγω του διαφορετικού τρόπου αποθήκευσης των δεδομένων μεταξύ των OpenCV και OpenGL και ορίζεται ότι η δισδιάστατη υφή είναι το αντικατοπτρισμένο κατά τον άξονα x στιγμιότυπο. Κατόπιν, ορίζεται ο μοναδιαίος 4x4 πίνακας μοντέλου - κάμερας και ο 4x4 πίνακας προοπτικής προβολής, όπως περιγράφηκε παραπάνω, και στη συνέχεια καθορίζονται οι συντεταγμένες μοντέλου και υφής των τεσσάρων κορυφών του ορθογωνίου.

Ο μετασχηματισμός παραθύρου παρατήρησης δεν ορίζεται, γεγονός το οποίο σημαίνει ότι η OpenGL θεωρεί τον προεπιλεγμένο μετασχηματισμό παραθύρου παρατήρησης, σύμφωνα με τον οποίο η επιφάνεια σχεδίασης είναι ολόκληρη η διαθέσιμη επιφάνεια του παραθύρου. Ωστόσο, στην περίπτωση που το μέγεθος του παραθύρου έχει μεταβληθεί από το χρήστη, η επιφάνεια σχεδίασης είναι αυτή που καθορίζεται από τη συνάρτηση που καλείται όταν ο χρήστης μεταβάλλει το μέγεθος του παραθύρου.



Εικόνα 5-18. Συντεταγμένες υφής των κορυφών του τετραγώνου.

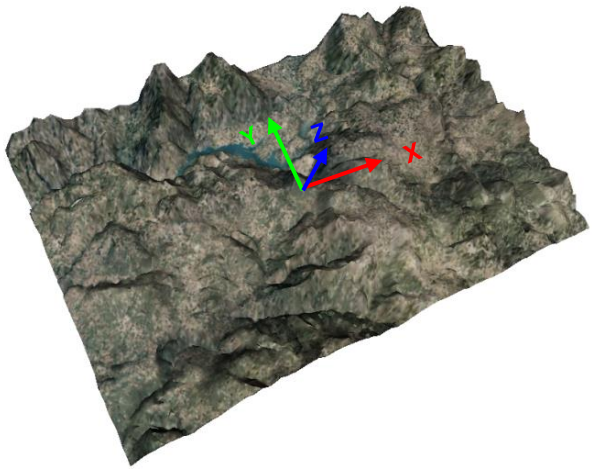
5.4.3.4 Σχεδίαση του τρισδιάστατου μοντέλου στο παράθυρο επαύξησης της πραγματικότητας

Τελικό βήμα της επαναληπτικής διαδικασίας είναι η σχεδίαση του τρισδιάστατου μοντέλου στη σωστή θέση, με την κατάλληλη κλίμακα και προοπτική και με σωστό προσανατολισμό στο παράθυρο «AR Window». Για το λόγο αυτό, αρχικά καθορίζεται ο πίνακας προοπτικής προβολής (βλ. σχέση 4.40) με βάση τα στοιχεία του εσωτερικού προσανατολισμού και τη θεώρηση τιμών για τις αποστάσεις των κοντινών και μακρινών επιπέδων αποκοπής, οι οποίες δεν επηρεάζουν το τελικό αποτέλεσμα εφόσον οι συντεταγμένες Z_{EYE} των κορυφών του μοντέλου στο σύστημα της κάμερας βρίσκονται μεταξύ του κοντινού και μακρινού επιπέδου αποκοπής. Τα στοιχεία του εσωτερικού προσανατολισμού είναι είτε αυτά που εισήγαγε ο χρήστης είτε τα προεπιλεγμένα από την εφαρμογή που προέκυψαν από τη βαθμονόμηση της κάμερας του υπολογιστή ή του κινητού τηλεφώνου, ανάλογα με την επιλογή του χρήστη για χρήση της κάμερας του υπολογιστή ή ενός αρχείου βίντεο για την επαύξηση του πραγματικού κόσμου.

Οι συντεταγμένες των κορυφών του τρισδιάστατου μοντέλου που βρίσκονται στο αρχείο OBJ δεν έχουν το εύρος τιμών [-1,1] και αναφέρονται σε ένα οποιοδήποτε σύστημα αναφοράς, το οποίο δεν ταυτίζεται με το επίγειο σύστημα αναφοράς που ορίστηκε στο πλαίσιο της εφαρμογής. Για το λόγο αυτό, γίνεται μία κανονικοποίηση των συντεταγμένων του ΨΜΕ στο εύρος [-1,1], ώστε το κέντρο του να τοποθετηθεί στο κέντρο του επίγειου

συστήματος συντεταγμένων (δηλαδή στο κέντρο του πρότυπου αντικειμένου, που είναι η ορθοφωτογραφία), το ΨΜΕ να αποκτήσει τις ίδιες διαστάσεις με την ορθοφωτογραφία, της οποίας οι συντεταγμένες των τεσσάρων κορυφών είναι στο εύρος [-1,1], και ο προσανατολισμός του να ταυτιστεί με τον προσανατολισμό της ορθοφωτογραφίας. Η παραπάνω διαδικασία πραγματοποιείται μέσω της βιβλιοθήκης GLM.

Στο σημείο αυτό κρίνεται σημαντική η επισήμανση μίας λεπτομέρειας. Στο νέο σύστημα συντεταγμένων του μοντέλου ο άξονας Y κατευθύνεται προς τα επάνω, ο άξονας X προς τα δεξιά και ο άξονας Z κατευθύνεται έξω από το μοντέλο (βλ. Εικόνα 5-19). Ωστόσο, το επίγειο σύστημα συντεταγμένων ορίστηκε με τον άξονα Y προς τα κάτω (βλ. Εικόνα 5-6, δεξιά). Ο ορισμός του επίγειου συστήματος συντεταγμένων με τον άξονα Y προς τα κάτω έγινε λόγω των διαφορετικών συστημάτων συντεταγμένων εικόνας των OpenCV και OpenGL. Ενώ στην OpenCV η αρχή του συστήματος συντεταγμένων της εικόνας είναι το επάνω αριστερά pixel, με αποτέλεσμα ο άξονας y να έχει φορά προς τα κάτω, στην OpenGL η αρχή του συστήματος συντεταγμένων της εικόνας είναι το κάτω αριστερά pixel, με αποτέλεσμα ο άξονας y να έχει φορά προς τα επάνω. Για την αντιστάθμιση της διαφοράς μεταξύ των δύο συστημάτων συντεταγμένων, ο άξονας Y του επίγειου συστήματος τοποθετείται με φορά αντίθετη σε σχέση με τον άξονα Y του συστήματος συντεταγμένων του μοντέλου. Έτσι, το ΨΜΕ τελικά σχεδιάζεται με σωστό προσανατολισμό, που είναι ίδιος με τον προσανατολισμό της ορθοφωτογραφίας στο εκάστοτε στιγμιότυπο. Ο ίδιος τρόπος ορισμού του επίγειου συστήματος συντεταγμένων ακολουθείται και στην Εφαρμογή 2 για τον ίδιο λόγο.



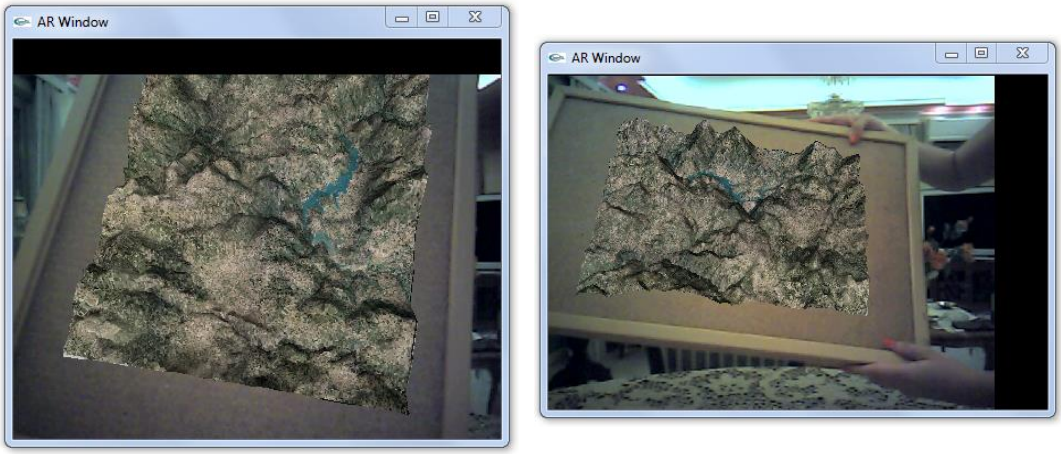
Εικόνα 5-19. Νέο σύστημα συντεταγμένων του μοντέλου, μετά το μετασχηματισμό που πραγματοποιείται μέσω της βιβλιοθήκης GLM.

Στη συνέχεια, ορίζεται ο πίνακας του μετασχηματισμού μοντέλου κάμερας (βλ. σχέση 4.2), τα στοιχεία του οποίου προκύπτουν από τα στοιχεία του πίνακα στροφής και του διανύσματος μετάθεσης, καθώς το επίγειο σύστημα συντεταγμένων ταυτίζεται με το σύστημα συντεταγμένων του μοντέλου. Ο πίνακας στροφής και το διάνυσμα μετάθεσης ορίζουν τη στροφή και τη μετάθεση του επίγειου συστήματος συντεταγμένων ώστε να ταυτιστεί με το φωτογραμμετρικό, όπως αυτό ορίζεται στην OpenGL.

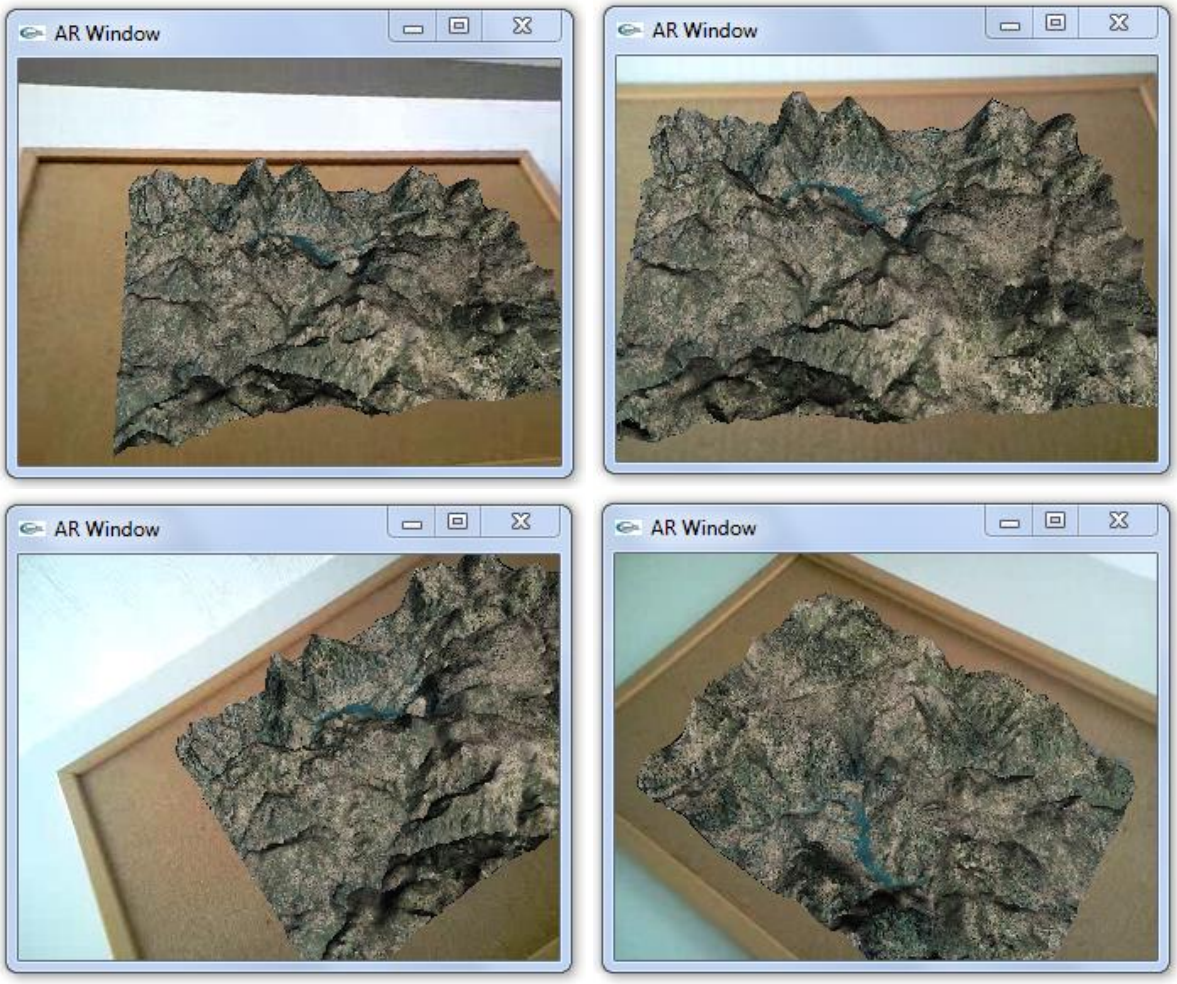
Στη συνέχεια, ενεργοποιείται ο φωτισμός του μοντέλου και επιλέγεται ο προκαθορισμένος από την OpenGL φωτισμός. Το τρισδιάστατο μοντέλο συνοδεύεται από μία εικόνα που αποτελεί την υφή του. Για το λόγο αυτό, γίνεται η απόδοση της υφής του μέσω της OpenGL με ανάλογη διαδικασία όπως περιγράφηκε προηγουμένως, με τη διαφορά ότι πλέον δεν εισάγονται οι συντεταγμένες μοντέλου και οι αντίστοιχες συντεταγμένες υφής, αφού έχουν ήδη διαβαστεί από τη βιβλιοθήκη GLM.

Στις Εικόνες 5-20, 5-21, 5-22 και 5-23 παρουσιάζονται στιγμιότυπα από το παράθυρο «AR Window», στο οποίο επαυξάνονται οι σκηνές των στιγμιότυπων της ζωντανής μετάδοσης της κάμερας και των αρχείων βίντεο. Στις Εικόνες 5-20 και 5-22 το μέγεθος της σχεδιαστικής επιφάνειας των παραθύρων που εικονίζονται είναι το ίδιο με το μέγεθος των στιγμιότυπων. Αντίθετα, το μέγεθος των παραθύρων των Εικόνων 5-21 και 5-23 έχει μεταβληθεί από το χρήστη. Στην περίπτωση αυτή, καλείται αυτόματα η συνάρτηση που έχει οριστεί να καλείται όταν μεταβάλλονται οι διαστάσεις του παραθύρου, ώστε ο λόγος των διαστάσεων των πλευρών του στιγμιότυπου να παραμένει σταθερός.

Εικόνα 5-20. Στιγμιότυπα από το παράθυρο «AR Window» της Εφαρμογής 1, στο οποίο επαυξάνονται οι σκηνές των στιγμιότυπων της ζωντανής μετάδοσης της κάμερας. Το μέγεθος της σχεδιαστικής επιφάνειας του παραθύρου είναι ίδιο με το μέγεθος των στιγμιότυπων της κάμερας του υπολογιστή.



Εικόνα 5-21. Στιγμιότυπα από το παράθυρο «AR Window» της Εφαρμογής 1, το μέγεθος του οποίου έχει μεταβληθεί από το χρήστη (περίπτωση επαύξησης της πραγματικότητας με χρήση της κάμερας του υπολογιστή σε πραγματικό χρόνο).



Εικόνα 5-22. Στιγμιότυπα από το παράθυρο «AR Window» της Εφαρμογής 1, στο οποίο επαυξάνονται οι σκηνές των αρχείων βίντεο. Το μέγεθος της σχεδιαστικής επιφάνειας του παραθύρου είναι ίδιο με το μέγεθος των στιγμιότυπων του αρχείου βίντεο.

Στην Εικόνα 5-24 παρουσιάζεται ενδεικτικά ένα στιγμιότυπο από την Εφαρμογή 1 κατά τη διεξαγωγή της επαναληπτικής διαδικασίας, στην περίπτωση που ο χρήστης είχε επιλέξει να

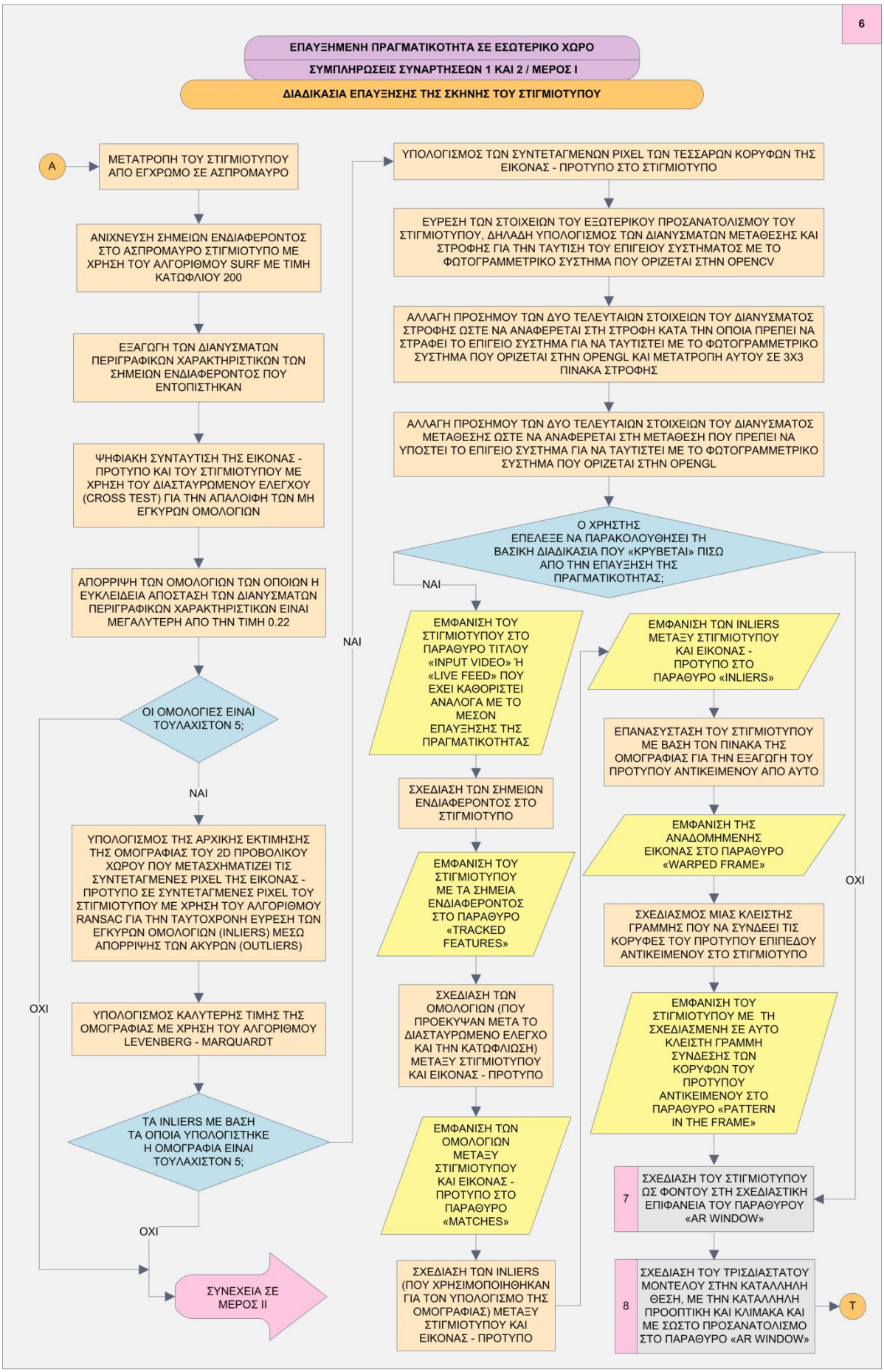
παρακολουθήσει τη βασική διαδικασία που «κρύβεται» πίσω από την επαύξηση της πραγματικότητας, για την περίπτωση επαύξησης των σκηνών ενός αρχείου βίντεο.

Εικόνα 5-23. Στιγμιότυπα από το παράθυρο «AR Window» της Εφαρμογής 1, το μέγεθος του οποίου έχει μεταβληθεί από το χρήστη (περίπτωση επαύξησης των σκηνών ενός αρχείου βίντεο).

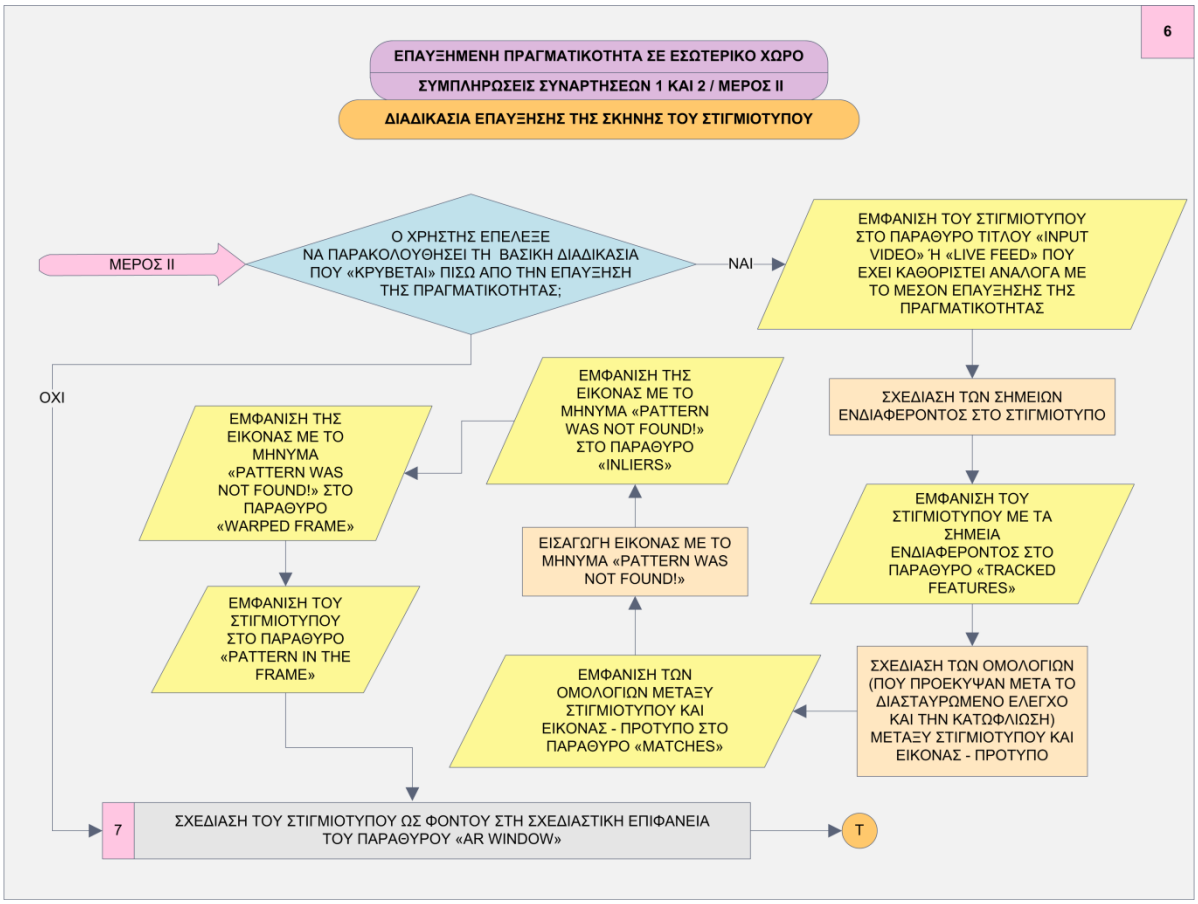
Εικόνα 5-24. Στιγμιότυπο από την Εφαρμογή 1, στην περίπτωση επαύξησης των σκηνών ενός αρχείου βίντεο.

Η διαδικασία που περιγράφηκε παραπάνω πραγματοποιείται στην περίπτωση που έχουν εντοπιστεί τουλάχιστον πέντε ομολογίες σημείων. Στην αντίθετη περίπτωση, σχεδιάζεται το στιγμιότυπο ως φόντο της σχεδιαστικής επιφάνειας του παραθύρου «AR Window», όπως ήδη περιγράφηκε, ενώ το τρισδιάστατο μοντέλο δεν σχεδιάζεται. Επίσης, στην περίπτωση αυτή μεταβάλλεται και το περιεχόμενο κάποιων παραθύρων που εμφανίζονται αν ο χρήστης το επέλεξε στο αρχικό στάδιο της εφαρμογής. Συγκεκριμένα, το παράθυρο «Inliers» δεν εικονίζει τις ομολογίες που επαληθεύουν τη σχέση ομογραφίας μεταξύ της εικόνας - πρότυπο και του στιγμιότυπου, καθώς είτε δεν υπολογίστηκε ο μετασχηματισμός ομογραφίας είτε υπολογίστηκε και οι έγκυρες αντιστοιχίσεις είναι λιγότερες από πέντε, οπότε θεωρείται ότι το πρότυπο αντικείμενο δεν εντοπίστηκε. Στην περίπτωση αυτή, εμφανίζεται στο παράθυρο «Inliers» μία εικόνα η οποία ενημερώνει το χρήστη ότι το πρότυπο αντικείμενο δεν βρέθηκε. Η ίδια εικόνα εμφανίζεται και στο παράθυρο «Warped Frame». Στο παράθυρο «Pattern in the Frame», στο οποίο απεικονίζεται το στιγμιότυπο με μία κλειστή γραμμή που ενώνει τις κορυφές του πρότυπου επίπεδου αντικειμένου, εφόσον αυτό δεν εντοπίστηκε φαίνεται μόνο το στιγμιότυπο χωρίς καμία επιπλέον γραμμή. Τα υπόλοιπα παράθυρα έχουν το ίδιο περιεχόμενο.

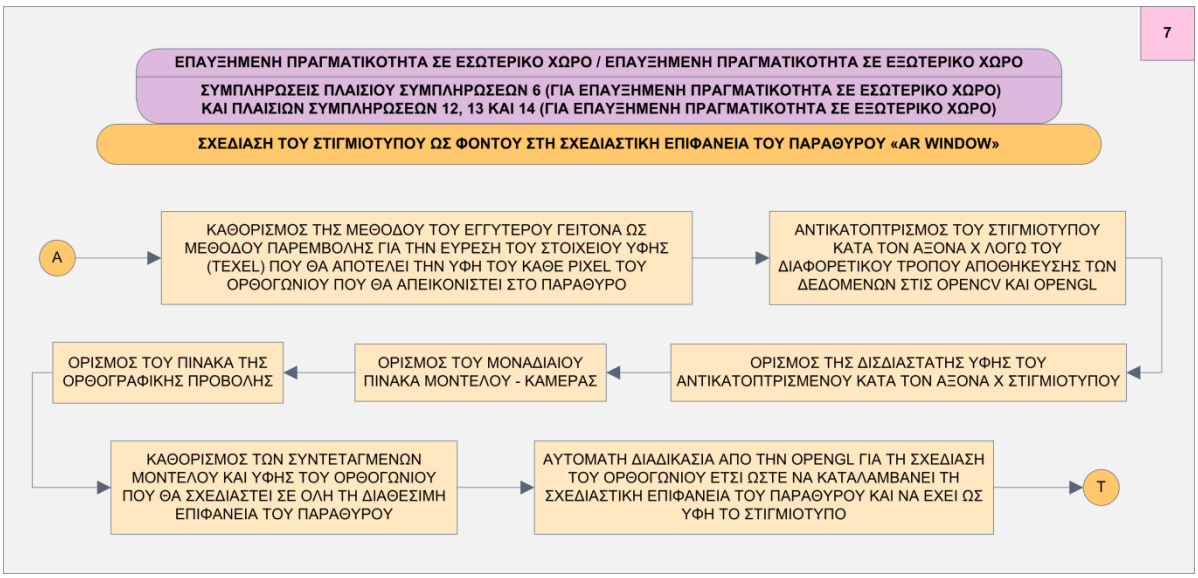
Στις Διαγραμματικές παρουσιάσεις 5-11 έως και 5-15 παρουσιάζονται τα διαγράμματα ροής της επαναληπτικής διαδικασίας που περιγράφηκε στις παραγράφους 5.4.3.1 έως και 5.4.3.4.



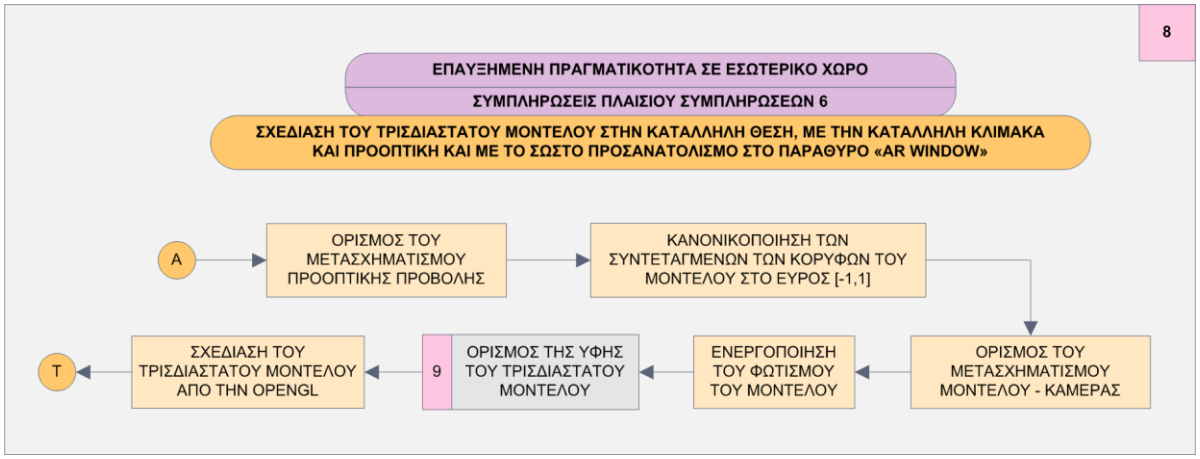
Διαγραμματική παρουσίαση 5-11. Εφαρμογή 1, πλαίσιο συμπληρώσεων 6, μέρος Ι.



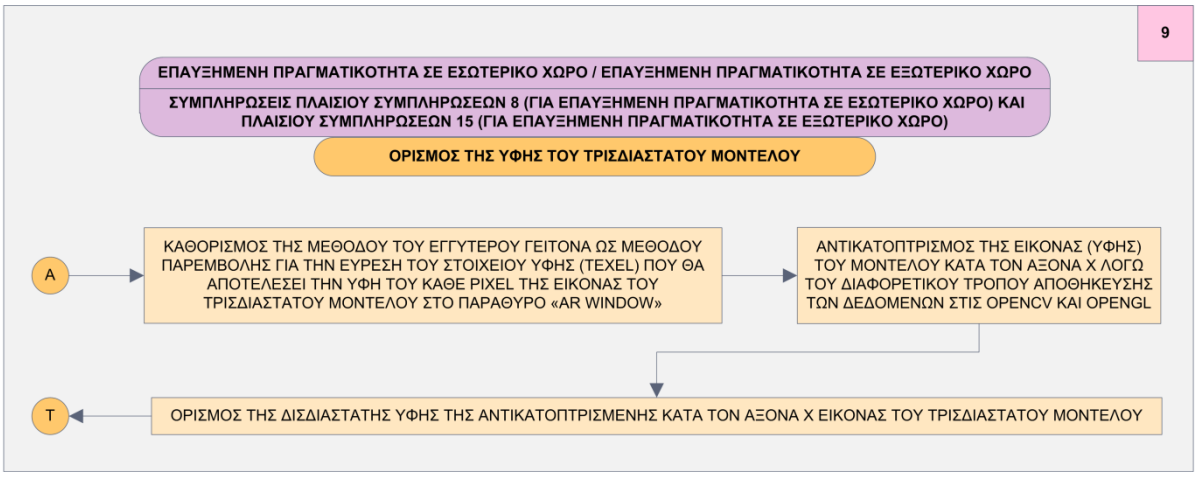
Διαγραμματική παρουσίαση 5-12. Εφαρμογή 1, πλαίσιο συμπληρώσεων 6, μέρος ΙΙ.



Διαγραμματική παρουσίαση 5-13. Εφαρμογή 1 - Εφαρμογή 2, πλαίσιο συμπληρώσεων 7.



Διαγραμματική παρουσίαση 5-14. Εφαρμογή 1, πλαίσιο συμπληρώσεων 8.



Διαγραμματική παρουσίαση 5-15. Εφαρμογή 1 - Εφαρμογή 2, πλαίσιο συμπληρώσεων 9.

5.5 Εφαρμογή 2 - Επαυξημένη Πραγματικότητα σε Εξωτερικό Χώρο

Στην εφαρμογή αυτή υποστηρίζονται δύο διαφορετικά μοντέλα, το καθένα από τα οποία επαυξάνει μία διαφορετική σκηνή. Η πρώτη σκηνή είναι η περιοχή μπροστά από τη νότια πλευρά του κτηρίου Καθαρτήρια που βρίσκεται στον πολιτιστικό πολυχώρο «Τεχνόπολις» του Δήμου Αθηναίων, στην περιοχή Γκάζι. Το μοντέλο το οποίο επαυξάνει τη νότια πλευρά του κτηρίου είναι ένα αγγείο, το οποίο τοποθετείται μπροστά από την πλευρά αυτή όταν το στιγμιότυπο του βίντεο εικονίζει το πρότυπο επίπεδο αντικείμενο. Το τελευταίο είναι το τμήμα της νότιας πλευράς μπροστά από το οποίο τοποθετείται το τρισδιάστατο αγγείο.

Η δεύτερη σκηνή είναι η βόρεια πλευρά του κτηρίου Βάιλερ – που είναι ευρέως γνωστό και ως κτήριο Μακρυγιάννη – και βρίσκεται δίπλα στο νέο Μουσείο της Ακρόπολης, βορειοανατολικά αυτού. Το τρισδιάστατο μοντέλο που την επαυξάνει είναι ένα άγαλμα. Αυτό τοποθετείται μπροστά από το μεσαίο παράθυρο της βόρειας πλευράς του ισογείου του

κτηρίου Μακρυγιάννη στην περίπτωση που στο στιγμιότυπο του βίντεο εικονίζεται το πρότυπο επίπεδο αντικείμενο, το οποίο είναι τμήμα της βόρειας πλευράς του κτηρίου, που περιλαμβάνει το συγκεκριμένο παράθυρο. Τόσο το άγαλμα όσο και το αγγείο διατέθηκαν από το Εργαστήριο Φωτογραμμετρίας της Σχολής Αγρονόμων και Τοπογράφων Μηχανικών του Εθνικού Μετσόβιου Πολυτεχνείου για την εκπόνηση της παρούσας εργασίας.

Λαμβάνοντας υπ' όψιν το γεγονός ότι ένας υπολογιστής δεν μεταφέρεται εύκολα σε εξωτερικό χώρο λόγω του μεγέθους του, η εφαρμογή αυτή δεν προγραμματίστηκε έτσι ώστε να υποστηρίζει την επαύξηση του φυσικού κόσμου σε πραγματικό χρόνο μέσω της κάμερας του ηλεκτρονικού υπολογιστή. Αντίθετα, επεξεργάζεται αρχεία βίντεο, με αποτέλεσμα να μην πληροί όλες τις προϋποθέσεις χαρακτηρισμού της ως εφαρμογής επαυξημένης πραγματικότητας, σύμφωνα με τον Azuma. Ωστόσο, στο πλαίσιο της παρούσας διπλωματικής εργασίας υιοθετείται αυτός ο χαρακτηρισμός. Ο χρήστης έχει τη δυνατότητα να δει μία προεπισκόπηση της εφαρμογής, επιλέγοντας ένα από τα δύο τρισδιάστατα μοντέλα, ή να εισαγάγει κάποιο δικό του αρχείο βίντεο. Στην περίπτωση αυτή, το αρχείο βίντεο θα πρέπει να έχει ληφθεί με κινητή συσκευή με λειτουργικό σύστημα Android έκδοσης 2.2 ή νεότερης, μέσω της εφαρμογής «Camera by Stela» που ανέπτυξα στο πλαίσιο του μαθήματος - θέματος «Εφαρμογές Διασυνδεδεμένων Ψηφιακών Συστημάτων» του 9ου εξαμήνου της Σχολής Αγρονόμων και Τοπογράφων Μηχανικών [162]. Η εφαρμογή αυτή επιτρέπει – εκτός των άλλων – τη λήψη βίντεο και την αποθήκευση σε αρχείο κειμένου (.txt) των τιμών του GPS (γεωγραφικό μήκος, γεωγραφικό πλάτος και γεωμετρικό υψόμετρο) που ορίζουν τη θέση της κινητής συσκευής στο γεωδαιτικό σύστημα WGS ΄84 κατά τη διάρκεια λήψης του βίντεο. Αν ο χρήστης επιλέξει να εισαγάγει ένα αρχείο βίντεο, πρέπει να εισαγάγει και το αντίστοιχο αρχείο κειμένου με τη θέση του κινητού κατά τη διάρκεια του βίντεο, με βάση την οποία καθορίζεται ποιο από τα δύο μοντέλα θα επαυξήσει το τοπίο που εικονίζεται στο βίντεο. Επίσης, του δίνεται η δυνατότητα, εφόσον το επιλέξει, να δει τη βασική διαδικασία που εκτελείται μέσω της βιβλιοθήκης OpenCV για την επαύξηση της πραγματικότητας και να πληκτρολογήσει τα στοιχεία του εσωτερικού προσανατολισμού της κάμερας της κινητής του συσκευής.

Η ονομασία «επαυξημένη πραγματικότητα σε εξωτερικό χώρο» ή «outdoor augmented reality» οφείλεται στο γεγονός ότι τα βίντεο που επαυξάνονται λαμβάνονται σε εξωτερικό χώρο μέσω κάποιας κινητής συσκευής. Στην ενότητα αυτή περιγράφεται η διαδικασία που ακολουθήθηκε κατά τον προγραμματισμό της εφαρμογής και παρουσιάζονται τα σχετικά αποτελέσματά της.

5.5.1 Αρχικά βήματα για τον καθορισμό των επιλογών του χρήστη

Αρχικό βήμα της εφαρμογής αποτελεί ο καθορισμός των επιλογών του χρήστη. Για το σκοπό αυτό, εμφανίζονται διαδοχικά μηνύματα και ο χρήστης καλείται να δηλώσει τις

επιλογές του μέσω πληκτρολόγησης των κατάλληλων χαρακτήρων. Παρακάτω παρουσιάζονται όλοι οι πιθανοί συνδυασμοί επιλογών του χρήστη.

- Επαύξηση των σκηνών ενός αρχείου βίντεο που εισάγεται από το χρήστη μαζί με το αρχείο με τις τιμές του GPS που το συνοδεύει (το οποίο καθορίζει αν το μοντέλο επαύξησης είναι το αγγείο ή το άγαλμα), εμφάνιση μόνο του παραθύρου απεικόνισης των επαυξημένων σκηνών, πληκτρολόγηση των στοιχείων του εσωτερικού προσανατολισμού της κάμερας με την οποία έγινε η λήψη του βίντεο από το χρήστη.
- Επαύξηση των σκηνών ενός αρχείου βίντεο που εισάγεται από το χρήστη μαζί με το αρχείο με τις τιμές του GPS που το συνοδεύει (το οποίο καθορίζει αν το μοντέλο επαύξησης είναι το αγγείο ή το άγαλμα), εμφάνιση μόνο του παραθύρου απεικόνισης των επαυξημένων σκηνών, μη εισαγωγή των στοιχείων του εσωτερικού προσανατολισμού της κάμερας με την οποία έγινε η λήψη του βίντεο από το χρήστη.
- Επαύξηση των σκηνών ενός αρχείου βίντεο που εισάγεται από το χρήστη μαζί με το αρχείο με τις τιμές του GPS που το συνοδεύει (το οποίο καθορίζει αν το μοντέλο επαύξησης είναι το αγγείο ή το άγαλμα), εμφάνιση όχι μόνο του παραθύρου επαύξησης της πραγματικότητας αλλά και παραθύρων που παρουσιάζουν τη βασική διαδικασία που διεξάγεται μέσω της βιβλιοθήκης OpenCV, πληκτρολόγηση των στοιχείων του εσωτερικού προσανατολισμού της κάμερας με την οποία έγινε η λήψη του βίντεο από το χρήστη.
- Επαύξηση των σκηνών ενός αρχείου βίντεο που εισάγεται από το χρήστη μαζί με το αρχείο με τις τιμές του GPS που το συνοδεύει (το οποίο καθορίζει αν το μοντέλο επαύξησης είναι το αγγείο ή το άγαλμα), εμφάνιση όχι μόνο του παραθύρου επαύξησης της πραγματικότητας αλλά και παραθύρων που παρουσιάζουν τη βασική διαδικασία που διεξάγεται μέσω της βιβλιοθήκης OpenCV, μη εισαγωγή των στοιχείων του εσωτερικού προσανατολισμού της κάμερας με την οποία έγινε η λήψη του βίντεο από το χρήστη.
- Επαύξηση των σκηνών ενός προεπιλεγμένου αρχείου βίντεο με το τρισδιάστατο μοντέλο του αγγείου, εμφάνιση μόνο του παραθύρου απεικόνισης των επαυξημένων σκηνών.
- Επαύξηση των σκηνών ενός προεπιλεγμένου αρχείου βίντεο με το τρισδιάστατο μοντέλο του αγγείου, εμφάνιση όχι μόνο του παραθύρου επαύξησης της πραγματικότητας αλλά και παραθύρων που παρουσιάζουν τη βασική διαδικασία που διεξάγεται μέσω της βιβλιοθήκης OpenCV.
- Επαύξηση των σκηνών ενός προεπιλεγμένου αρχείου βίντεο με το τρισδιάστατο μοντέλο του αγάλματος, εμφάνιση μόνο του παραθύρου απεικόνισης των επαυξημένων σκηνών.

- Επαύξηση των σκηνών ενός προεπιλεγμένου αρχείου βίντεο με το τρισδιάστατο μοντέλο του αγάλματος, εμφάνιση όχι μόνο του παραθύρου επαύξησης της πραγματικότητας αλλά και παραθύρων που παρουσιάζουν τη βασική διαδικασία που διεξάγεται μέσω της βιβλιοθήκης OpenCV.

Οι επιλογές του χρήστη ελέγχονται για την ορθότητά τους και σε περίπτωση λανθασμένης επιλογής ο χρήστης καλείται εκ νέου να πληκτρολογήσει το σωστό χαρακτήρα.

Στη συνέχεια αναφέρεται ο τρόπος καθορισμού του τρισδιάστατου μοντέλου αν ο χρήστης επιλέξει να επαυξήσει τις σκηνές ενός βίντεο, του οποίου το όνομα πληκτρολογεί. Στην περίπτωση αυτή καλείται να εισαγάγει – εκτός από το βίντεο – και το αρχείο με τις τιμές του GPS που το συνοδεύει. Σε κάθε γραμμή του αρχείου με τις τιμές του GPS αναγράφονται τέσσερα πεδία, τα οποία διαχωρίζονται μεταξύ τους με το χαρακτήρα «tab». Το πρώτο πεδίο είναι η ώρα στην οποία αναφέρονται οι τιμές του GPS, υπό τη μορφή ΩΩ:ΛΛ:ΔΔ, όπου τα γράμματα ΩΩ δηλώνουν την ώρα, ως ΛΛ συμβολίζονται τα λεπτά και ως ΔΔ τα δευτερόλεπτα. Το επόμενο πεδίο του αρχείου κειμένου είναι η τιμή του γεωγραφικού μήκους λ, ακολουθεί το γεωγραφικό πλάτος φ και έπεται στο τέλος το γεωμετρικό υψόμετρο h. Για παράδειγμα, μία γραμμή το αρχείου GPS μπορεί να είναι η ακόλουθη:

18:57:24 23.77942786579352 37.97571910141754 224.0

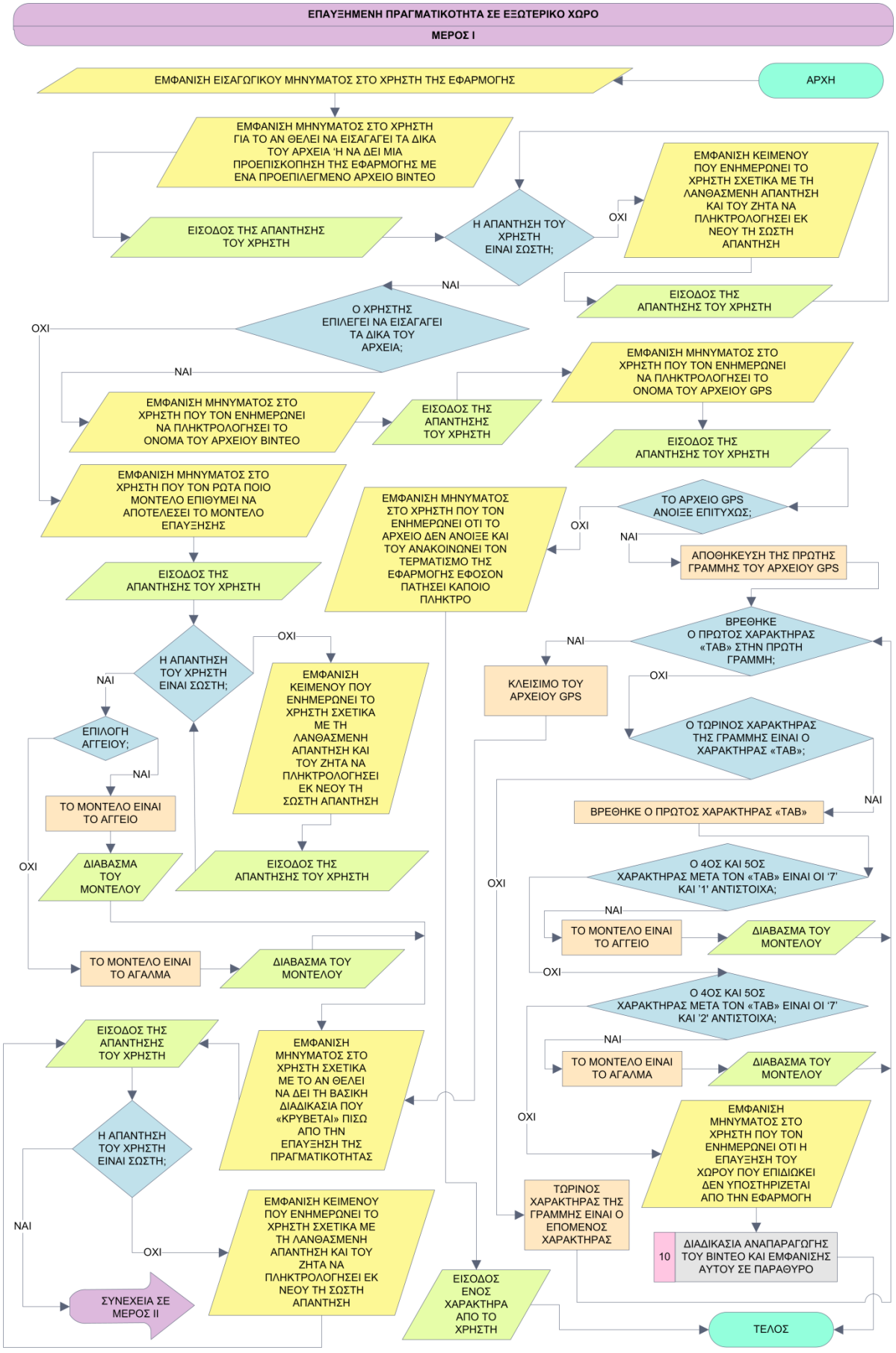
Η επιλογή του μοντέλου το οποίο θα επαυξήσει το περιεχόμενο του βίντεο που εισήγαγε ο χρήστης γίνεται με βάση τη θέση της κινητής συσκευής κατά τη διάρκεια λήψης του βίντεο. Έτσι, όταν ο χρήστης βρίσκεται στον πολυχώρο «Τεχνόπολις», οι αρχικοί χαρακτήρες που αποθηκεύονται στο αρχείο κειμένου και αναφέρονται στο γεωγραφικό μήκος είναι οι 23.71, ενώ οι αντίστοιχοι για το γεωγραφικό πλάτος είναι οι 37.97. Αντίστοιχα, αν αυτός βρίσκεται πλησίον του κτηρίου Μακρυγιάννη, οι αρχικοί χαρακτήρες που αναφέρονται στο γεωγραφικό μήκος είναι οι 23.72 και οι αντίστοιχοι για το γεωγραφικό πλάτος είναι οι 37.96. Με βάση το γεγονός ότι οι δύο περιοχές είναι μακριά η μία από την άλλη, με αποτέλεσμα να μην υπάρχει πιθανότητα να απεικονίζονται και οι δύο στο ίδιο αρχείο βίντεο, η ανάγνωση μόνο της πρώτης γραμμής του αρχείου κειμένου με τις τιμές του GPS μπορεί να καθορίσει ποια από τις δύο περιοχές εικονίζεται στο βίντεο και επομένως ποιο είναι το κατάλληλο τρισδιάστατο μοντέλο το οποίο θα πρέπει να διαβαστεί από το πρόγραμμα. Έτσι, το μοντέλο θα είναι το αγγείο στην περίπτωση που τα δεκαδικά ψηφία του γεωγραφικού μήκους είναι οι αριθμοί 7 και 1, και το άγαλμα στην περίπτωση που αυτά είναι τα ψηφία 7 και 2.

Όταν ο χρήστης πληκτρολογήσει το όνομα του αρχείου κειμένου, γίνεται προσπάθεια από το πρόγραμμα να ανοίξει το αρχείο αυτό. Αν το τελευταίο δεν ανοίξει επιτυχώς, εμφανίζεται κατάλληλο μήνυμα και η εφαρμογή τερματίζεται. Εφόσον το αρχείο με τις τιμές του GPS ανοίξει επιτυχώς, εντοπίζεται και αποθηκεύεται η πρώτη γραμμή του. Για την εύρεση του γεωγραφικού μήκους, το οποίο θα καθορίσει το μοντέλο που θα επαυξήσει το περιεχόμενο

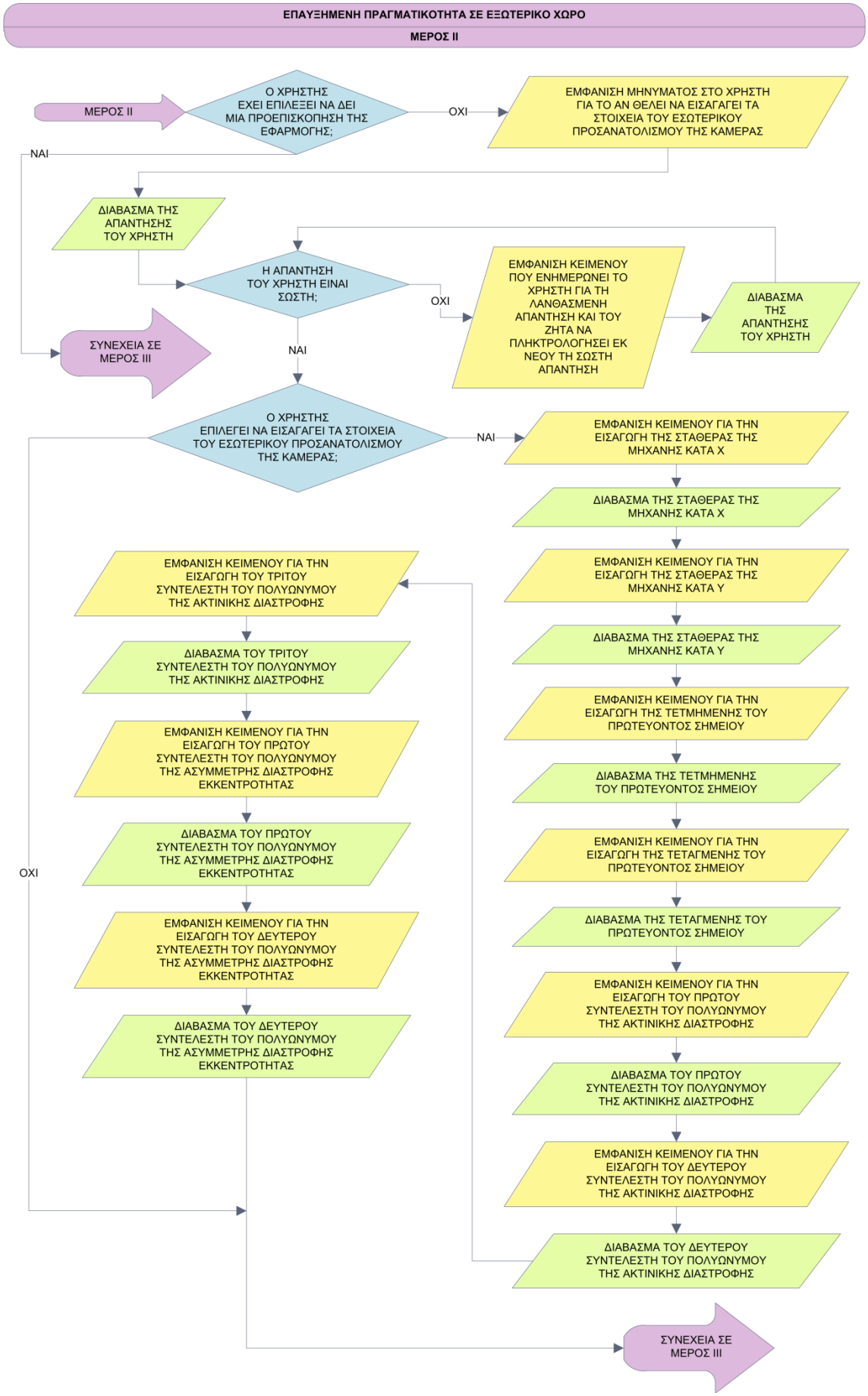
του βίντεο, αρκεί να βρεθεί ο πρώτος χαρακτήρας «tab» στην πρώτη γραμμή. Όταν αυτός βρεθεί, τότε ο τέταρτος και πέμπτος χαρακτήρας σε σχέση με το χαρακτήρα «tab» είναι εκείνοι οι οποίοι θα καθορίσουν το μοντέλο επαύξησης. Έτσι, αν ο τέταρτος χαρακτήρας είναι ο χαρακτήρας «7» και ο πέμπτος είναι ο χαρακτήρας «1», διαβάζεται μέσω της βιβλιοθήκης GLM το τρισδιάστατο μοντέλο του αγγείου (βλ. Εικόνα 5-5, κέντρο), ενώ αν ο τέταρτος και πέμπτος χαρακτήρας μετά τον «tab» είναι οι «7» και «2» αντίστοιχα, διαβάζεται μέσω της βιβλιοθήκης αυτής το τρισδιάστατο μοντέλο του αγάλματος (βλ. Εικόνα 5-5, δεξιά). Και στις δύο περιπτώσεις, αφού εντοπίστηκε ο πρώτος χαρακτήρας «tab» στην πρώτη γραμμή του αρχείου GPS, το τελευταίο κλείνει.

Αν κατά την εξέταση του περιεχομένου της πρώτης γραμμής του GPS ο τέταρτος και πέμπτος χαρακτήρας μετά τον «tab» δεν είναι οι αριθμοί «7» και «1» ούτε οι «7» και «2», ο χρήστης έχει εισαγάγει βίντεο από περιοχή που δεν υποστηρίζεται από την εφαρμογή. Στην περίπτωση αυτή, μπορεί να παρακολουθήσει το βίντεο που εισήγαγε, χωρίς κάποιου είδους επαύξηση.

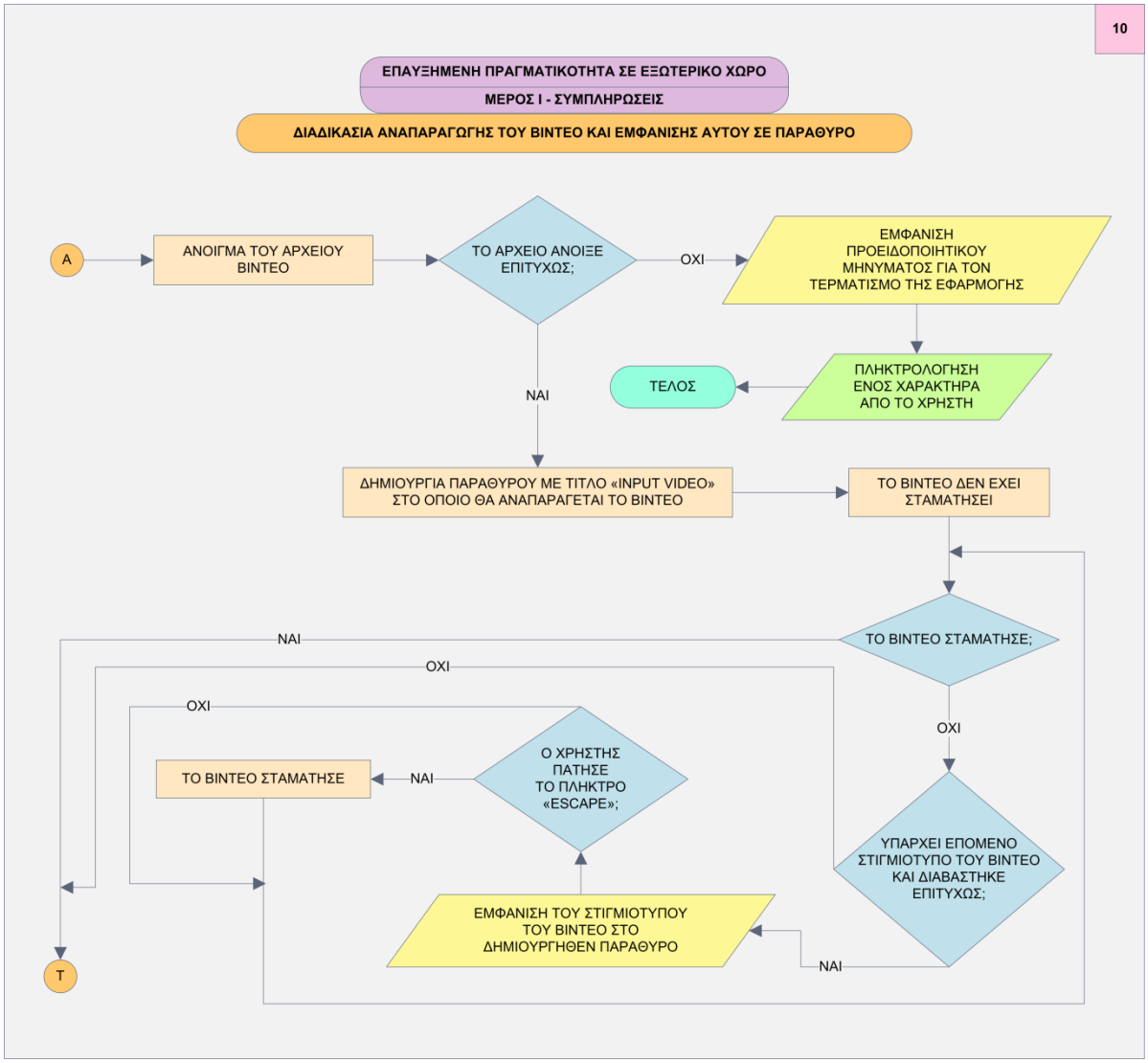
Στις Διαγραμματικές παρουσιάσεις 5-16, 5-17 και 5-18 παρουσιάζονται τα βήματα της Εφαρμογής 2 για την περάτωση της διαδικασίας που αναφέρθηκε στην υποενότητα αυτή. Τέλος, το Παράρτημα Δ αποτελεί εγχειρίδιο χρήσης της Εφαρμογής 2, στο οποίο εικονίζονται τα μηνύματα που εμφανίζονται κατά την εκτέλεσή της, αναφέρονται οδηγίες για τη σωστή εισαγωγή των επιλογών του χρήστη, ώστε να γίνει κατανοητός ο τρόπος αλληλεπίδρασης του χρήστη με την εφαρμογή, και παρουσιάζονται κάποιες λεπτομέρειες σχετικές με τον τρόπο λειτουργίας της.



Διαγραμματική παρουσίαση 5-16. Εφαρμογή 2, μέρος Ι.



Διαγραμματική παρουσίαση 5-17. Εφαρμογή 2, μέρος ΙΙ.



Διαγραμματική παρουσίαση 5-18. Εφαρμογή 2, πλαίσιο συμπληρώσεων 10.

5.5.2 Ενδιάμεσα βήματα πριν την επαναληπτική διαδικασία επαύξησης της σκηνής κάθε στιγμιότυπου

Αντίστοιχα με την Εφαρμογή 1, τον καθορισμό των επιλογών του χρήστη ακολουθεί η δημιουργία του 3x3 πίνακα της μηχανής, με τα στοιχεία του εσωτερικού προσανατολισμού της κάμερας, και του 5x1 διανύσματος με τους συντελεστές των πολυωνύμων της ακτινικής διαστροφής και της ασύμμετρης διαστροφής εκκεντρότητας. Ο πίνακας της μηχανής και το διάνυσμα με τους συντελεστές της διαστροφής δημιουργούνται από:

- τα στοιχεία που δίνει ο χρήστης, εφόσον επέλεξε να πληκτρολογήσει τα στοιχεία του εσωτερικού προσανατολισμού της κάμερας, ή
- τα στοιχεία του εσωτερικού προσανατολισμού της κάμερας του κινητού τηλεφώνου Samsung, όπως προέκυψαν από τη βαθμονόμηση που πραγματοποιήθηκε, εφόσον δεν έγινε εισαγωγή των στοιχείων του εσωτερικού προσανατολισμού από το χρήστη.

Στη συνέχεια, επιχειρείται η αποθήκευση των διαστάσεων του βίντεο, για το μετέπειτα καθορισμό του μεγέθους του παραθύρου επαύξησης της πραγματικότητας. Για το λόγο αυτό, διαβάζεται το πρώτο στιγμιότυπο του αρχείου βίντεο, το οποίο είναι:

- αυτό που εισήγαγε ο χρήστης, μέσω πληκτρολόγησης του ονόματός του, ή
- το προεπιλεγμένο βίντεο που εικονίζει τη βόρεια πλευρά του κτηρίου Μακρυγιάννη, αν ο χρήστης επέλεξε να παρακολουθήσει μία προεπισκόπηση της εφαρμογής και το άγαλμα ως μοντέλο επαύξησης, ή
- το προεπιλεγμένο βίντεο που εικονίζει τη νότια πλευρά του κτηρίου Καθαρτήρια, αν ο χρήστης επέλεξε να παρακολουθήσει μία προεπισκόπηση της εφαρμογής και το αγγείο ως μοντέλο επαύξησης.

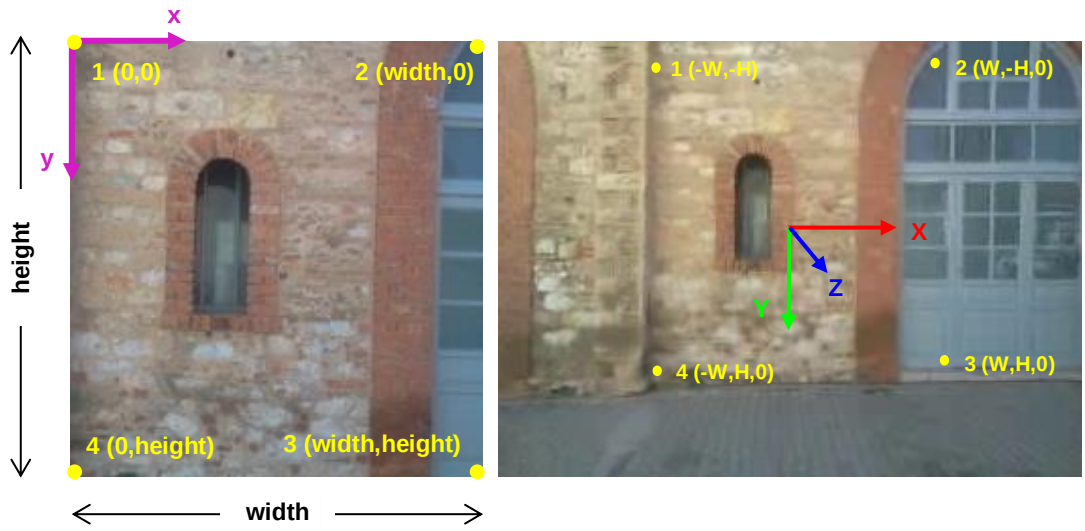
Αν το αρχείο βίντεο άνοιξε επιτυχώς, διαβάζεται το πρώτο στιγμιότυπο και υπολογίζονται οι διαστάσεις του. Αντίθετα, αν το βίντεο δεν άνοιξε επιτυχώς (συνήθως λόγω λανθασμένης πληκτρολόγησης του ονόματος από το χρήστη) εμφανίζεται κατάλληλο προειδοποιητικό μήνυμα για τον τερματισμό της εφαρμογής (βλ. Παράρτημα Δ).

Στην περίπτωση που το αρχείο βίντεο άνοιξε επιτυχώς, ακολουθούν τα βήματα που απαιτούνται για την αρχικοποίηση και δημιουργία ενός παραθύρου για την επαύξηση της σκηνής μέσω της βιβλιοθήκης freeglut, όπως περιγράφηκαν στην υποενότητα 5.4.2, στο πλαίσιο της Εφαρμογής 1. Τα βήματα αυτά είναι η αρχικοποίηση της βιβλιοθήκης freeglut, ο ορισμός διαφόρων παραμέτρων για την απεικόνιση της επαυξημένης σκηνής στο παράθυρο που θα δημιουργηθεί (χρήση χρωματικού μοντέλου RGB και διπλού «buffer), ο καθορισμός του μεγέθους του παραθύρου και της θέσης στην οποία θα εμφανίζεται στην οθόνη και τελικά η δημιουργία του παραθύρου με τίτλο «AR Window». Στη συνέχεια, αρχικοποιείται η βιβλιοθήκη GLEW. Στην περίπτωση που η αρχικοποίηση δεν είναι επιτυχής, εμφανίζεται κατάλληλο μήνυμα που ειδοποιεί για τον επερχόμενο τερματισμό της εφαρμογής (βλ. Παράρτημα Δ).

Ακολούθως, καθορίζεται η συνάρτηση η οποία θα καλείται από τη βιβλιοθήκη freeglut κάθε φορά που απαιτείται ο σχεδιασμός της σκηνής και η συνάρτηση που θα εκτελείται συνεχώς, όταν δεν λαμβάνονται «γεγονότα» προς επεξεργασία. Ανάλογα με την Εφαρμογή 1, καθορίζεται η κλήση της ίδιας συνάρτησης, όταν απαιτείται επανασχεδιασμός της σκηνής αλλά και συνεχώς όταν δεν λαμβάνονται «γεγονότα» προς επεξεργασία. Πρόκειται για τη συνάρτηση η οποία επαυξάνει την πραγματικότητα και η οποία θα καλείται συνεχώς, για κάθε νέο στιγμιότυπο του βίντεο. Ωστόσο, η τελευταία διαφέρει ανάλογα με το τρισδιάστατο μοντέλο επαύξησης και έτσι καθορίζεται διαφορετική συνάρτηση αν το τρισδιάστατο μοντέλο είναι το αγγείο ή το άγαλμα αντίστοιχα. Οι συναρτήσεις αυτές παρουσιάζονται αναλυτικά στην υποενότητα 5.5.3.

Επίσης, καθορίζεται η συνάρτηση η οποία θα καλείται κάθε φορά που ο χρήστης πατά κάποιο πλήκτρο, σύμφωνα με την οποία αν ο χρήστης πατήσει το πλήκτρο «Escape» η εφαρμογή θα τερματιστεί. Επιπλέον, καθορίζεται η συνάρτηση η οποία θα καλείται κάθε φορά που ο χρήστης μεταβάλλει το μέγεθος του παραθύρου «AR Window», η οποία μεταβάλλει και το μέγεθος της επαυξημένης (ή μη) σκηνής με τέτοιο τρόπο ώστε να διατηρείται ο λόγος των διαστάσεων της σκηνής σταθερός και ίσος με τον αρχικό λόγο αυτών. Η συνάρτηση αυτή παρουσιάστηκε στην παράγραφο 5.4.2 (βλ. Διαγραμματική παρουσίαση 5-7).

Ανάλογα με το τρισδιάστατο μοντέλο (αγγείο ή άγαλμα), εισάγεται στο πρόγραμμα η κατάλληλη εικόνα - πρότυπο και υπολογίζονται οι διαστάσεις της, δηλαδή το πλάτος και το ύψος της σε pixels. Στη συνέχεια, δημιουργείται το διάνυσμα με τις συντεταγμένες των τεσσάρων κορυφών του επίπεδου πρότυπου αντικειμένου στο σύστημα συντεταγμένων της εικόνας - πρότυπο και το διάνυσμα με τις συντεταγμένες των τεσσάρων κορυφών του επίπεδου πρότυπου αντικειμένου στο επίγειο σύστημα συντεταγμένων. Το τελευταίο ορίστηκε όπως περιγράφηκε στην υποενότητα 5.4.2 για την Εφαρμογή 1. Στις Εικόνες 5-25 και 5-26 φαίνονται τα συστήματα συντεταγμένων εικόνας και χώρου (επίγειο) και οι συντεταγμένες των τεσσάρων κορυφών του πρότυπου αντικειμένου στα δύο αυτά συστήματα, για τις δύο περιπτώσεις των μοντέλων επαύξησης.



Εικόνα 5-25. Σύστημα συντεταγμένων εικόνας [αριστερά] και επίγειο [δεξιά] και συντεταγμένες των τεσσάρων κορυφών του πρότυπου αντικειμένου (τμήμα της νότιας πλευράς του κτηρίου Καθαρτήρια) στα δύο συστήματα, για την περίπτωση που το μοντέλο επαύξησης είναι το αγγείο (Εφαρμογή 2).

Όσον αφορά στα πρότυπα αντικείμενα που ορίστηκαν στην Εφαρμογή 2, αυτά παρεκκλίνουν από την επιπεδότητα, καθώς έχουν μικρό ανάγλυφο, που δεν ξεπερνά τα 20cm. Η απόσταση αυτή, συγκρινόμενη με την απόσταση λήψης της εικόνας - πρότυπο και των στιγμιότυπων του βίντεο, δεν είναι αμελητέα. Ωστόσο, έγινε επιλογή των συγκεκριμένων κτηρίων για τη διερεύνηση των τελικών αποτελεσμάτων, στην περίπτωση που το πρότυπο αντικείμενο παρουσιάζει μικρό ανάγλυφο. Τελικά, το αποτέλεσμα επαύξησης των σκηνών

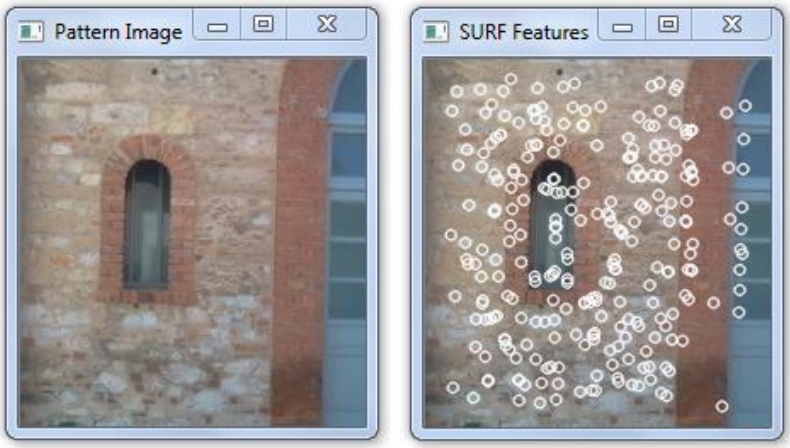
του βίντεο των περιοχών αυτών ήταν ικανοποιητικό και ανάλογο με την περίπτωση του επίπεδου πρότυπου αντικειμένου της ορθοφωτογραφίας.

Εικόνα 5-26. Σύστημα συντεταγμένων εικόνας [αριστερά] και επίγειο [δεξιά] και συντεταγμένες των τεσσάρων γωνιών του πρότυπου αντικειμένου (τμήμα της βόρειας πλευράς του κτηρίου Μακρυγιάννη) στα δύο συστήματα, για την περίπτωση που το μοντέλο επαύξησης είναι το άγαλμα (Εφαρμογή 2).

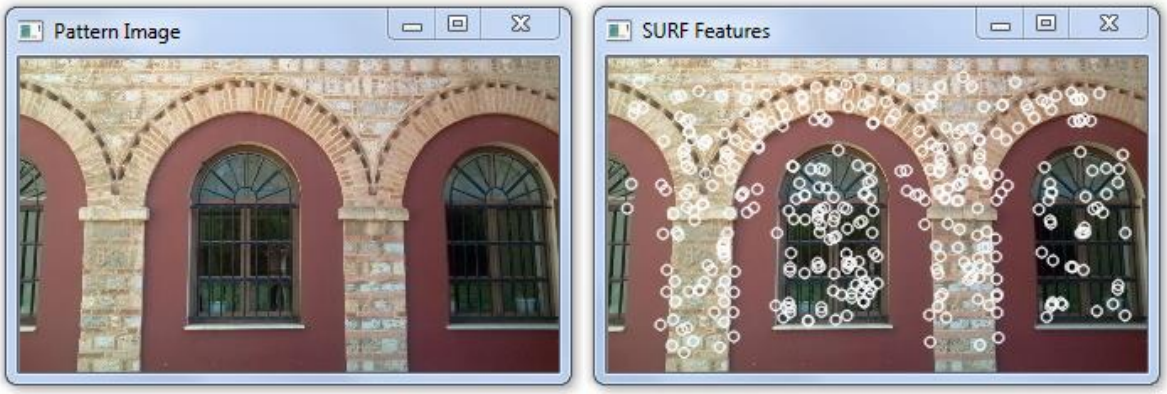
Στην περίπτωση που ο χρήστης είχε επιλέξει να παρακολουθήσει τη βασική διαδικασία που διεξάγεται προκειμένου να επιτευχθεί η επαύξηση της πραγματικότητας, εμφανίζεται η κατάλληλη εικόνα - πρότυπο σε ένα παράθυρο που δημιουργείται με χρήση της OpenCV με τίτλο «Pattern Image». Όταν ο χρήστης πατήσει οποιοδήποτε πλήκτρο, αυτό κλείνει και το πρόγραμμα συνεχίζει με το επόμενο βήμα, που είναι η ανίχνευση των σημείων ενδιαφέροντος με χρήση του αλγορίθμου SURF στην εικόνα - πρότυπο, μετά από μετατροπή της σε ασπρόμαυρη. Αποτέλεσμα της διαδικασίας αυτής, όπως περιγράφηκε στην προηγούμενη υποενότητα, είναι η δομή «KeyPoint» με τις ιδιότητες κάθε σημείου ενδιαφέροντος. Στην περίπτωση που ο χρήστης είχε επιλέξει να δει τη βασική διαδικασία που απαιτείται να διεξαχθεί προκειμένου να επιτευχθεί η επαύξηση της πραγματικότητας, δημιουργείται με χρήση της βιβλιοθήκης OpenCV ένα παράθυρο με τίτλο «SURF Features», στο οποίο εμφανίζεται η έγχρωμη εικόνα - πρότυπο με τα σημεία ενδιαφέροντος τα οποία εντοπίστηκαν. Το παράθυρο αυτό παραμένει ανοικτό όταν ο χρήστης δεν πατά κάποιο πλήκτρο, ενώ κλείνει και συνεχίζει η διαδικασία επαύξησης της πραγματικότητας όταν πατήσει οποιοδήποτε πλήκτρο. Στις Εικόνες 5-27 και 5-28 παρουσιάζονται τα παράθυρα με την εικόνα - πρότυπο και με τα εντοπισμένα χαρακτηριστικά σημεία σε αυτή, για τις δύο περιπτώσεις της Εφαρμογής 2 (κτήριο Καθαρτήρια και κτήριο Μακρυγιάννη αντίστοιχα).

Μετά την ανίχνευση των σημείων πραγματοποιείται η περιγραφή τους με χρήση του SURF, διαδικασία η οποία υπολογίζει έναν πίνακα v γραμμών και 64 στηλών, όπου v ο αριθμός των σημείων ενδιαφέροντος, ενώ ο αριθμός 64 είναι η διάσταση του διανύσματος περιγραφικών χαρακτηριστικών κάθε σημείου. Η διαδικασία αυτή διεξάγεται με δεδομένη τη δομή «KeyPoint» και την ασπρόμαυρη εικόνα. Οι παράμετροι που καθορίστηκαν ως δεδομένα εισόδου για τον αλγόριθμο SURF είναι οι ίδιες με εκείνες που ορίστηκαν και στην Εφαρμογή 1, εκτός από το κατώφλι, το οποίο μετά από δοκιμές τέθηκε 100 για την εικόνα -

πρότυπο του τμήματος της νότιας πλευράς του κτηρίου Καθαρτήρια στον πολυχώρο «Τεχνόπολις» και 400 για το τμήμα της βόρειας πλευράς του κτηρίου Μακρυγιάννη.



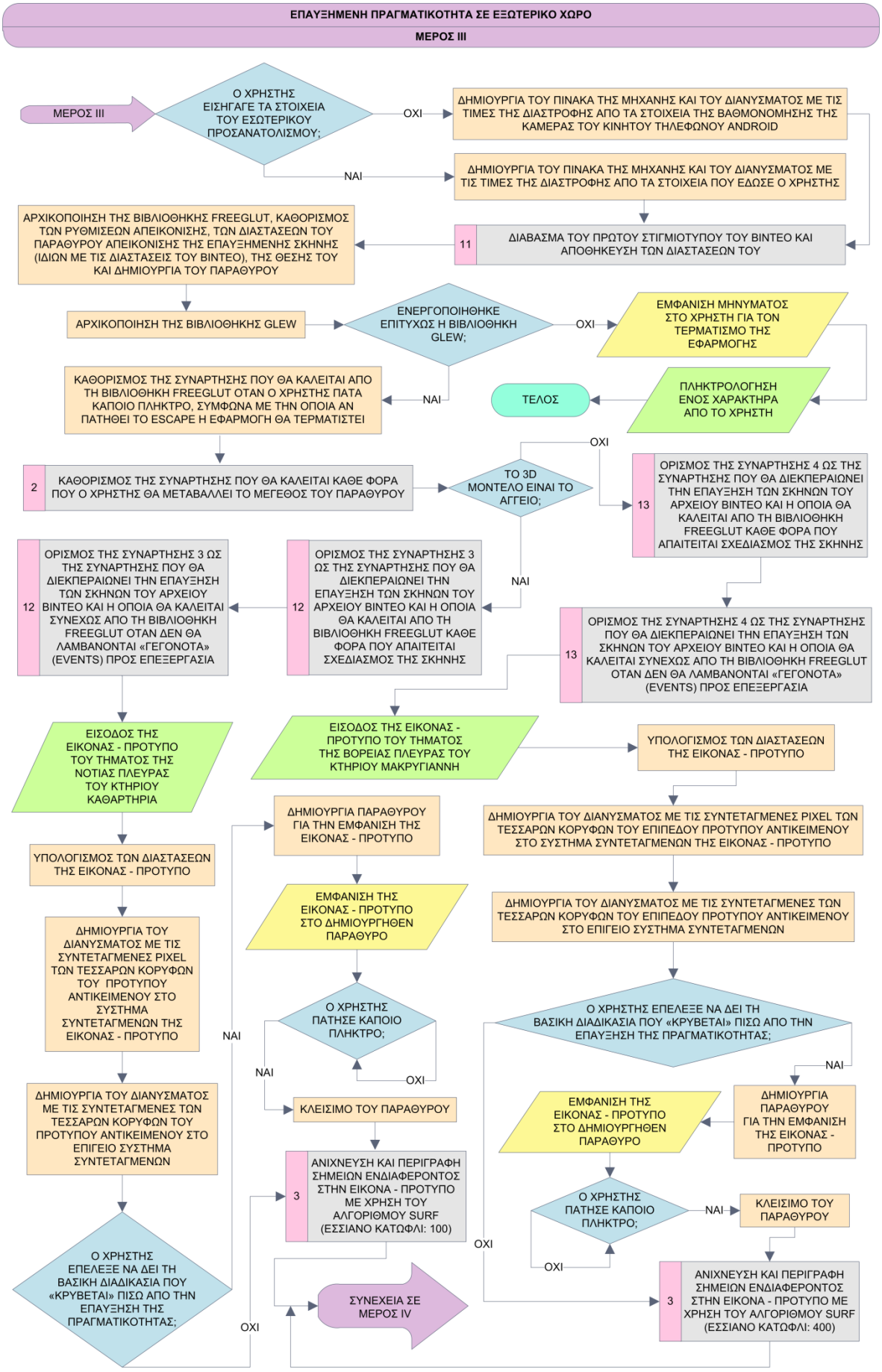
Εικόνα 5-27. Εικόνα - πρότυπο του τμήματος της νότιας πλευράς του κτηρίου Καθαρτήρια [αριστερά] και εντοπισμένα σημεία ενδιαφέροντος με χρήση του αλγορίθμου SURF [δεξιά] για την Εφαρμογή 2.



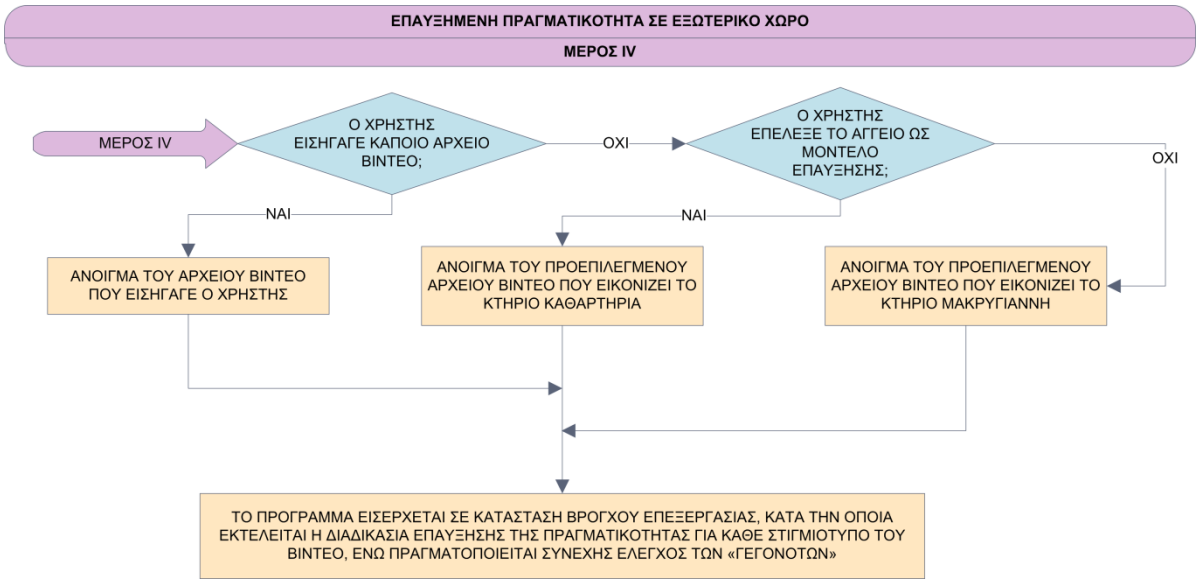
Εικόνα 5-28. Εικόνα - πρότυπο του τμήματος του κτηρίου Μακρυγιάννη [αριστερά] και εντοπισμένα σημεία ενδιαφέροντος με χρήση του αλγορίθμου SURF [δεξιά] για την Εφαρμογή 2.

Ακολούθως, αν ο χρήστης είχε εισαγάγει κάποιο αρχείο βίντεο, αυτό ανοίγει. Διαφορετικά, αν είχε επιλέξει να παρακολουθήσει μία προεπισκόπηση της εφαρμογής με το αγγείο ως μοντέλο επαύξησης, ανοίγει το προεπιλεγμένο από την εφαρμογή αρχείο βίντεο από τον πολυχώρο «Τεχνόπολις» που εικονίζει το κτήριο Καθαρτήρια. Αντίθετα, αν είχε επιλέξει το άγαλμα ως μοντέλο επαύξησης, ανοίγει το προεπιλεγμένο αρχείο βίντεο που εικονίζει το κτήριο Μακρυγιάννη. Και στις τρεις περιπτώσεις, μετά από τη διαδικασία αυτή εκτελείται η επαναληπτική διαδικασία επαύξησης της σκηνής κάθε στιγμιότυπου, η οποία ταυτόχρονα εξετάζει την αλληλεπίδραση του χρήστη με τις μονάδες εισόδου, λαμβάνοντας τα αντίστοιχα «γεγονότα» και εκτελώντας τις κατάλληλες ενέργειες. Αυτές είναι το κλείσιμο της εφαρμογής όταν πατηθεί το πλήκτρο «Escape», και η κατάλληλη μεταβολή του μεγέθους της σκηνής που εικονίζεται στο παράθυρο αυτό, όταν μεταβάλλεται το μέγεθος του παραθύρου.

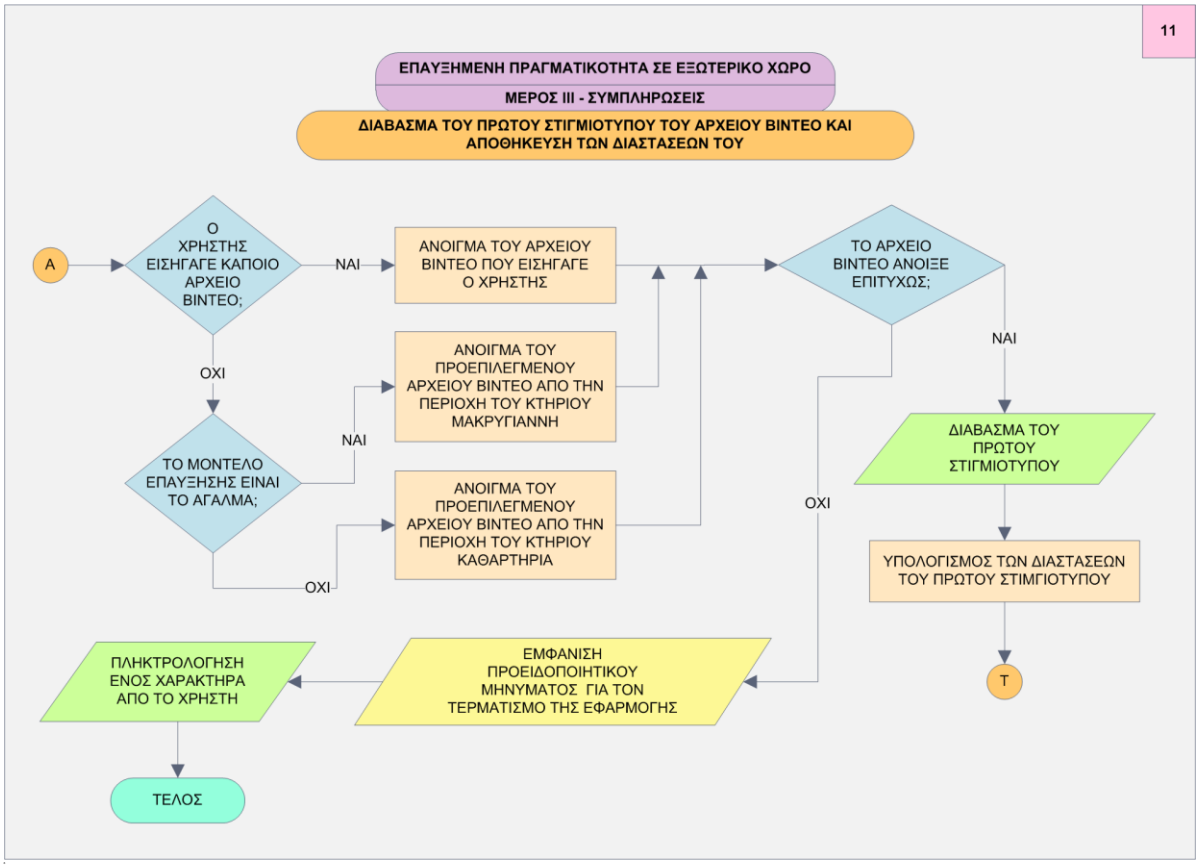
Στις Διαγραμματικές παρουσιάσεις 5-19, 5-20 και 5-21 παρουσιάζεται το διάγραμμα ροής της διαδικασίας που περιγράφηκε στην παρούσα υποενότητα.



Διαγραμματική παρουσίαση 5-19. Εφαρμογή 2, μέρος ΙΙΙ.



Διαγραμματική παρουσίαση 5-20. Εφαρμογή 2, μέρος IV.



Διαγραμματική παρουσίαση 5-21. Εφαρμογή 2, πλαίσιο συμπληρώσεων 11.

5.5.3 Επαναληπτική διαδικασία επαύξησης της σκηνής κάθε στιγμιότυπου

Όπως αναφέρθηκε στην υποενότητα 5.5.2, ανάλογα με το μοντέλο επαύξησης της πραγματικότητας ορίζεται διαφορετική συνάρτηση που θα καλείται επαναληπτικά για όλα τα

στιγμιότυπα του βίντεο. Οι διαφορές των συναρτήσεων αυτών εντοπίζονται στο κατώφλι για την ανίχνευση των σημείων ενδιαφέροντος στο εκάστοτε στιγμιότυπο από τον αλγόριθμο SURF και στη σχεδίαση του τρισδιάστατου μοντέλου του αγγείου και του αγάλματος. Στις παραγράφους που ακολουθούν περιγράφονται τα κοινά τμήματα των δύο συναρτήσεων, και στα σημεία στα οποία διαφέρουν παρουσιάζεται η διαφορετική αντιμετώπιση της επαύξησης της πραγματικότητας όταν το τρισδιάστατο μοντέλο είναι το αγγείο ή το άγαλμα αντίστοιχα.

5.5.3.1 Υπολογισμός του εξωτερικού προσανατολισμού του στιγμιότυπου

Αρχικό βήμα της επαναληπτικής διαδικασίας είναι το διάβασμα του εκάστοτε στιγμιότυπου. Αν αυτό δεν διαβάστηκε, επειδή πιθανότατα το βίντεο τελείωσε, η εφαρμογή τερματίζεται. Διαφορετικά, ακολουθεί η επεξεργασία του στιγμιότυπου για την επαύξηση της σκηνής. Αρχικό βήμα της διαδικασίας αυτής είναι η μετατροπή του στιγμιότυπου σε ασπρόμαυρο, ώστε να διεξαχθεί η ανίχνευση και περιγραφή των σημείων ενδιαφέροντος από τον αλγόριθμο SURF. Το κατώφλι που ορίζεται είναι το ίδιο με αυτό που ορίστηκε για την ανίχνευση των σημείων ενδιαφέροντος στην εικόνα - πρότυπο και στην περίπτωση που το μοντέλο επαύξησης είναι το αγγείο τίθεται 100, ενώ αν πρόκειται για το άγαλμα τίθεται 400.

Επόμενο βήμα αποτελεί η ψηφιακή συνταύτιση της εικόνας - πρότυπο και του στιγμιότυπου, με δεδομένους τους πίνακες με τα διανύσματα περιγραφικών χαρακτηριστικών των σημείων ενδιαφέροντος της εικόνας - πρότυπο και του στιγμιότυπου. Στην εφαρμογή αυτή επιλέγεται να διεξαχθεί η συνταύτιση με χρήση του ελέγχου του λόγου των αποστάσεων των εγγύτερων γειτόνων (ratio test) ως μεθόδου απαλοιφής των μη έγκυρων ομολογιών (βλ. παράγραφο 2.3.3.1). Για το λόγο αυτό, ορίζεται κατά τη διαδικασία της συνταύτισης για κάθε σημείο ενδιαφέροντος του στιγμιότυπου να επιστρέφονται τα δύο κοντινότερα σε αυτό σημεία ενδιαφέροντος της εικόνας - πρότυπο, με βάση την ευκλείδεια απόσταση των διανυσμάτων περιγραφικών χαρακτηριστικών τους. Από τη διαδικασία αυτή, η οποία υλοποιείται μέσω συνάρτησης της βιβλιοθήκης OpenCV, προκύπτει μία δομή «DMatch» που περιλαμβάνει για κάθε σημείο ενδιαφέροντος του στιγμιότυπου τέσσερα χαρακτηριστικά για τον πλησιέστερο γείτονά του και άλλα τέσσερα χαρακτηριστικά για το δεύτερο πλησιέστερο γείτονά του. Τα χαρακτηριστικά αυτά είναι:

- ένας δείκτης που φανερώνει τη θέση του σημείου στη δομή «Keypoint» του στιγμιότυπου
- ένας δείκτης για τη θέση του σημείου στην εικόνα - πρότυπο
- ένας δείκτης που χαρακτηρίζει την εικόνα πρότυπο (μη χρήσιμος για την εφαρμογή)
- η ευκλείδεια απόσταση των διανυσμάτων περιγραφικών χαρακτηριστικών των σημείων.

Η δομή αυτή θα μπορούσε να γίνει κατανοητή ως ένας πίνακας n γραμμών, όπου n ο αριθμός των σημείων ενδιαφέροντος του στιγμιότυπου, και 2 στηλών, εκ των οποίων η

πρώτη στήλη περιλαμβάνει τα τέσσερα προαναφερθέντα χαρακτηριστικά που αφορούν στον πλησιέστερο γείτονα και η δεύτερη περιλαμβάνει τα τέσσερα προαναφερθέντα χαρακτηριστικά του δεύτερου πλησιέστερου γείτονα.

Ο λόγος - κατώφλι βρέθηκε μέσω διαδοχικών δοκιμών ίσος με 1/1.5. Ο λόγος αυτός δείχνει ότι η ευκλείδεια απόσταση των διανυσμάτων περιγραφικών χαρακτηριστικών του σημείου ενδιαφέροντος του στιγμιότυπου και του δεύτερου εγγύτερου γείτονά του στην εικόνα - πρότυπο πρέπει να είναι τουλάχιστον μιάμιση φορά μεγαλύτερη από την απόσταση των διανυσμάτων περιγραφικών χαρακτηριστικών του σημείου ενδιαφέροντος του στιγμιότυπου και του πρώτου εγγύτερου γείτονά του στην εικόνα - πρότυπο, ώστε ο τελευταίος να θεωρηθεί ομόλογος του σημείου ενδιαφέροντος του στιγμιότυπου. Έτσι, αν ο λόγος των αποστάσεων των εγγύτερων γειτόνων (NNDR), ο οποίος δίνεται από τη σχέση (2.18), είναι μικρότερος από το λόγο 1/1.5, τότε θεωρείται ότι το ομόλογο του εικονοσημείου του στιγμιότυπου είναι ο εγγύτερος γείτονάς του στην εικόνα πρότυπο. Διαφορετικά θεωρείται ότι για το συγκεκριμένο εικονοσημείο του στιγμιότυπου δεν έχει εντοπιστεί το ομόλογό του στην εικόνα - πρότυπο.

Οι ομολογίες που προκύπτουν από την παραπάνω διαδικασία περιέχουν αρκετά χονδροειδή λάθη, τα οποία οφείλονται στο γεγονός ότι δεν έχει οριστεί κάποιο κατώφλι για την αποδοχή ή την απόρριψή τους και ο μόνος έλεγχος ο οποίος γίνεται είναι το ratio test. Για το λόγο αυτό, θεωρείται ότι το ζεύγος σημείων είναι ομόλογο εφόσον η ευκλείδεια απόσταση των διανυσμάτων περιγραφικών χαρακτηριστικών τους είναι μικρότερη της τιμής 0.20, η οποία βρέθηκε μετά από δοκιμές. Με βάση αυτή τη θεώρηση απορρίπτονται οι μη έγκυρες ομολογίες.

Επόμενο βήμα αποτελεί ο υπολογισμός του δισδιάστατου προβολικού μετασχηματισμού μεταξύ της εικόνας - πρότυπο και του στιγμιότυπου του βίντεο, με δεδομένες τις ευρεθείσες ομολογίες σημείων. Η εύρεση μίας αρχικής εκτίμησης του πίνακα του μετασχηματισμού αυτού γίνεται μέσω του αλγορίθμου RANSAC, ο οποίος ταυτόχρονα απορρίπτει τις εναπομένουσες μη έγκυρες ομολογίες, ενώ η λύση αυτή βελτιώνεται μέσω του αλγορίθμου βελτιστοποίησης Levenberg-Marquardt, με χρήση όλων των έγκυρων ομολογιών που εντοπίστηκαν από το RANSAC. Η διαδικασία αυτή πραγματοποιείται στην περίπτωση που οι διατιθέμενες ομολογίες σημείων είναι τουλάχιστον πέντε, για το λόγο ο οποίος αναφέρθηκε στην παράγραφο 5.4.3.1. Στην περίπτωση εύρεσης τουλάχιστον πέντε ομολογιών, υπολογίζεται ο 3x3 πίνακας της ομογραφίας που συνδέει τις συντεταγμένες της εικόνας - πρότυπο με τις συντεταγμένες του πρότυπου επίπεδου αντικειμένου στο στιγμιότυπο, όπως περιγράφεται στην παράγραφο 5.4.3.1.

Εφόσον ο πίνακας του προβολικού μετασχηματισμού υπολογίστηκε με χρήση τουλάχιστον πέντε ομολογιών, θεωρείται ότι το πρότυπο αντικείμενο έχει εντοπιστεί στο στιγμιότυπο και μπορεί να πραγματοποιηθεί η διαδικασία επαύξησής του. Στην περίπτωση αυτή,

υπολογίζονται οι συντεταγμένες των τεσσάρων κορυφών της εικόνας - πρότυπο στο στιγμιότυπο, με βάση τον πίνακα της ομογραφίας, ο οποίος μετασχηματίζει τις συντεταγμένες της εικόνας - πρότυπο σε συντεταγμένες στιγμιότυπου, όπως περιγράφηκε στην παράγραφο 5.4.3.1 (βλ. σχέσεις (5.8) και (5.9)).

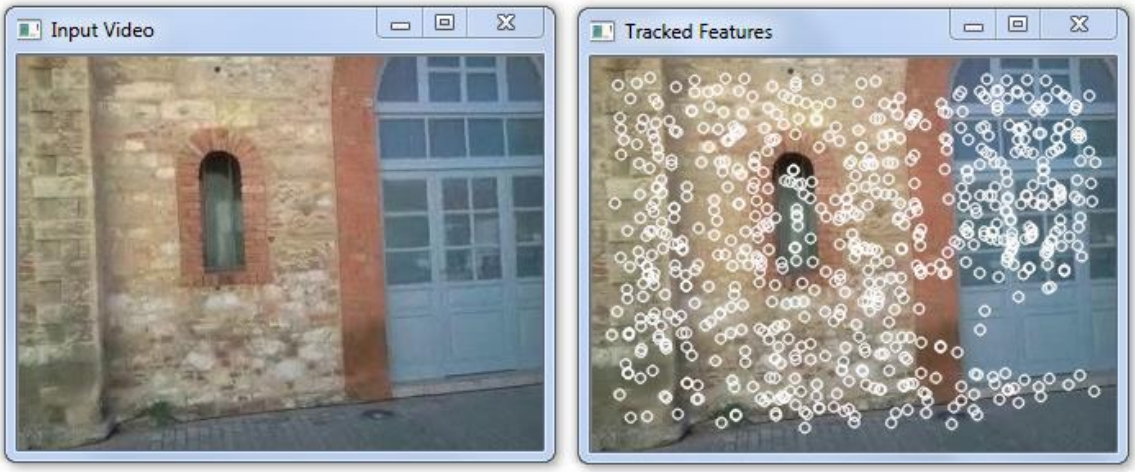
Στη συνέχεια, υπολογίζεται ο εξωτερικός προσανατολισμός του συγκεκριμένου στιγμιότυπου, με δεδομένα τις συντεταγμένες pixel των τεσσάρων κορυφών του πρότυπου επίπεδου αντικειμένου στο στιγμιότυπο, τις συντεταγμένες αυτών στο επίγειο σύστημα αναφοράς και τα στοιχεία του εσωτερικού προσανατολισμού. Ο υπολογισμός των στοιχείων του εξωτερικού προσανατολισμού πραγματοποιείται μέσω συνάρτησης της βιβλιοθήκης OpenCV, η οποία υπολογίζει το διάνυσμα μετάθεσης **t** και το διάνυσμα στροφής **r,** τα οποία ορίζουν τη μετάθεση και τη στροφή που πρέπει να υποστεί το επίγειο σύστημα συντεταγμένων για να ταυτιστεί με το φωτογραμμετρικό, όπως ορίζεται στη βιβλιοθήκη OpenCV (βλ. Εικόνα 4-2, δεξιά). Όπως αναφέρθηκε στην παράγραφο 5.4.3.1, προκειμένου τα αποτελέσματα αυτά να χρησιμοποιηθούν για την κατασκευή του πίνακα του μετασχηματισμού μοντέλου - κάμερας, το διάνυσμα στροφής και το διάνυσμα μετάθεσης μετασχηματίζονται σύμφωνα με τις σχέσεις (5.10) και (5.11), ώστε να αναφέρονται στον ορισμό του φωτογραμμετρικού συστήματος της OpenGL. Το διάνυσμα στροφής στη συνέχεια μετατρέπεται σε 3x3 πίνακα στροφής, σύμφωνα με τις σχέσεις (3.111) και (3.112).

5.5.3.2 Εμφάνιση των κατάλληλων παραθύρων μέσω της OpenCV

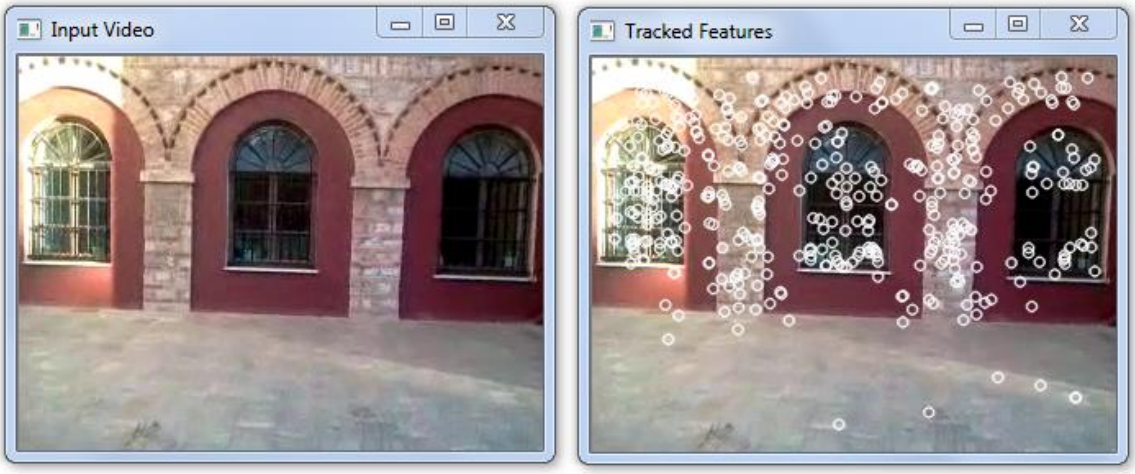
Προτού προχωρήσει η διαδικασία επαύξησης της πραγματικότητας με τη βοήθεια της OpenGL, προγραμματίζεται η εμφάνιση κάποιων παραθύρων μέσω της OpenCV, με περιεχόμενο ανάλογο με το αντίστοιχο των παραθύρων της Εφαρμογής 1, εφόσον ο χρήστης είχε επιλέξει να παρακολουθήσει τη βασική διαδικασία που ακολουθείται προκειμένου να επαυξηθεί η πραγματικότητα. Συγκεκριμένα, δημιουργείται ένα παράθυρο με τίτλο «Input Video» και σε αυτό εμφανίζεται το αντίστοιχο στιγμιότυπο. Επίσης, δημιουργείται ένα παράθυρο με τίτλο «Tracked Features» όπου απεικονίζεται το στιγμιότυπο, στο οποίο έχουν σχεδιαστεί τα σημεία ενδιαφέροντος που εντοπίστηκαν με τον αλγόριθμο SURF. Στις Εικόνες 5-29 και 5-30 φαίνονται τυχαίες σκηνές των παραθύρων αυτών για την περίπτωση που το βίντεο εικονίζει τη σκηνή που θα επαυξηθεί με το αγγείο και για την περίπτωση που το μοντέλο επαύξησης είναι το άγαλμα αντίστοιχα.

Επίσης, εμφανίζονται οι ομολογίες των σημείων που προέκυψαν από τη διαδικασία της συνταύτισης μετά τον έλεγχο του λόγου των αποστάσεων των εγγύτερων γειτόνων και την αποδοχή εκείνων για τις οποίες η απόσταση των διανυσμάτων περιγραφικών χαρακτηριστικών των αντίστοιχων σημείων είναι μικρότερη του ορισθέντος κατωφλίου (0.20). Το παράθυρο που δημιουργείται για το σκοπό αυτό έχει τίτλο «Matches». Ωστόσο, η ομογραφία δεν υπολογίζεται από το σύνολο αυτών των ομολογιών αλλά από τα inliers, που

επαληθεύουν την ευρεθείσα γεωμετρική σχέση μεταξύ της εικόνας - πρότυπο και του στιγμιότυπου. Οι έγκυρες αυτές συνταυτίσεις εμφανίζονται σε παράθυρο με τίτλο «Inliers». Στις Εικόνες 5-31 και 5-32 παρουσιάζονται τα δύο παραπάνω παράθυρα για τυχαία στιγμιότυπα των αρχείων βίντεο των δύο διαφορετικών περιοχών.



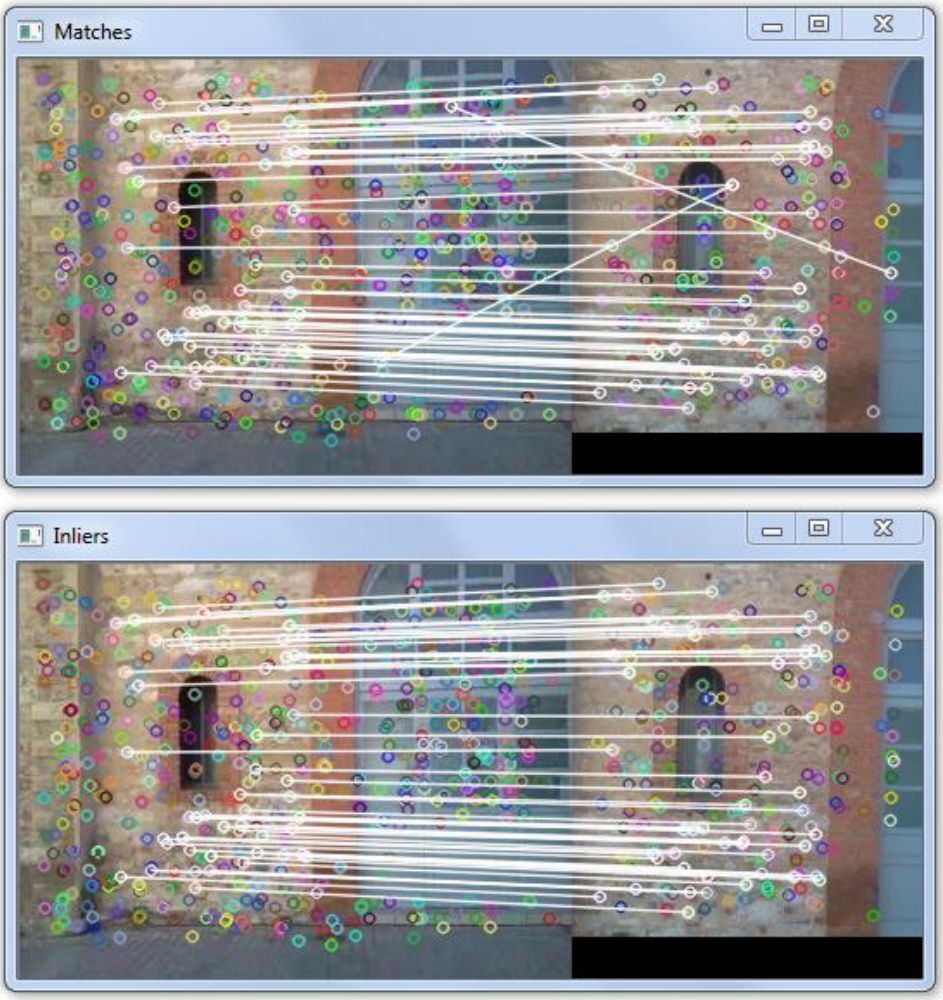
Εικόνα 5-29. Παράθυρο «Input Video», με ένα τυχαίο στιγμιότυπο του βίντεο [αριστερά], και παράθυρο «Tracked Features», με τα σημεία ενδιαφέροντος που εντοπίστηκαν στο ίδιο στιγμιότυπο [δεξιά] για την περίπτωση του βίντεο που θα επαυξηθεί με το αγγείο (Εφαρμογή 2).



Εικόνα 5-30. Παράθυρο «Input Video», με ένα τυχαίο στιγμιότυπο του βίντεο [αριστερά], και παράθυρο «Tracked Features», με τα σημεία ενδιαφέροντος που εντοπίστηκαν στο ίδιο στιγμιότυπο [δεξιά] για την περίπτωση του βίντεο που θα επαυξηθεί με το άγαλμα (Εφαρμογή 2).

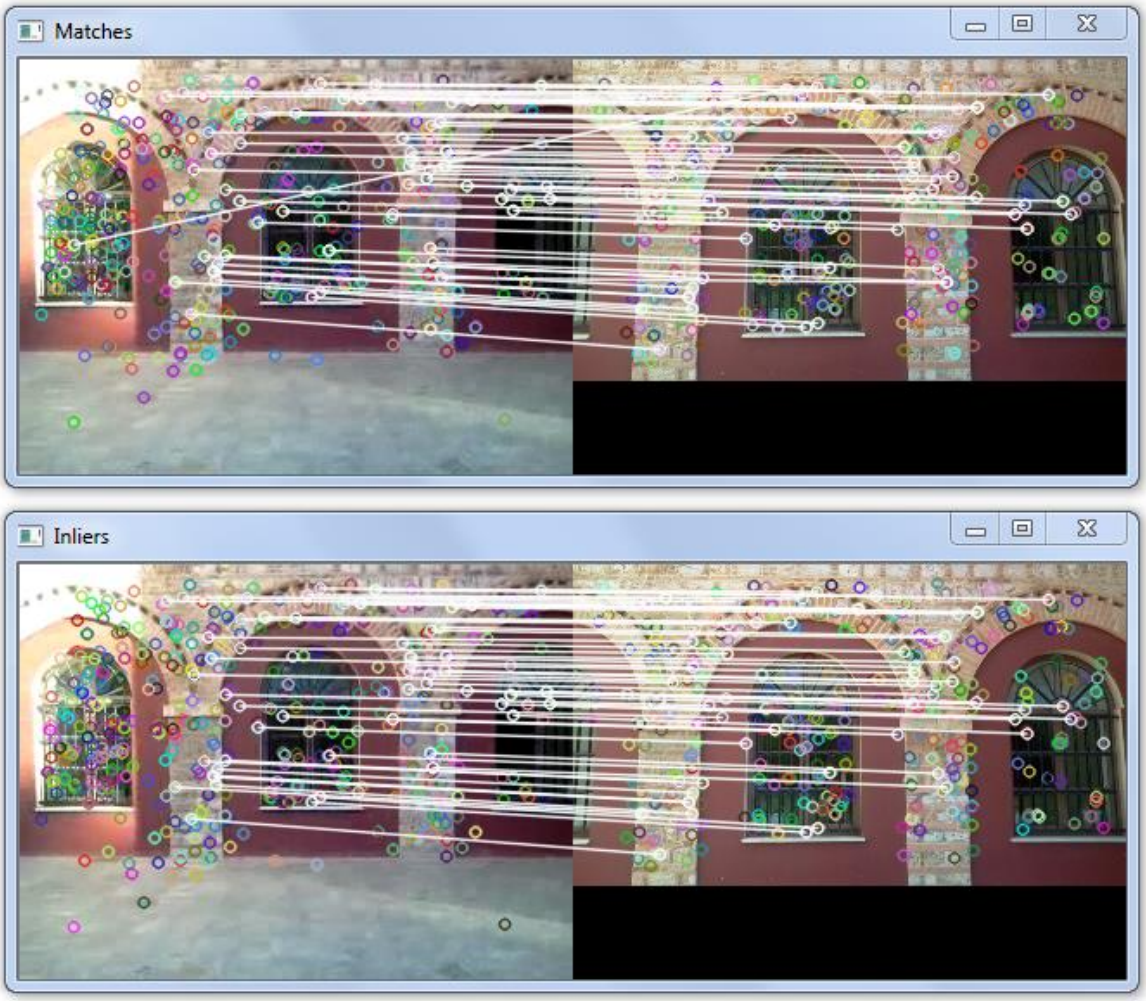
Προκειμένου να καταστεί εύκολος ο έλεγχος της ορθότητας του πίνακα της ομογραφίας από το χρήστη, πραγματοποιείται αναδόμηση του στιγμιότυπου με βάση το μετασχηματισμό αυτό, ώστε το αναδομημένο στιγμιότυπο να δείχνει το πρότυπο επίπεδο αντικείμενο ή τμήμα του, το οποίο εντοπίστηκε στο στιγμιότυπο, όπως περιγράφηκε στην παράγραφο 5.4.3.2. Το αναδομημένο στιγμιότυπο παρουσιάζεται σε παράθυρο με τίτλο «Warped Frame». Στις Εικόνες 5-33 και 5-34, δεξιά παρουσιάζεται το παράθυρο αυτό για το αντίστοιχο στιγμιότυπο του βίντεο που εικονίζεται στο παράθυρο «Input Video» στις ίδιες εικόνες, αριστερά, για τις

περιπτώσεις επαύξησης του βίντεο με τα δύο διαφορετικά τρισδιάστατα μοντέλα, το αγγείο και το άγαλμα αντίστοιχα.

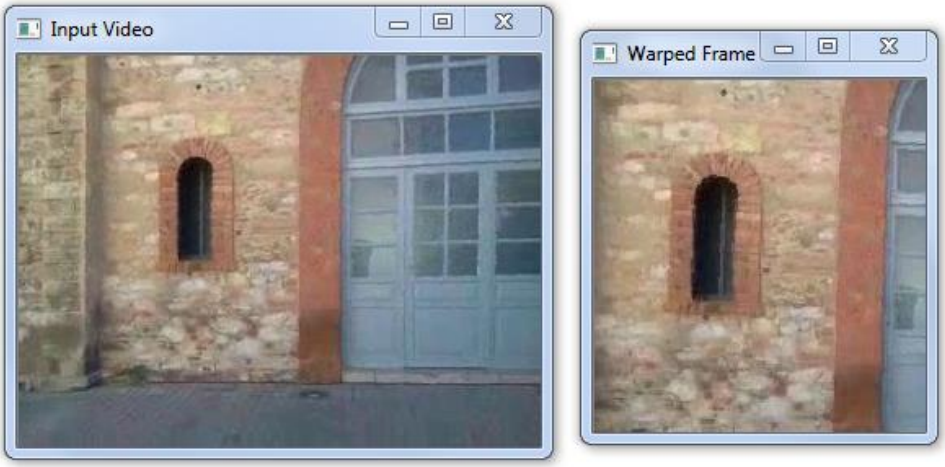


Εικόνα 5-31. Παράθυρο «Matches» με τις ομολογίες που προκύπτουν μετά τον έλεγχο του λόγου των αποστάσεων των εγγύτερων γειτόνων και την κατωφλίωση και αντίστοιχο παράθυρο «Inliers» με τις έγκυρες ομολογίες που προέκυψαν από τον αλγόριθμο RANSAC και χρησιμοποιήθηκαν για τον υπολογισμό του πίνακα της ομογραφίας, για μία τυχαία σκηνή ενός αρχείου βίντεο της Εφαρμογής 2 που θα επαυξηθεί με το τρισδιάστατο μοντέλου του αγγείου.

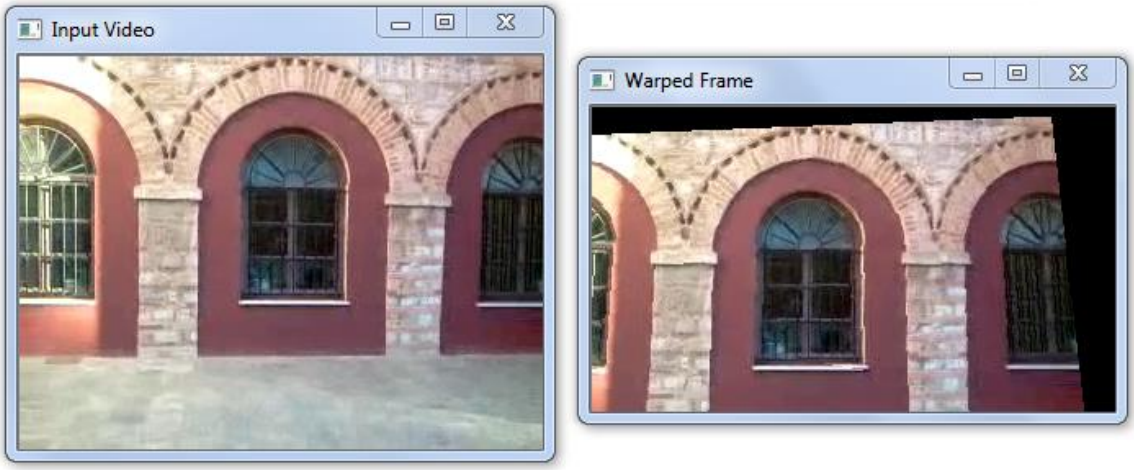
Τέλος, δημιουργείται ένα ακόμη παράθυρο το οποίο δείχνει το κάθε στιγμιότυπο του βίντεο με μία κλειστή κόκκινη γραμμή σχεδιασμένη σε αυτό, η οποία συνδέει τις τέσσερις κορυφές του πρότυπου αντικειμένου που εντοπίστηκαν στο στιγμιότυπο (βλ. Εικόνες 5-35 και 5-36). Αυτές είχαν ήδη υπολογιστεί με βάση το μετασχηματισμό ομογραφίας και τις συντεταγμένες pixel των τεσσάρων κορυφών της εικόνας - πρότυπο.



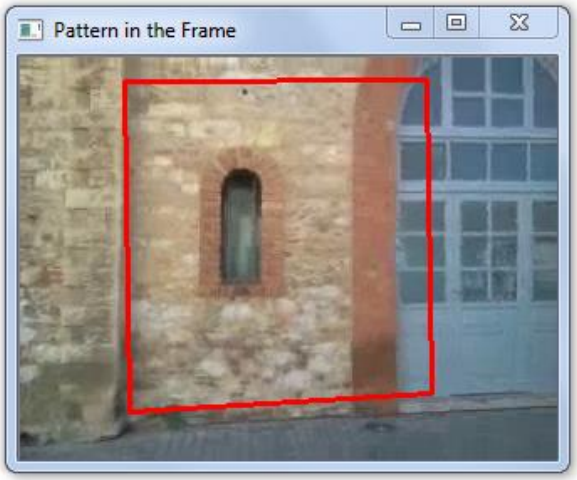
Εικόνα 5-32. Παράθυρο «Matches» με τις ομολογίες που προκύπτουν μετά τον έλεγχο του λόγου των αποστάσεων των εγγύτερων γειτόνων και την κατωφλίωση και αντίστοιχο παράθυρο «Inliers» με τις έγκυρες ομολογίες που προέκυψαν από τον αλγόριθμο RANSAC και χρησιμοποιήθηκαν για τον υπολογισμό του πίνακα της ομογραφίας, για μία τυχαία σκηνή ενός αρχείου βίντεο της Εφαρμογής 2 που θα επαυξηθεί με το τρισδιάστατο μοντέλο του αγάλματος.



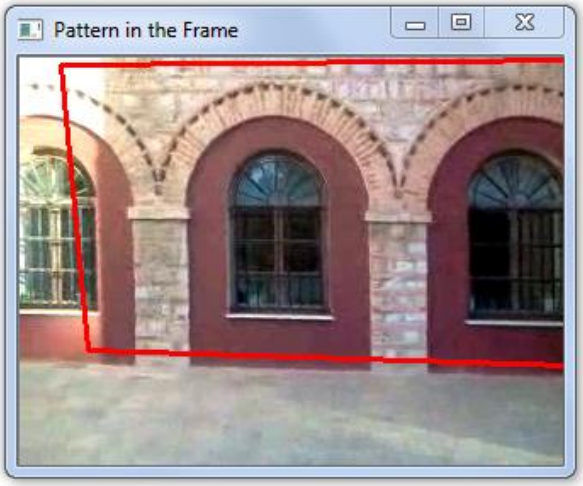
Εικόνα 5-33. Παράθυρο «Input Video» [αριστερά], που εικονίζει ένα τυχαίο στιγμιότυπο του βίντεο, και παράθυρο «Warped Frame» [δεξιά], που εικονίζει το αντίστοιχο αναδομημένο στιγμιότυπο που δείχνει το πρότυπο αντικείμενο, στην περίπτωση που το βίντεο θα επαυξηθεί με το αγγείο (Εφαρμογή 2).



Εικόνα 5-34. Παράθυρο «Input Video» [αριστερά], που εικονίζει ένα τυχαίο στιγμιότυπο του βίντεο, και παράθυρο «Warped Frame» [δεξιά], που εικονίζει το αντίστοιχο αναδομημένο στιγμιότυπο που δείχνει το πρότυπο αντικείμενο, στην περίπτωση που το βίντεο θα επαυξηθεί με το άγαλμα (Εφαρμογή 2).



Εικόνα 5-35. Παράθυρο «Pattern in the Frame» της Εφαρμογής 2, με ένα τυχαίο στιγμιότυπο του βίντεο της περιοχής που θα επαυξηθεί με το αγγείο. Στο στιγμιότυπο έχει σχεδιαστεί μία κόκκινη κλειστή γραμμή που ενώνει τις τέσσερις κορυφές του πρότυπου αντικειμένου που εντοπίστηκαν.



Εικόνα 5-36. Παράθυρο «Pattern in the Frame» της Εφαρμογής 2, με ένα τυχαίο στιγμιότυπο του βίντεο της περιοχής που θα επαυξηθεί με το άγαλμα. Στο στιγμιότυπο έχει σχεδιαστεί μία κόκκινη γραμμή που ενώνει τις τέσσερις κορυφές του πρότυπου αντικειμένου που εντοπίστηκαν.

5.5.3.3 Σχεδίαση της επαυξημένης σκηνής

Προτελευταίο βήμα της επαναληπτικής διαδικασίας είναι η σχεδίαση του στιγμιότυπου του βίντεο ως φόντου στη σχεδιαστική επιφάνεια του παραθύρου «AR Window». Καθώς η διαδικασία αυτή περιγράφηκε στην παράγραφο 5.4.3.3 (βλ. Διαγραμματική παρουσίαση 5-13), δεν αναφέρεται εκ νέου στην παράγραφο αυτή.

Τελικό βήμα της επαναληπτικής διαδικασίας είναι η σχεδίαση του τρισδιάστατου μοντέλου στην επιθυμητή θέση και με το σωστό τρόπο στο παράθυρο «AR Window». Για το λόγο αυτό, αρχικά καθορίζεται ο πίνακας προοπτικής προβολής (βλ. σχέση 4.40) με βάση τα στοιχεία του εσωτερικού προσανατολισμού και τη θεώρηση τιμών για τις αποστάσεις των κοντινών και μακρινών επιπέδων αποκοπής, οι οποίες δεν επηρεάζουν το τελικό αποτέλεσμα με την προϋπόθεση ότι οι συντεταγμένες Z_{EYE} των κορυφών του μοντέλου στο σύστημα της κάμερας βρίσκονται μεταξύ του κοντινού και του μακρινού επιπέδου αποκοπής. Τα στοιχεία του εσωτερικού προσανατολισμού είναι είτε αυτά που εισήγαγε ο χρήστης είτε τα προεπιλεγμένα από την εφαρμογή που προέκυψαν από τη βαθμονόμηση της κάμερας του κινητού τηλεφώνου.

Οι συντεταγμένες των κορυφών του τρισδιάστατου μοντέλου που ορίζονται στο αρχείο OBJ δεν έχουν το εύρος τιμών [-1,1] και αναφέρονται σε ένα σύστημα συντεταγμένων το οποίο δεν ταυτίζεται με το επίγειο σύστημα συντεταγμένων που ορίστηκε στο πλαίσιο της εφαρμογής. Για το λόγο αυτό, γίνεται μία κανονικοποίηση των συντεταγμένων του τρισδιάστατου μοντέλου στο εύρος [-1,1]. Αυτή πραγματοποιείται μέσω της βιβλιοθήκης GLM. Με την κανονικοποίηση αυτή, το κέντρο του τοποθετείται στο κέντρο του πρότυπου αντικειμένου, το μοντέλο αποκτά τον ίδιο προσανατολισμό με το πρότυπο αντικείμενο και η μεγαλύτερη διάστασή του αποκτά το ίδιο μέγεθος με τη μεγαλύτερη διάσταση του πρότυπου αντικειμένου. Ωστόσο, επειδή δεν είναι επιθυμητή η ταύτιση του μεγέθους των τρισδιάστατων μοντέλων με το μέγεθος του πρότυπου αντικειμένου, οι κανονικοποιημένες συντεταγμένες του πολλαπλασιάζονται με έναν παράγοντα κλίμακας. Στην περίπτωση του αγγείου, ως παράγοντας κλίμακας ορίζεται ο αριθμός 0.6, ενώ στην περίπτωση του αγάλματος, οι νέες συντεταγμένες του πολλαπλασιάζονται με τον αριθμό 0.3.

Αν ο πίνακας του μετασχηματισμού μοντέλου - κάμερας οριστεί με βάση τα στοιχεία του εξωτερικού προσανατολισμού (βλ. σχέση (4.2)), το κέντρο του μοντέλου θα τοποθετηθεί στο κέντρο του πρότυπου αντικειμένου και επιπλέον το μοντέλο θα έχει τον ίδιο προσανατολισμό με το πρότυπο αντικείμενο. Ωστόσο, στην εφαρμογή αυτή δεν είναι επιθυμητή η τοποθέτηση του μοντέλου στο κέντρο του επίπεδου πρότυπου αντικειμένου, ούτε η ακριβής ταύτιση του προσανατολισμού του με τον προσανατολισμό του πρότυπου αντικειμένου. Συνεπώς, ο πίνακας $\mathbf{M}_{\mathbf{MODELVIEW}}$ δεν ορίζεται μόνο από τις παραμέτρους του εξωτερικού προσανατολισμού. Αντίθετα, ορίζεται ως το γινόμενο του πίνακα που προκύπτει από τα στοιχεία του εξωτερικού προσανατολισμού της λήψης με το γινόμενο ενός πίνακα

μετάθεσης και ενός πίνακα στροφής. Συγκεκριμένα, ορίζεται ένας πίνακας μετάθεσης, ώστε το τρισδιάστατο μοντέλο να τοποθετηθεί στη σωστή θέση, και ένας πίνακας στροφής, ώστε αυτό να τοποθετηθεί με τον επιθυμητό προσανατολισμό. Ο πίνακας μετάθεσης, τόσο στην περίπτωση του τρισδιάστατου μοντέλου του αγγείου, όσο και στην περίπτωση του αγάλματος έχει την εξής μορφή:

$$\mathbf{M}_{\mathbf{TRANSLATION}} = \begin{bmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{5.21}$$

η οποία δηλώνει μετάθεση του μοντέλου κατά τους άξονες X και Y.

Ο πίνακας στροφής στην περίπτωση του τρισδιάστατου μοντέλου του αγγείου λαμβάνει τη μορφή:

$$\mathbf{M}_{\mathbf{ROTATION}} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{5.22}$$

γεγονός το οποίο σημαίνει ότι το μοντέλο περιστρέφεται γύρω από τον άξονα Z κατά γωνία θ, αριστερόστροφα. Αντίστοιχα, στην περίπτωση του τρισδιάστατου μοντέλου του αγάλματος, ο πίνακας στροφής λαμβάνει την ακόλουθη μορφή, που υποδηλώνει αριστερόστροφη στροφή γύρω από τον άξονα X, κατά γωνία φ:

$$\mathbf{M}_{\mathbf{ROTATION}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi & 0 \\ 0 & \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{5.23}$$

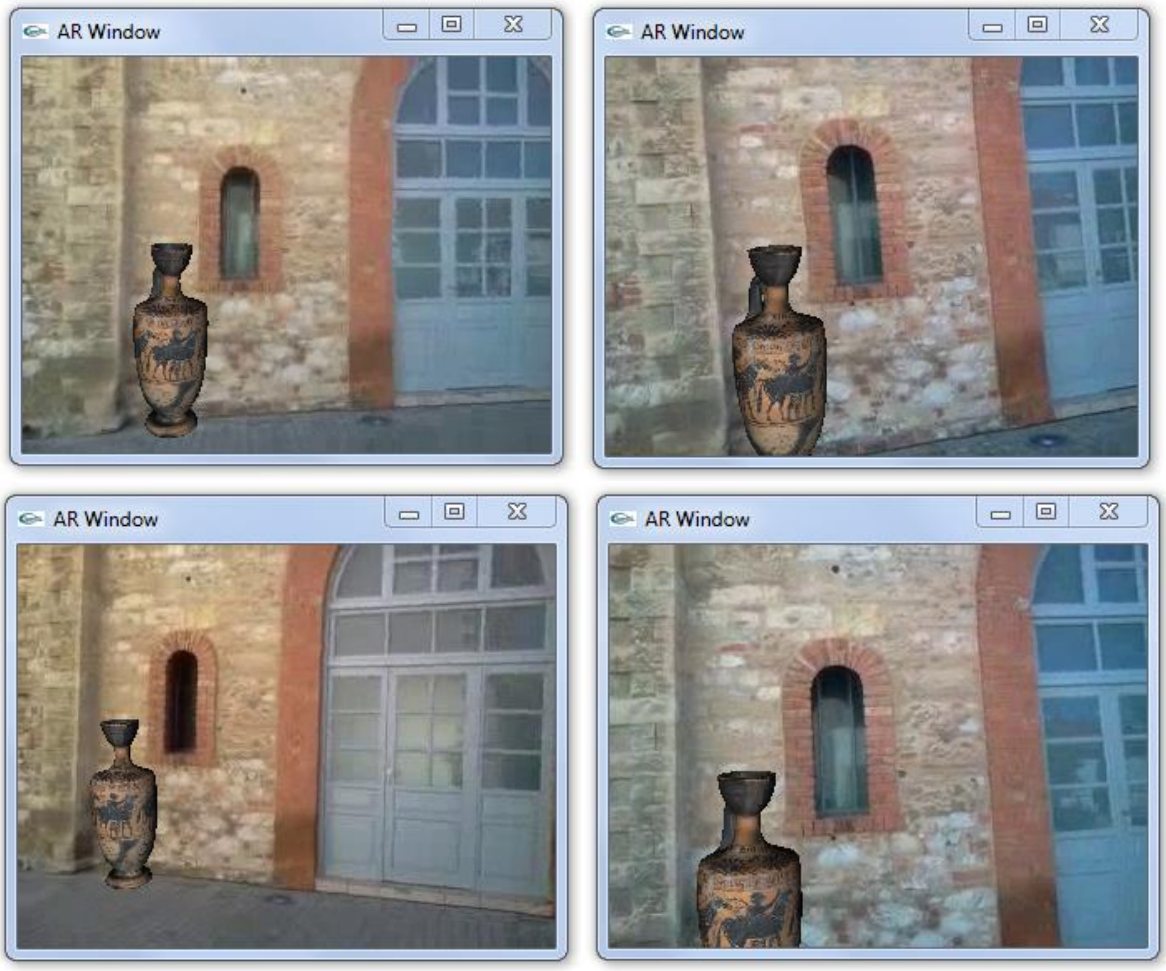
Συνεπώς, στο πλαίσιο της εφαρμογής αυτής, ο πίνακας του μετασχηματισμού μοντέλου - κάμερας δίνεται από τη σχέση:

$$\mathbf{M}_{\mathbf{MODELVIEW}} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \mathbf{M}_{\mathbf{TRANSLATION}} \cdot \mathbf{M}_{\mathbf{ROTATION}} \tag{5.24}$$

Μέσω της παραπάνω διαδικασίας, το κέντρο τόσο του αγγείου, όσο και του αγάλματος τοποθετείται κάτω αριστερά σε σχέση με το κέντρο των αντίστοιχων πρότυπων αντικειμένων και με τον επιθυμητό προσανατολισμό.

Μετά τον ορισμό των πινάκων του μετασχηματισμού προοπτικής προβολής και του μετασχηματισμού μοντέλου - κάμερας, ενεργοποιείται ο φωτισμός του μοντέλου και επιλέγεται ο προκαθορισμένος από την OpenGL φωτισμός.

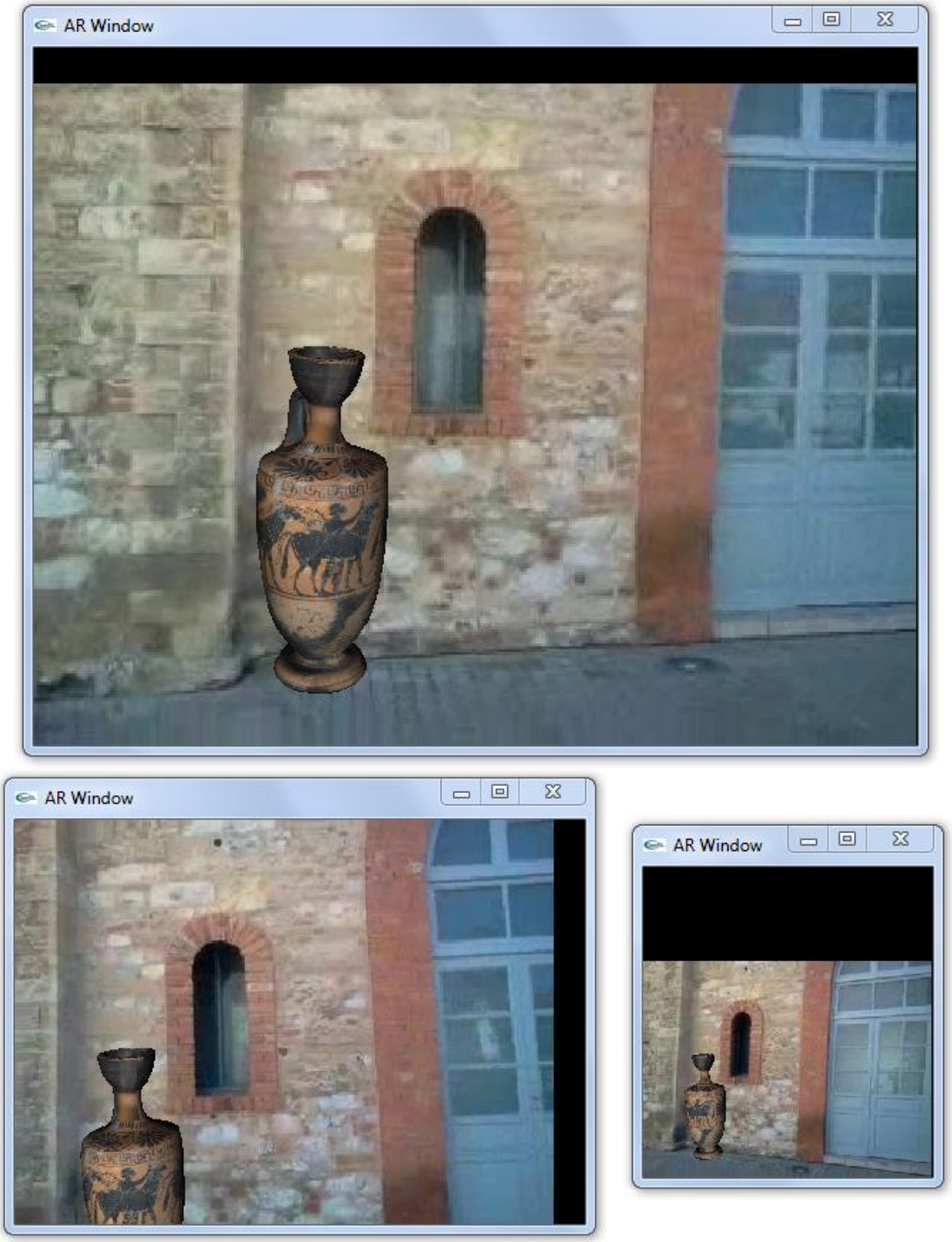
Όσον αφορά στο αγγείο, αυτό συνοδεύεται από μία εικόνα που αποτελεί την υφή του. Για το λόγο αυτό, γίνεται η απόδοση της υφής του μέσω της OpenGL (βλ. Διαγραμματική παρουσίαση 5-15). Ωστόσο, το άγαλμα δεν συνοδεύεται από κάποια εικόνα, με αποτέλεσμα το βήμα αυτό να παραλείπεται στην περίπτωσή του.



Εικόνα 5-37. Στιγμιότυπα από το παράθυρο «AR Window» της Εφαρμογής 2, στο οποίο επαυξάνεται το πραγματικό περιεχόμενο του βίντεο με το τρισδιάστατο μοντέλο του αγγείου. Το μέγεθος της σχεδιαστικής επιφάνειας των παραθύρων είναι ίδιο με το μέγεθος των στιγμιότυπων του βίντεο.

Στις Εικόνες 5-37 έως και 5-40 παρουσιάζονται στιγμιότυπα από το παράθυρο «AR Window», όπου επαυξάνονται οι σκηνές του βίντεο της περιοχής μπροστά από τη νότια πλευρά του κτηρίου Καθαρτήρια του πολυχώρου «Τεχνόπολις» και εκείνης μπροστά από τη βόρεια πλευρά του κτηρίου Μακρυγιάννη. Στις Εικόνες 5-37 και 5-39 το μέγεθος της σχεδιαστικής επιφάνειας των παραθύρων που εικονίζονται είναι το ίδιο με το μέγεθος των στιγμιότυπων του βίντεο. Αντίθετα, το μέγεθος των παραθύρων των Εικόνων 5-38 και 5-40 έχει μεταβληθεί από το χρήστη. Στην περίπτωση αυτή, καλείται αυτόματα η συνάρτηση που

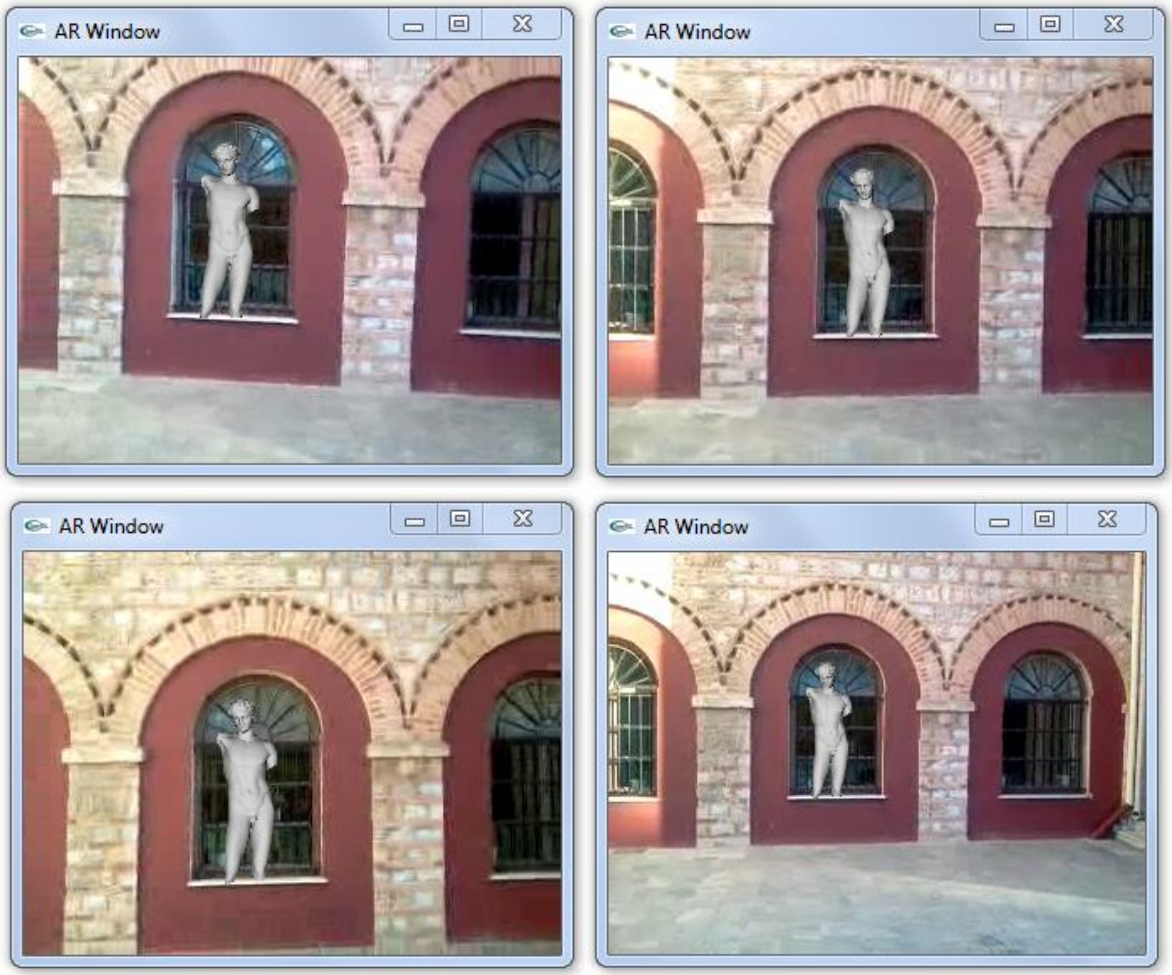
έχει οριστεί να καλείται όταν μεταβάλλονται οι διαστάσεις του παραθύρου, η οποία διατηρεί το λόγο των διαστάσεων των πλευρών του στιγμιότυπου.



Εικόνα 5-38. Στιγμιότυπα από το παράθυρο «AR Window» της Εφαρμογής 2, το μέγεθος του οποίου έχει μεταβληθεί από το χρήστη (περίπτωση επαύξησης του περιεχομένου του βίντεο με το αγγείο).

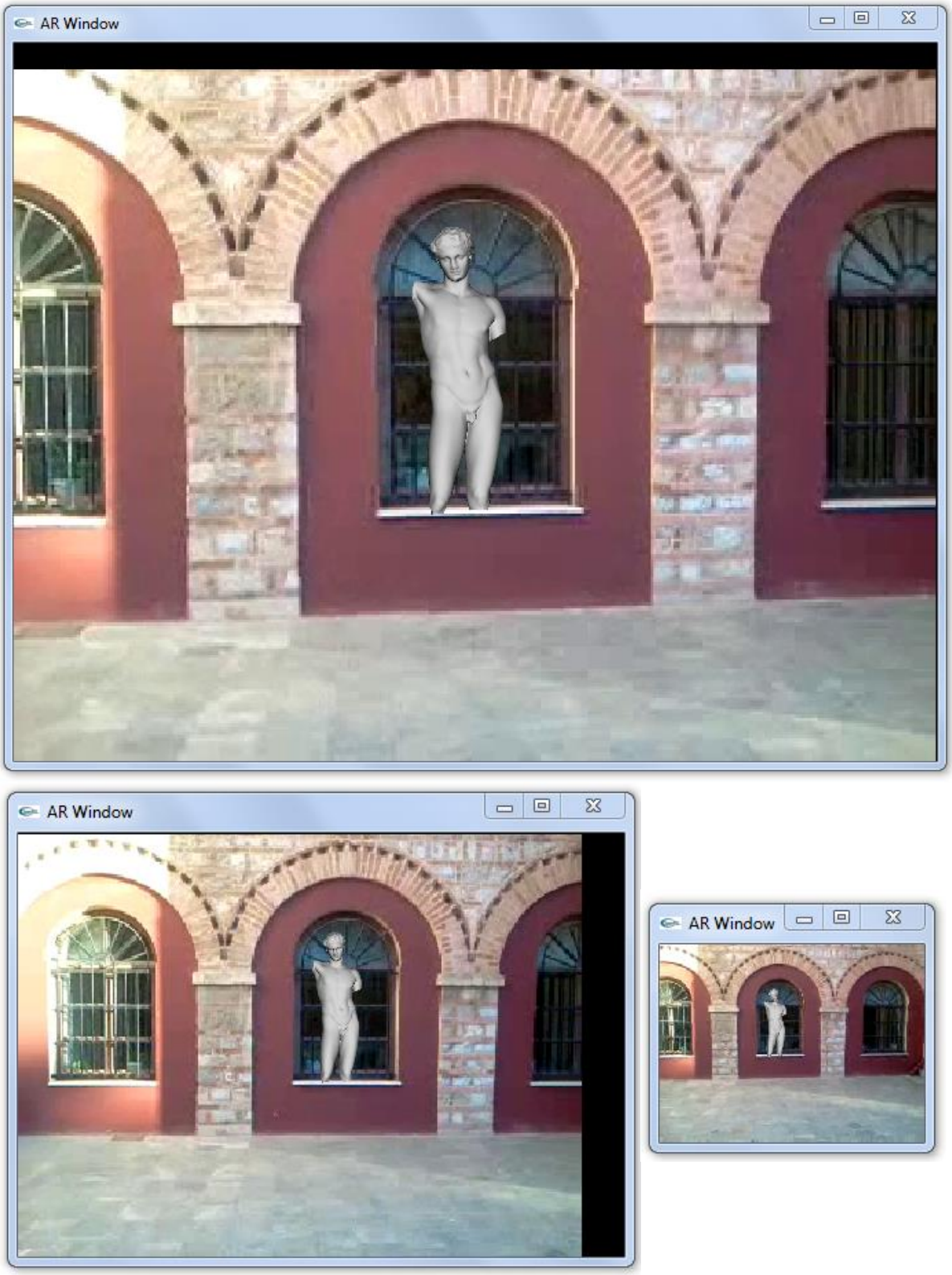
Η διαδικασία που περιγράφηκε παραπάνω πραγματοποιείται στην περίπτωση που έχουν εντοπιστεί τουλάχιστον πέντε ομολογίες σημείων. Διαφορετικά, σχεδιάζεται το στιγμιότυπο

ως φόντο στη σχεδιαστική επιφάνεια του παραθύρου «AR Window» και η σχεδίαση του τρισδιάστατου μοντέλου παραλείπεται. Επίσης, στην περίπτωση αυτή μεταβάλλεται και το περιεχόμενο των παραθύρων «Inliers», «Warped Frame» και «Pattern in the Frame», που εμφανίζονται αν ο χρήστης το επέλεξε στο αρχικό στάδιο της εφαρμογής. Συγκεκριμένα, στα δύο πρώτα εμφανίζεται μία εικόνα που ενημερώνει το χρήστη ότι το πρότυπο επίπεδο αντικείμενο δεν εντοπίστηκε, ενώ στο τελευταίο απεικονίζεται το στιγμιότυπο χωρίς να είναι σχεδιασμένη η κόκκινη γραμμή που ενώνει τις τέσσερις κορυφές του πρότυπου αντικειμένου.

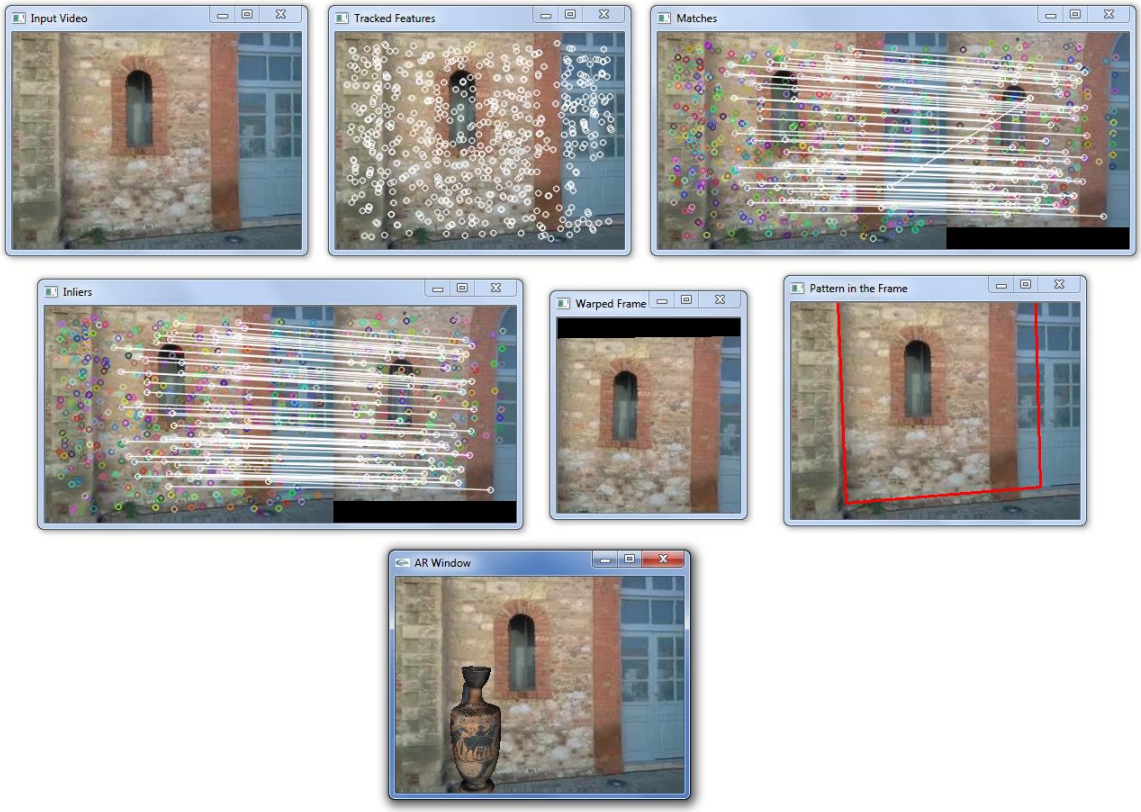


Εικόνα 5-39. Στιγμιότυπα από το παράθυρο «AR Window» της Εφαρμογής 2, στο οποίο επαυξάνεται το πραγματικό περιεχόμενο του βίντεο με το τρισδιάστατο μοντέλο του αγάλματος. Το μέγεθος της σχεδιαστικής επιφάνειας των παραθύρων είναι ίδιο με το μέγεθος των στιγμιότυπων του βίντεο.

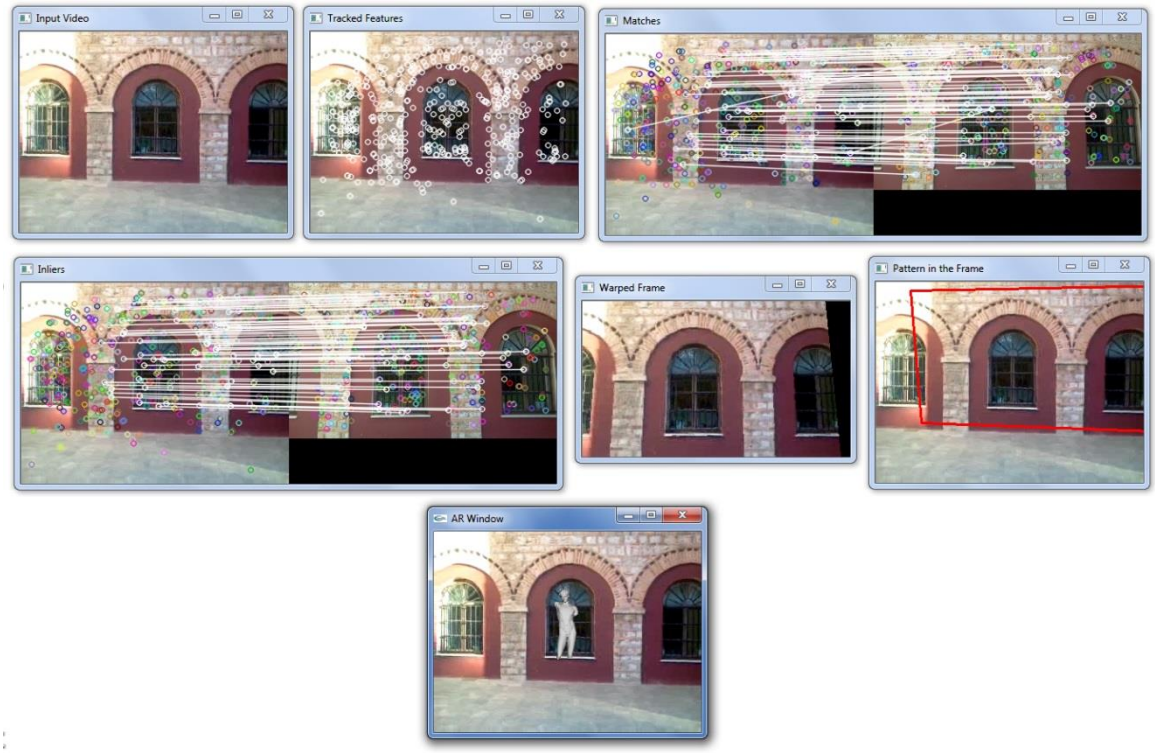
Στις Εικόνες 5-41 και 5-42 παρουσιάζονται ενδεικτικά δύο στιγμιότυπα από την Εφαρμογή 2 κατά τη διεξαγωγή της επαναληπτικής διαδικασίας, στην περίπτωση που ο χρήστης είχε επιλέξει να παρακολουθήσει τη βασική διαδικασία που «κρύβεται» πίσω από την επαύξηση της πραγματικότητας, για την περίπτωση του τρισδιάστατου μοντέλου του αγγείου και του αγάλματος αντίστοιχα. Τέλος, στις Διαγραμματικές παρουσιάσεις 5-22 έως και 5-26 παρουσιάζονται τα διαγράμματα ροής για την επαναληπτική διαδικασία που περιγράφηκε στις παραγράφους 5.5.3.1 έως και 5.5.3.3.



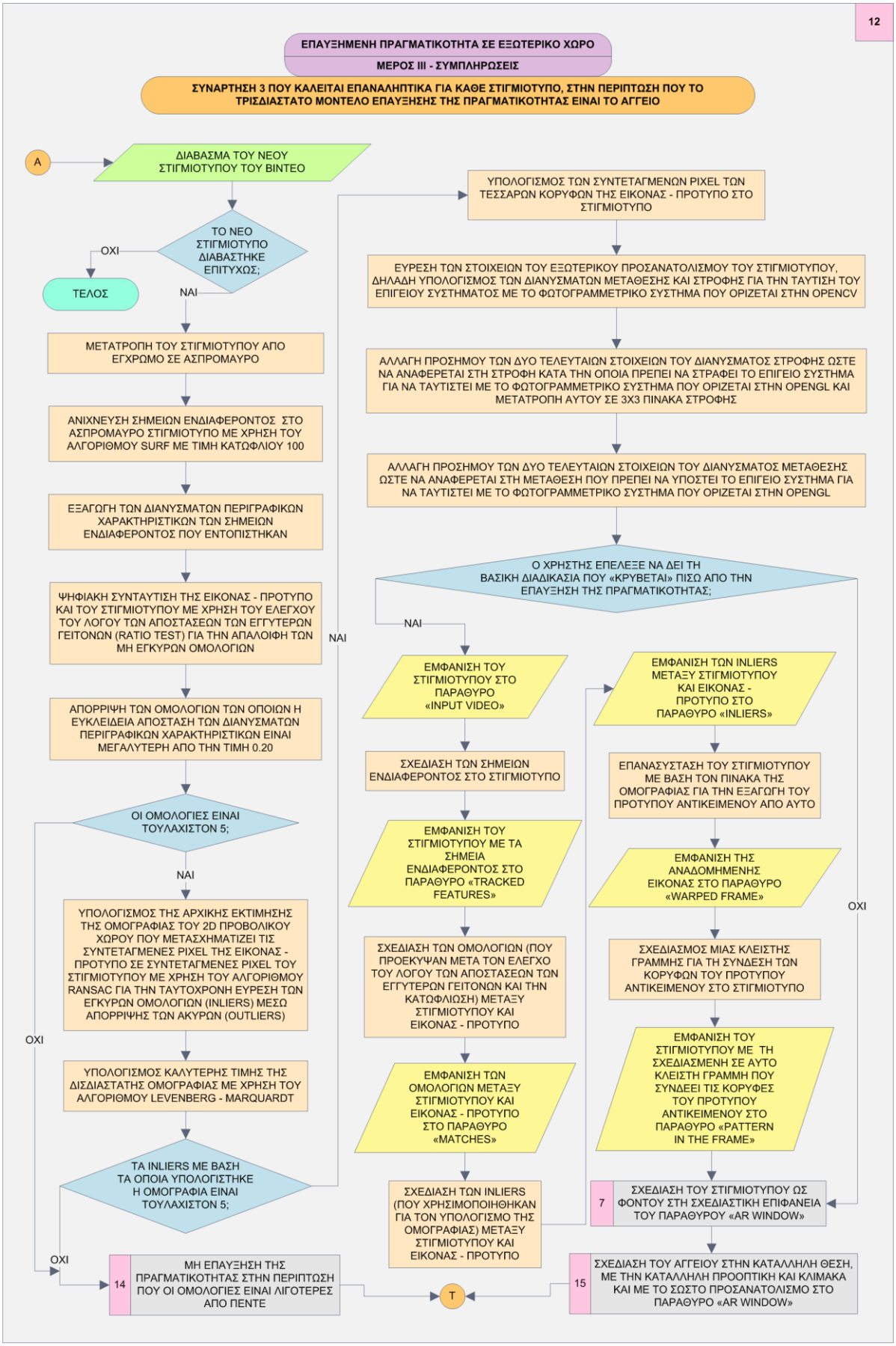
Εικόνα 5-40. Στιγμιότυπα από το παράθυρο «AR Window» της Εφαρμογής 2, το μέγεθος του οποίου έχει μεταβληθεί από το χρήστη (περίπτωση επαύξησης του περιεχομένου του βίντεο με το άγαλμα).



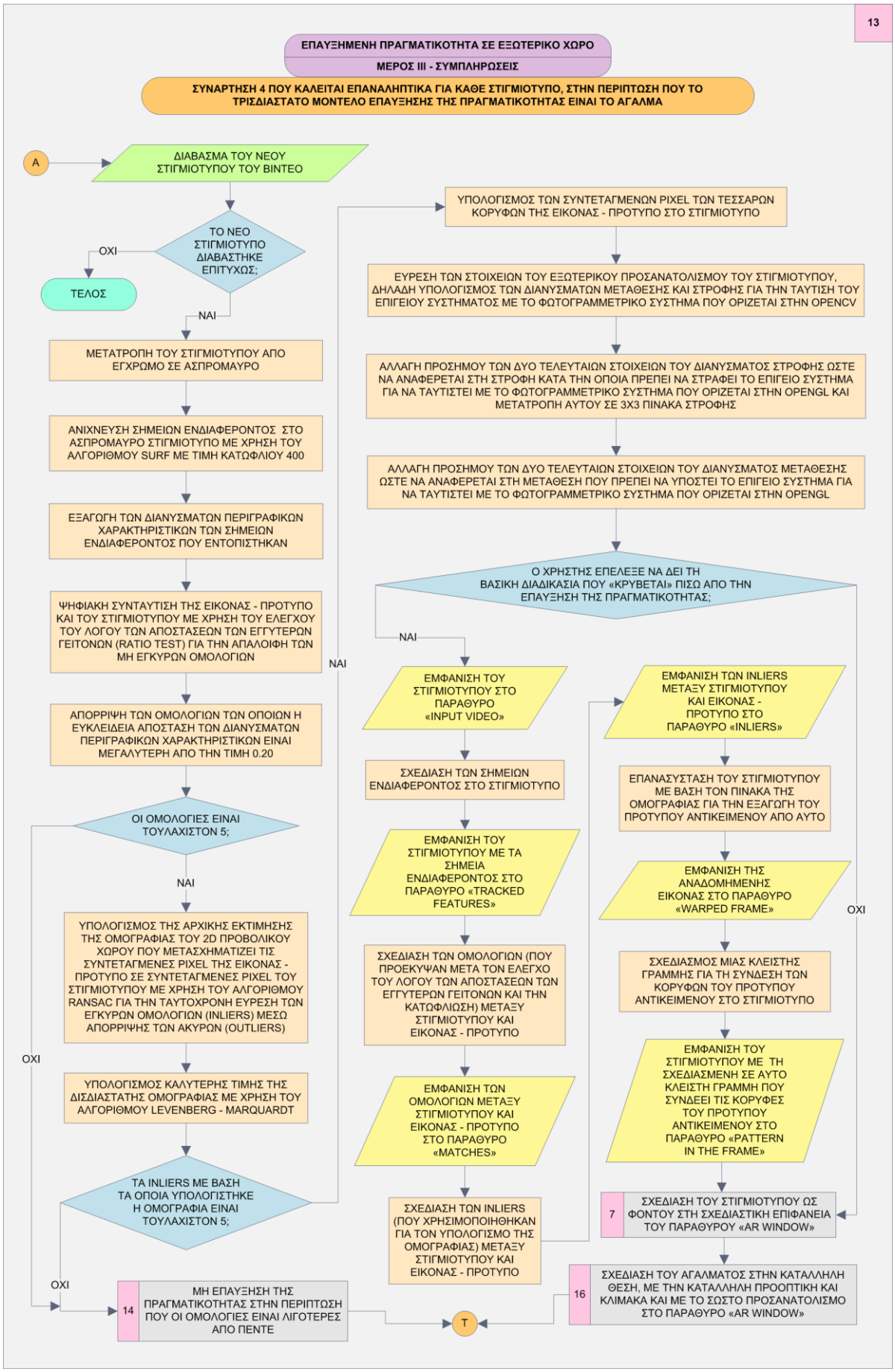
Εικόνα 5-41. Στιγμιότυπο από την Εφαρμογή 2, στην περίπτωση επαύξησης της περιοχής μπροστά από τη νότια πλευρά του κτηρίου Καθαρτήρια.



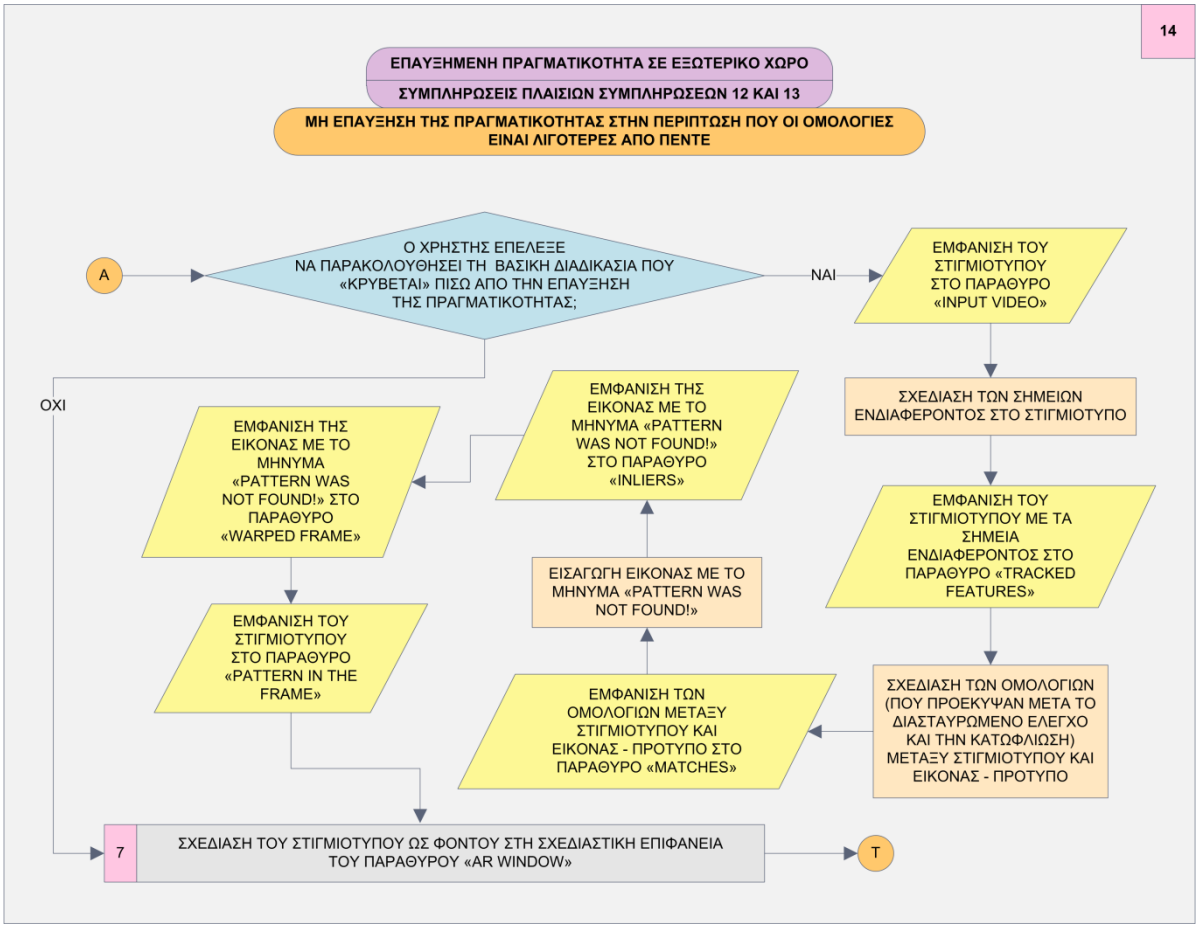
Εικόνα 5-42. Στιγμιότυπο από την Εφαρμογή 2, στην περίπτωση επαύξησης της βόρειας πλευράς του κτηρίου Μακρυγιάννη.



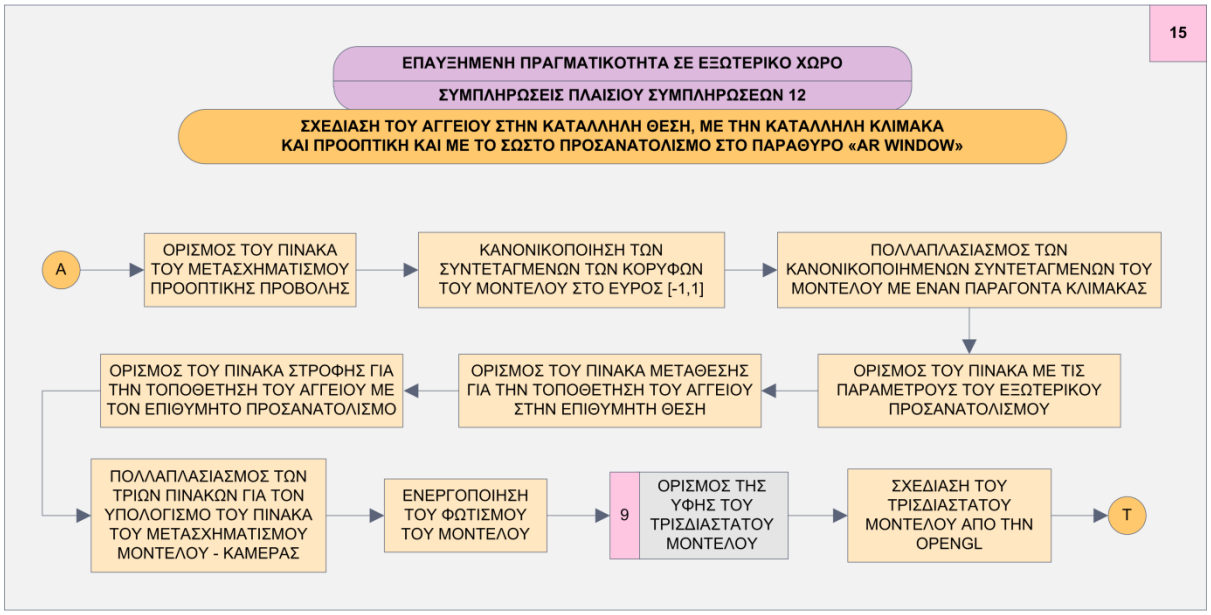
Διαγραμματική παρουσίαση 5-22. Εφαρμογή 2, πλαίσιο συμπληρώσεων 12.



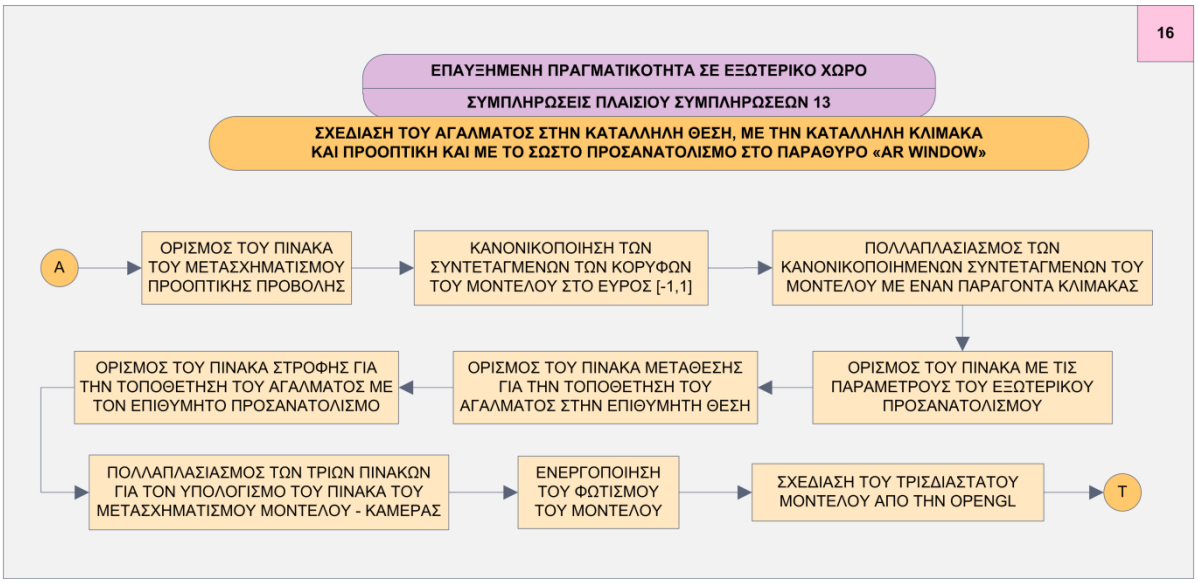
Διαγραμματική παρουσίαση 5-23. Εφαρμογή 2, πλαίσιο συμπληρώσεων 13.



Διαγραμματική παρουσίαση 5-24. Εφαρμογή 2, πλαίσιο συμπληρώσεων 14.



Διαγραμματική παρουσίαση 5-25. Εφαρμογή 2, πλαίσιο συμπληρώσεων 15.



Διαγραμματική παρουσίαση 5-26. Εφαρμογή 2, πλαίσιο συμπληρώσεων 16.

ΣΥΜΠΕΡΑΣΜΑΤΑ

Σκοπός της παρούσας διπλωματικής εργασίας είναι η ανάπτυξη εφαρμογών επαυξημένης πραγματικότητας βάσει επίπεδου προτύπου, με αξιοποίηση μεθόδων και αλγορίθμων της φωτογραμμετρίας και της όρασης υπολογιστών. Στο πλαίσιο αυτό, έγινε μία περιγραφή της έννοιας της επαυξημένης πραγματικότητας, παρουσιάστηκε το θεωρητικό υπόβαθρο των διαδικασιών οι οποίες χρησιμοποιήθηκαν για την περάτωση των εφαρμογών και έγινε αναφορά στο προγραμματιστικό μέρος των εφαρμογών (περιβάλλον ανάπτυξης, γλώσσα προγραμματισμού και χρησιμοποιούμενες βιβλιοθήκες) ώστε στο τελευταίο κεφάλαιο της εργασίας να παρουσιαστούν τα αποτελέσματα των εφαρμογών και η ακριβής διαδικασία που ακολουθήθηκε. Έχοντας ολοκληρώσει την ανάπτυξη των εφαρμογών, κρίνεται αναγκαία η επισήμανση κάποιων παρατηρήσεων και συμπερασμάτων καθώς επίσης και διαφόρων προτάσεων για μελλοντική ενασχόληση και βελτίωση αυτών.

Τα αποτελέσματα των εφαρμογών κρίνονται ικανοποιητικά. Τόσο η εφαρμογή επαυξημένης πραγματικότητας εσωτερικού χώρου όσο και η αντίστοιχη εξωτερικού χώρου αναγνωρίζουν επιτυχώς το πρότυπο επίπεδο αντικείμενο όταν αυτό τοποθετηθεί στο οπτικό πεδίο της κάμερας, ακόμη και στην περίπτωση κατά την οποία απεικονίζεται στη σκηνή ένα τμήμα του. Επίσης, η αναγνώρισή του γίνεται ανεξάρτητα από τις συνθήκες φωτισμού, τον προσανατολισμό του πρότυπου αντικειμένου και το μέγεθός του. Η σωστή αναγνώρισή του – η οποία οφείλεται κυρίως στο χρησιμοποιούμενο αλγόριθμο ανίχνευσης και περιγραφής των σημείων ενδιαφέροντος – σε συνδυασμό με την ορθότητα των αποτελεσμάτων της βαθμονόμησης είναι τα δύο βασικότερα στοιχεία που κρίνουν την επιτυχία επαύξησης του πραγματικού κόσμου και τη ρεαλιστικότητα των επαυξημένων σκηνών. Εξάλλου, η ρεαλιστική επαύξηση της πραγματικότητας, δηλαδή η ενσωμάτωση του εικονικού μοντέλου στον τρισδιάστατο χώρο έτσι ώστε να μην ξεχωρίζει από την πραγματική σκηνή αλλά να αποτελεί ένα με αυτή, είναι ένας από τους σημαντικότερους στόχους των εφαρμογών επαυξημένης πραγματικότητας, ο οποίος επιτεύχθηκε σε υψηλό βαθμό στην παρούσα εργασία.

Μη ρεαλιστική είναι η επαύξηση της σκηνής στην περίπτωση που ένα αντικείμενο του πραγματικού κόσμου τοποθετηθεί μπροστά από τη θέση στην οποία προορίζεται να τοποθετηθεί το εικονικό μοντέλο, προς την πλευρά της κάμερας. Η ιδανική λύση θα ήταν η απόκρυψη του τμήματος του τρισδιάστατου μοντέλου επαύξησης που κρύβεται από το

πραγματικό αντικείμενο. Ωστόσο, στις εφαρμογές που προγραμματίστηκαν, κάθε στιγμιότυπο της πραγματικής σκηνής ορίζεται ως εικόνα που καλύπτει τη σχεδιαστική επιφάνεια του παραθύρου επαύξησης της πραγματικότητας, αποτελώντας το φόντο του, χωρίς να εξάγεται από αυτό η πληροφορία της τρίτης διάστασης. Η τελευταία θα μπορούσε να χρησιμοποιηθεί στον ορισμό των αποκρύψεων. Συνεπώς, μία ενδιαφέρουσα μελλοντική επέκταση των εφαρμογών είναι η κατασκευή σε πραγματικό χρόνο του χάρτη βάθους (depth map) για κάθε στιγμιότυπο, για τον ορισμό των τμημάτων των εικονικών μοντέλων τα οποία αποκρύπτονται από τα αντικείμενα του πραγματικού κόσμου.

Οι εφαρμογές που αναπτύχθηκαν έχουν ικανοποιητικά αποτελέσματα στην περίπτωση που το πρότυπο αντικείμενο είναι επίπεδο ή έχει μικρό ανάγλυφο σε σχέση με την απόσταση λήψης, λόγω της γεωμετρικής σχέσης του δισδιάστατου προβολικού μετασχηματισμού που θεωρήθηκε ότι συνδέει την εικόνα - πρότυπο και το εκάστοτε στιγμιότυπο. Συνεπώς, μία μελλοντική επέκταση των δυνατοτήτων των εφαρμογών θα μπορούσε να είναι η επαύξηση μη επίπεδων σκηνών του πραγματικού κόσμου, ανεξάρτητα από τη γεωμετρία τους. Η λύση σε ένα τέτοιο πρόβλημα μπορεί να προκύψει από την επιπολική γεωμετρία.

Η εφαρμογή επαυξημένης πραγματικότητας εσωτερικού χώρου, η οποία υποστηρίζει τη συμπλήρωση της ορθοφωτογραφίας με την πληροφορία της τρίτης διάστασης, ταιριάζει με το μέσο που χρησιμοποιήθηκε για την επαύξηση σε πραγματικό χρόνο, δηλαδή την κάμερα του υπολογιστή, καθώς ο χρήστης μπορεί εύκολα να τοποθετήσει την εκτυπωμένη ορθοφωτογραφία μπροστά από την κάμερα του υπολογιστή του και να εξετάσει το ανάγλυφο της συγκεκριμένης περιοχής. Αντίθετα, δεν θα ήταν διατεθειμένος να μεταφέρει το φορητό υπολογιστή του στον πολυχώρο «Τεχνόπολις» ή στην περιοχή του νέου Μουσείου της Ακρόπολης για να παρακολουθήσει σε ζωντανό χρόνο την επαύξηση της πραγματικότητας στις δύο αυτές τοποθεσίες. Για το λόγο αυτό, η Εφαρμογή 2 δεν προγραμματίστηκε ώστε να είναι πραγματικού χρόνου και προϋποθέτει από το χρήστη να έχει εγγράψει ένα βίντεο μέσω της κινητής συσκευής του και εν συνεχεία να παρακολουθήσει, από το χώρο στον οποίο βρίσκεται ο υπολογιστής του, τις επαυξημένες σκηνές. Ωστόσο, θα ήταν προτιμότερη η επαύξηση της πραγματικότητας σε πραγματικό χρόνο μέσω μίας κινητής συσκευής χειρός. Έτσι, μελλοντικός στόχος είναι η τροποποίηση των εφαρμογών ώστε να εκτελούνται και σε ένα κινητό τηλέφωνο ή σε ένα tablet υπολογιστή, δηλαδή σε συσκευές χειρός που μεταφέρονται εύκολα και μπορούν να χρησιμοποιηθούν χωρίς δυσκολία από χρήστες εφαρμογών επαυξημένης πραγματικότητας. Άλλη πιθανή βελτίωση - διαφοροποίηση των εφαρμογών θα μπορούσε να είναι η εισαγωγή από το χρήστη της εικόνας - πρότυπο και του τρισδιάστατου μοντέλου και ο καθορισμός από αυτόν της επιθυμητής θέσης του τελευταίου σε σχέση με το πρότυπο αντικείμενο. Παράλληλα, η υποστήριξη της δυνατότητας βαθμονόμησης της κάμερας, με την οποία λαμβάνονται τα στιγμιότυπα προς επαύξηση, στην αρχή των εφαρμογών, μέσω λήψης βίντεο μίας σκακιέρας, θα ήταν μία ακόμη χρήσιμη συμπλήρωση των εφαρμογών.

Η συμβολή των εφαρμογών επαυξημένης πραγματικότητας και η χρήση τους σε διάφορους τομείς έχει ήδη παρουσιαστεί στο πρώτο κεφάλαιο της παρούσας εργασίας. Όσον αφορά στις συγκεκριμένες εφαρμογές, η Εφαρμογή 1 μπορεί να χρησιμοποιηθεί για την οπτικοποίηση πληροφοριών τρίτης διάστασης και για την εξέταση του αναγλύφου μίας περιοχής με τρόπο παραστατικό, χωρίς να υφίσταται η ανάγκη τρισδιάστατης εκτύπωσης του ΨΜΕ. Μία πιθανή επέκταση της εφαρμογής θα μπορούσε να περιλαμβάνει την εισαγωγή διαφορετικών ΨΜΕ και πρότυπων εικόνων και την αυτόματη αναγνώριση της συγκεκριμένης ορθοφωτογραφίας ή άλλου είδους χάρτη που βρίσκεται στο οπτικό πεδίο της κάμερας, προκειμένου να επαυξηθεί με το κατάλληλο ΨΜΕ. Έτσι, θα μπορούσε να χρησιμοποιηθεί στη χαρτογραφία, καθώς παρέχει ένα ρεαλιστικό και καινοτόμο τρόπο αναπαράστασης της τοπογραφίας μίας περιοχής. Επίσης, θα μπορούσε να δώσει σε μαθητές τη δυνατότητα να εξετάζουν το ανάγλυφο περιοχών που απεικονίζονται στα σχολικά βιβλία γεωγραφίας τους, με τη χρήση της οθόνης ενός υπολογιστή. Επιπλέον, θα μπορούσε να χρησιμοποιηθεί από τουρίστες για την τρισδιάστατη παρατήρηση περιοχών που παρουσιάζουν ιδιαίτερο ενδιαφέρον και απεικονίζονται σε τουριστικούς χάρτες.

Η χρησιμότητα της Εφαρμογής 2 θα μπορούσε να συνοψισθεί στην εξέταση αν τα συγκεκριμένα μοντέλα θα μπορούσαν να τοποθετηθούν στις καθορισμένες από την εφαρμογή θέσεις, από αισθητικής άποψης. Ωστόσο, τόσο τα μοντέλα όσο και οι περιοχές που χρησιμοποιήθηκαν είναι ενδεικτικά και περισσότερη σημασία δόθηκε όχι σε αυτά, αλλά στη διαδικασία που ακολουθήθηκε. Έτσι, ανεξάρτητα από τα συγκεκριμένα μοντέλα και τις αντίστοιχες περιοχές, ενδιαφέρουσες χρήσεις της εφαρμογής θα μπορούσαν να είναι η επαύξηση ενός κτηρίου με τμήμα του το οποίο υπήρχε στο παρελθόν αλλά σήμερα έχει καταστραφεί, ή η τοποθέτηση ενός κτηρίου, που δεν έχει ακόμη κατασκευαστεί, στη θέση στην οποία σχεδιάζεται να τοποθετηθεί, για να εξεταστεί αν είναι συμβατό με τον περιβάλλοντα χώρο. Τέλος, λαμβάνοντας υπ΄ όψιν τη μελλοντική επέκταση της εφαρμογής για την υποστήριξη ενός τρισδιάστατου πρότυπου αντικειμένου, ενδιαφέρουσα χρήση της θα μπορούσε να είναι η επαύξηση ενός ημι-κατεστραμμένου αγάλματος ή αρχαίου μνημείου με το τμήμα το οποίο λείπει από τη σημερινή κατάστασή του, για την οπτικοποίηση της μορφής του, όπως ήταν στο παρελθόν.

ΑΝΑΦΟΡΕΣ

[1] *Wikipedia, The Free Encyclopedia* [Ηλεκτρονικό]. Διαθέσιμο: en.wikepedia.org/wiki/Augmented_reality. [Πρόσβαση: 4 Μαΐου 2013].

[2] Δ. Νικολαΐδης, «Επαυξημένη Πραγματικότητα: Πολλαπλασιάζοντας τις δυνατότητες των αισθήσεων», *Περισκόπιο της Επιστήμης,* τεύχος 270, Μάρτιος 2003.

[3] R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier and B. MacIntyre, «Recent Advances in Augmented Reality», *IEEE Computer Graphics and Applications,* vol. 21, no. 6, pp. 34-47, November/December 2001.

[4] G. Kipper and J. Rampolla, *Augmented Reality. An Emerging Technologies Guide to AR*, USA: Syngress, 2013.

[5] R. Azuma, «A Survey of Augmented Reality», *Presence: Teleoperators and Virtual Environments,* vol 6, no. 4, pp. 355-385, August 1997.

[6] O. Bimber and R. Raskar, *Spatial Augmented Reality: Merging Real and Virtual Worlds*, Wellesley, MA: A K Peters Ltd, 2005.

[7] Δ. Δούκογλου και Π. Θεολόγου, «Ανάπτυξη Παιχνιδιού Σοβαρού Σκοπού Αναπαραγωγής Αθλητικών Γεγονότων», Διπλωματική Εργασία, Εθνικό Μετσόβιο Πολυτεχνείο, Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών, Αθήνα, 2011.

[8] P. Milgram, H. Takemura, A. Utsumi and F. Kishino, «Augmented Reality: A class of displays on the reality - virtuality continuum», *SPIE Proceedings: Telemanipulator and Telepresence Technologies,* vol. 2351, pp. 282-292, 1994.

[9] F. Baum, *The Master Key*, Indianapolis: Bowen-Merrill Company, 1901.

[10] I. Sutherland, «A head-mounted three dimensional display», in *Proc. Fall Joint Computer Conference*, San Francisco, California, 1968.

[11] *Wikipedia, The Free Encyclopedia* [Ηλεκτρονικό]. Διαθέσιμο: en.wikipedia.org/wiki/The_Sword_of_Damocles_(virtual_reality). [Πρόσβαση: 5 Μαΐου 2013].

[12] *Wikipedia, The Free Encyclopedia* [Ηλεκτρονικό]. Διαθέσιμο: en.wikipedia.org/wiki/Videoplace. [Πρόσβαση: 5 Μαΐου 2013].

[13] *The Digital Age - Myron Krueger* [Ηλεκτρονικό]. Διαθέσιμο: http://thedigitalage.pbworks.com/w/page/22039083/Myron%Krueger. [Πρόσβαση: 5 Μαΐου 2013].

[14] G. Levin. «Computer Vision for Artists and Designers: Pedagogic Tools and Techniques for Novice Programmers», *FLONG* [Ηλεκτρονικό]. Διαθέσιμο: http://www.flong.com/texts/essays/essay_cvad. [Πρόσβαση: 5 Μαΐου 2013].

[15] W. Piekarski, «Interactive 3d Modelling in Outdoor Augmented Reality Worlds», Ph.D. Dissertation, University of South Australia, School of Computer and Information Science, Division of Information Technology, Engineering, and the Environment, Adelaide, South Australia, February 2004.

[16] D. Dixon. (Φεβρουάριος 2010). «Augmented Reality Goes Mobile», *Manifest Technology* [Ηλεκτρονικό]. Διαθέσιμο: www.manifest-tech.com/society/augmented_reality.htm. [Πρόσβαση: 5 Μαΐου 2013].

[17] T. Mullen, «Getting Started with Augmented Reality» in *Prototyping Augmented Reality*, Indianapolis, Indiana, John Wiley & Sons, Inc., 2011, pp. 1-12.

[18] L. Rosenberg, «The Use of Virtual Fixtures as Perceptual Overlays to Enhance Operator Performance in Remote Environments», Air Force Material Command, Wright-Patterson Air Force Base, Ohio, 1992.

[19] S. Feiner, B. MacIntyre and D. Seligmann. *KARMA: Knowledge-based Augmented Reality for Maintenance Assistance* [Ηλεκτρονικό]. Διαθέσιμο: monet.cs.columbia.edu/projects/karma/karma.html. [Πρόσβαση: 5 Μαίου 2013].

[20] «Augmented Reality», *CSISWiki* [Ηλεκτρονικό]. Διαθέσιμο: wiki.csisdmz.ul.ie/wiki/Augmented_reality. [Πρόσβαση: 6 Μαΐου 2013].

[21] «History of Mobile Augmented Reality», *Institute for Computer Graphics and Vision (ICG)* [Ηλεκτρονικό]. Διαθέσιμο**:** https://www.icg.tugraz.at/~daniel/HistoryOfMobileAR. [Πρόσβαση: 6 Μαΐου 2013].

[22] *Mobile AR History* [Ηλεκτρονικό]. Διαθέσιμο: https://www.google.gr/url?sa=t&rct=j&q=&esrc=s&source=web&cd=3&ved=0CEMQFjAC&url=https%3A%2F%2Fnoppa.aalto.fi%2Fnoppa%2Fkurssi%2Ft-111.5700%2Fluennot%2FT-111_5700_lecture_slides_-

_part_2.pdf&ei=xQQbUsjoEc-2hAf7-YDYBQ&usg=AFQjCNFU3X3STXpHRCLmBKY875mS2c5BJQ. [Πρόσβαση: 6 Μαΐου 2013].

[23] J. Rekimoto. «CybernеticHuman: From Human Computer Interaction to Human Computer Integration», *Microsoft* [Ηλεκτρονικό]. Διαθέσιμο: http://research.microsoft.com/en-us/events/chinahci2010/junrekimoto.pdf. [Πρόσβαση: 6 Μαΐου 2013].

[24] J. Rekimoto and Y. Ayatsuka, «CyberCode: Designing Augmented Reality Environments with Visual Tags», Sony Computer Science Laboratories Inc., Tokyo.

[25] B. Thomas, D. Victor, P. Wayne, H. David και G. Bernard, «A Wearable Computer System with Augmented Reality to Support Terrestrial Navigation», in *Proc. 2nd International Symposium on Wearable Computers*, Pittsburgh, PA, USA, 1998.

[26] *Total Immersion* [Ηλεκτρονικό]. Διαθέσιμο: https://community.t-immersion.com. [Πρόσβαση: 7 Μαΐου 2013].

[27] *HitLab - ARToolKit* [Ηλεκτρονικό]. Διαθέσιμο: http://www.hitl.washington.edu/artoolkit. [Πρόσβαση: 7 Μαΐου 2013].

[28] *Wikipedia, The Free Encyclopedia* [Ηλεκτρονικό]. Διαθέσιμο: http://en.wikipedia.org/wiki/ARToolKit. [Πρόσβαση: 7 Μαΐου 2013].

[29] «First Person Indoor/Outdoor Augmented Reality Application: ARQuake», *Personal and Ubiquitous Computing,* vol. 6, no. 1, pp. 75-86, 2002.

[30] R. Kooper and B. MacIntyre, «Browsing the Real-World Wide Web Browser: Maintaining Awareness of Virtual Information in an AR Information Space», *International Journal of Human - Computer Interaction,* vol. 16, no. 3, pp. 425-446, 2003.

[31] *ISMAR 2002* [Ηλεκτρονικό]. Διαθέσιμο: http://ismar2002.vgtc.org. [Πρόσβαση: 7 Μαΐου 2013].

[32] *Wikipedia, The Free Encyclopedia* [Ηλεκτρονικό]. Διαθέσιμο: http://en.wikipedia.org/wiki/International_Symposium_on_Mixed_and_Augmented_Reality. [Πρόσβαση: 7 Μαΐου 2013].

[33] R. Raskar, J. Van Baar, P. Beardsley, T. Willwacher, S. Rao and C. Forlines, «iLamps: Geometrically Aware and Self-Configuring Projectors», in *Proc. ACM SIGGRAPH Conference*, San Diego, 2003.

[34] M. Möhring, C. Lessig and B. Oliver, «Video See-Through AR on Consumer Cell-Phones», in *Proc. 3d IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR 04)*, Arlington, VA, USA, 2004.

[35] «The Invisible Train», *Institute for Computer Graphics and Vision (ICG)* [Ηλεκτρονικό]. Διαθέσιμο: http://studierstube.icg.tugraz.at/invisible_train. [Πρόσβαση: 8 Μαΐου 2013].

[36] *Nokia Research Center* [Ηλεκτρονικό]. Διαθέσιμο: http://research.nokia.com/page/219. [Πρόσβαση: 8 Μαΐου 2013].

[37] «Nokia Mara», *TechBlog* [Ηλεκτρονικό]. Διαθέσιμο: http://techblog.gr/techno/nokia-mara. [Πρόσβαση: 8 Μαΐου 2013].

[38] «Augmented Reality at Wellington Zoo», *The Inspiration Room* [Ηλεκτρονικό]. Διαθέσιμο: http://theinspirationroom.com/daily/2007/augmented-reality-at-wellington-zoo. [Πρόσβαση: 8 Μαΐου 2013].

[39] *Wikipedia, The Free Encyclopedia* [Ηλεκτρονικό]. Διαθέσιμο: http://en.wikipedia.org/wiki/Wikitude. [Πρόσβαση: 8 Μαΐου 2013].

[40] *Layar* [Ηλεκτρονικό]. Διαθέσιμο: http://www.layar.com/what-is-layar. [Πρόσβαση: 8 Μαΐου 2013].

[41] B. Goldiez, «Techniques for Assessing and Imroving Performance in Navigation and Wayfinding Using Mobile Augmented Reality», Ph.D. Dissertation, University of Central Florida, College of Arts and Sciences, Orlando, 2004.

[42] *Medical Augmented Reality Blog* [Ηλεκτρονικό]. Διαθέσιμο: http://medicalaugmentedreality.com/201107medical-augmented-reality-blog. [Πρόσβαση: 8 Μαΐου 2013].

[43] *MPEG AR* [Ηλεκτρονικό]. Διαθέσιμο: http://wg11.sc29.org/augmentedReality/?page_id=1154. [Πρόσβαση: 8 Μαΐου 2013].

[44] R. Behringer, «Augmented Reality», in *Encyclopedia of Computer Science and Technology*, vol. 45, no. 3, Allen Kent - James G. Williams, 2002, pp. 45-57.

[45] Ι. Λεβέντη, «Υπηρεσίες Βάσει Θέσης και Επαυξημένη Πραγματικότητα», Μεταπτυχιακή Εργασία, Εθνικό Μετσόβιο Πολυτεχνείο, Σχολή Αγρονόμων και Τοπογράφων Μηχανικών, Μεταπτυχιακό Πρόγραμμα Γεωπληροφορική, Αθήνα, 2012.

[46] *Defense Indystry Daily* [Ηλεκτρονικό]. Διαθέσιμο: http://www.defenseindustrydaily.com/First-Production-HMDS-Orders-for-F-35-Pilots-05549. [Πρόσβαση: 10 Μαΐου 2013].

[47] *University of CanterBury - Library Blogs* [Ηλεκτρονικό]. Διαθέσιμο: http://blogs.libr.canterbury.ac.nz/libnews.php?itemid=8423. [Πρόσβαση: 9 Μαΐου 2013].

[48] *Wikipedia, The Free Encyclopedia* [Ηλεκτρονικό]. Διαθέσιμο: http://en.wikipedia.org/wiki/1st_%26_Ten_(graphics_system). [Πρόσβαση: 9 Μαΐου 2013].

[49] *Turbulence* [Ηλεκτρονικό]. Διαθέσιμο: http://turbulence.org/blog/2010/10/03/live-stage-we-ar-in-moma-nyc. [Πρόσβαση: 9 Μαΐου 2013].

[50] *Konstruct* [Ηλεκτρονικό]. Διαθέσιμο: http://apps.augmatic.co.uk/konstruct. [Πρόσβαση: 9 Μαΐου 2013].

[51] J. Pair, J. Wilson, J. Chastine and M. Gandy, «The Duran Duran Project: The Augmented Reality Toolkit in Live Performance», in *Proc. 1st IEEE International Augmented Reality Toolkit Workshop*, Darmstadt, Germany, 2002.

[52] V. Vlahakis, N. Ioannidis, J. Karigiannis, M. Tsotros, M. Gounaris, D. Stricker, T. Gleue, P. Daehne and L. Almeida, «Archeoguide: An Augmented Reality Guide for Archeological Sites», *IEEE Computer Graphics and Applications*, vol. 22, no. 5, pp. 52-60, 2002.

[53] *Layar* [Ηλεκτρονικό]. Διαθέσιμο: http://www.layar.com/blog/2012/11/21/nissan-campaign-top-advertising-award. [Πρόσβαση: 10 Μαΐου 2013].

[54] *Wikipedia, The Free Encyclopedia* [Ηλεκτρονικό]. Διαθέσιμο: https://en.wikipedia.org/wiki/Head-mounted_display. [Πρόσβαση: 11 Μαΐου 2013].

[55] *Wikipedia, The Free Encyclopedia* [Ηλεκτρονικό]. Διαθέσιμο: https://en.wikipedia.org/wiki/Helmet_mounted_display. [Πρόσβαση: 11 Μαΐου 2013].

[56] O. Bilber and R. Raskar, *Modern Approaches to Augmented Reality*, Massachusetts: Mitsubishi Electric Research Laboratories Inc, 2006.

[57] J. Rolland, F. Biocca, F. Hamza-Lup, Y. Ha and R. Martins, «Development of Head-Mounted Projection Displays for Distributed, Collaborative, Augmented Reality Applications», *Presence,* vol. 14, no. 5, pp. 528-549, 2005.

[58] H. Hua, A. Girardot, C. Gao and J. Rolland, «Engineering of head-mounted projective displays», *APPLIED OPTICS,* vol. 39, no. 22, pp. 3814-3824, 2000.

[59] *Wikipedia, The Free Encyclopedia* [Ηλεκτρονικό]. Διαθέσιμο: http://en.wikipedia.org/wiki/Virtual_retinal_display. [Πρόσβαση: 11 Μαΐου 2013].

[60] D. Wolk. «Future of the Screen: Terminator-Style Augmented-Reality Glasses», *Wired* [Ηλεκτρονικό]. Διαθέσιμο: http://www.wired.com/dualperspectives/article/news/2009/09/dp_displays_wired0901. [Πρόσβαση: 11 Μαΐου 2013].

[61] *Wikipedia, The Free Encyclopedia* [Ηλεκτρονικό]. Διαθέσιμο: http://en.wikipedia.org/wiki/EyeTap. [Πρόσβαση: 12 Μαΐου 2013].

[62] G. R. Briones. (17 Ιουλίου 2012). *Ubergizmo* [Ηλεκτρονικό]. Διαθέσιμο: http://www.ubergizmo.com/2012/07/eyetap-director-allegedly-assaulted-by-mcdonald-employees. [Πρόσβαση: 12 Μαΐου 2013].

[63] *Wikipedia, The Free Encyclopedia* [Ηλεκτρονικό]. Διαθέσιμο: http://en.wikipedia.org/wiki/Google_Glass. [Πρόσβαση: 11 Μαΐου 2013].

[64] *Vuzix* [Ηλεκτρονικό]. Διαθέσιμο: http://www.vuzix.com/augmented-reality/docs/SMART-Glasses-Brochure.pdf. [Πρόσβαση: 11 Μαΐου 2013].

[65] *Wikipedia, The Free Encyclopedia* [Ηλεκτρονικό]. Διαθέσιμο: http://en.wikipedia.org/wiki/Bionic_contact_lens. [Πρόσβαση: 12 Μαΐου 2013].

[66] B. Parviz. (1 Σεπτεμβρίου 2009). «Augmented Reality in a Contact Lens», *IEEE Spectrum* [Ηλεκτρονικό]. Διαθέσιμο: http://spectrum.ieee.org/biomedical/bionics/augmented-reality-in-a-contact-lens. [Πρόσβαση: 12 Μαΐου 2013].

[67] «Breakthrough in Augmented Reality Contact Lens», *Science Daily* [Ηλεκτρονικό]. Διαθέσιμο: http://www.sciencedaily.com/releases/2012/12/121205090931.htm. [Πρόσβαση: 12 Μαΐου 2013].

[68] D. Wagner, «Handheld Augmented Reality», Dissertation, Graz University of Technology, Institute for Computer Graphics and Vision, Graz, Austria, 2007.

[69] D. L. Baggio, E. Shervin, D. M. Escrivá, K. Ievgen, N. Mahmood, J. Saragih and R. Shilkrot, *Mastering OpenCV with Practical Computer Vision Problems*, Birmingham: Packt Publishing, 2012.

[70] A. Olwal and A. D. Wilson, «SurfaceFusion: Unobtrusive Tracking of Everyday Objects in Tangible User Interfaces», in *Proc. 34th Canadian Graphics Interface Conference*, Windsor, Canada, 2008.

[71] *Microsoft Research* [Ηλεκτρονικό]. Διαθέσιμο: http://research.microsoft.com/en-us/projects/lightspace. [Πρόσβαση: 14 Μαΐου 2013].

[72] Δ. Τζιώνας, «Ευέλικτοι Αλγόριθμοι Ανίχνευσης Markers σε Εικόνα - Εφαρμογή στην Επαυξημένη Πραγματικότητα», Διπλωματική Εργασία, Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης, Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών, Θεσσαλονίκη, 2009.

[73] U. Neumann, S. You, Y. Cho, J. Lee J. Park, «Natural Feature Tracking for Augmented Reality», in *Proc. International Symposium on Mixed Realities*, Tokyo, Japan, 1999.

[74] Γ. Καρράς, «Διαφάνειες Μαθήματος Φωτογραμμετρία ΙΙΙ», Εθνικό Μετσόβιο Πολυτεχνείο, Σχολή Αγρονόμων και Τοπογράφων Μηχανικών, Αθήνα, 2012.

[75] R. Szeliski, *Computer Vision: Algorithms and Applications*, London: Springer, 2011.

[76] *Wikipedia, The Free Encyclopedia* [Ηλεκτρονικό]. Διαθέσιμο: http://en.wikipedia.org/wiki/Feature_detection_(computer_vision). [Πρόσβαση: 1 Ιουνίου 2013].

[77] M. S. Koul, «Computer Vision: Edge Detection», Lecture Slides, The University of Texas at Arlington, Department of Electrical Engineering, Texas.

[78] Σ. Φωτόπουλος. 2005. «Ανίχνευση Ακμών - Περιγραμμάτων», *Πανεπιστήμιο Πατρών, ΔΠΜΣ Ηλεκτρονική και Επεξεργασία της Πληροφορίας* [Ηλεκτρονικό]. Διαθέσιμο: http://www.hep.upatras.gr/class/download/psi_epe_iko/kef4.pdf. [Πρόσβαση: 2 Ιουνίου 2013].

[79] J. Canny, «A Computational Approach to Edge Detection», *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 8, no. 6, pp. 679-698, 1986.

[80] Ν. Τσαπατσούλης. 2005. «Ψηφιακή Επεξεργασία Εικόνας: Κατάτμηση Εικόνων: Ανίχνευση Ακμών και Κατάτμηση με Κατωφλίωση», *Πανεπιστήμιο Πειραιά, Τμήμα Διδακτικής της Τεχνολογίας και Ψηφιακών Συστημάτων* [Ηλεκτρονικό]. Διαθέσιμο: http://www2.cs.ucy.ac.cy/~nicolast/courses/teds150/lectures/IP10.pdf. [Πρόσβαση: 2 Ιουνίου 2013].

[81] W. Förstner, «A Feature Based Correspondence Algorithm for Image Matching», *International Archives of Photogrammetry,* vol. 26, no. 3, pp. 150-166, 1986.

[82] K. Grauman and B. Leibe (2011), *Synthesis Lectures on Computer Vision: Visual Object Recognition* [Ηλεκτρονικό], Διαθέσιμο: http://cs.gmu.edu/~kosecka/cs482/grauman-recognition-draft-27-01-11.pdf. [Πρόσβαση: 19 Μαϊου, 2013].

[83] Κ. Αδάμ, «Γεωμετρία του στερεοζεύγους από βαθμονομημένες και από μη βαθμονομημένες μηχανές», Διπλωματική Εργασία, Εθνικό Μετσόβιο Πολυτεχνείο, Σχολή Αγρονόμων και Τοπογράφων Μηχανικών, Αθήνα, 2011.

[84] T. Tuytelaars, «Local Invariant Features: What? Why? When? How?», Tutorial ECCV, Slides, 2006.

[85] H. Bay, T. Tuytelaars and L. Van Gool, «Speeded-Up Robust Features (SURF)», in *Proc. 9th European Conference on Computer Vision*, Graz, Austria, 2006.

[86] H. Bay, A. Ess, T. Tuytelaars and L. Van Gool, «Speeded-Up Robust Features (SURF)», *Computer Vision and Image Understanding,* vol. 110, no. 3, pp. 346-359, 2008.

[87] A. Xu and G. Namit, «SURF: Speeded-Up Robust Features», COMP 558 - Project Report, McGill University, Montreal, Quebec, Canada, 2008.

[88] C. Carson, S. Belongie, H. Greenspan and J. Malik, «Region-Based Image Querying», in *Proc. Workshop on Content-Based Access of Image and Video Libraries*, Washington, DC, USA, 1997.

[89] *Wikipedia, The Free Encylopedia* [Ηλεκτρονικό]. Διαθέσιμο: https://en.wikipedia.org/wiki/Blob_detection. [Πρόσβαση: 21 Μαΐου 2013].

[90] P. Viola and M. Jones, «Rapid Object Detection using a Boosted Cascade of Simple Features», in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Kauai, HI, USA, 2001.

[91] *Wikipedia, The Free Encyclopedia* [Ηλεκτρονικό]. Διαθέσιμο: http://en.wikipedia.org/wiki/Hessian_matrix. [Πρόσβαση: 20 Μαΐου 2013].

[92] *Wikipedia, The Free Encyclopedia* [Ηλεκτρονικό]. Διαθέσιμο: http://en.wikipedia.org/wiki/Gaussian_blur. [Πρόσβαση: 20 Μαΐου 2013].

[93] D. Lowe, «Distinctive Image Features from Scale-Invariant Keypoints», *International Journal of Computer Vision,* vol. 60, no. 2, pp. 91-110, 2004.

[94] Η. Καλησπεράκης, «Η Επιπολική Γεωμετρία στον Προβολικό και τον Ευκλείδειο Χώρο», Διδακτορική Διατριβή, Εθνικό Μετσόβιο Πολυτεχνείο, Σχολή Αγρονόμων και Τοπογράφων Μηχανικών, Αθήνα, 2010.

[95] C. Evans. (18 Ιανουαρίου 2009). *Notes on the OpenSURF Library* [Ηλεκτρονικό]. Διαθέσιμο: http://www.cs.bris.ac.uk/Publications/Papers/2000970.pdf [Πρόσβαση: 5 Ιουνίου 2013].

[96] S. Ameling. *An Introduction to Local Feature Extraction, Description and Tracking in Matching Applications* [Ηλεκτρονικό]. Διαθέσιμο: ftp://tracking.mat.ucsb.edu/projects/external_docs/tracking/intro_to_matching.pdf [Πρόσβαση: 5 Ιουνίου 2013].

[97] A. Neubeck and L. Van Gool, «Efficient Non-Maximum Suppression», in *Proc. 18th International Conference on Pattern Recognition*, Hong Kong, China, 2006.

[98] S. Smith. (15 Μαρτίου 2011). «Speeded-Up Robust Features: SURF», *University of Utah - Scientific Computing and Imaging Institute* [Ηλεκτρονικό]. Διαθέσιμο: http://www.sci.utah.edu/~fletcher/CS7960/slides/Scott.pdf. [Πρόσβαση: 3 Αυγούστου 2013].

[99] M. Brown and D. Lowe, «Invariant Features from Interest Point Groups», in *Proc. 13th British Machine Vision Conference*, Cardiff, UK, 2002.

[100] E. Oyallon and R. Julien, «An analysis and implementation of the SURF method, and its comparison to SIFT», *IPOL Journal - Image Processing On Line*, 2013.

[101] D. Fleet and A. Jepson, «Feature Descriptors, Detection and Matching», Slides, University of Toronto, Computer Science, Toronto, 2011.

[102] D. Poole, Linear Algebra: A Modern Introduction, Canada: Brooks/Cole Cengage Learning, 2011.

[103] M. Fischler and R. Bolles, «Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography», *Communications of the ACM*, vol. 24, no. 6, pp. 381-395, 1981.

[104] J. C. McGlone, E. M. Mikhail and J. Bethel, *Manual of Photogrammetry*, 5th ed., USA: American Society for Photogrammetry and Remote Sensing (ASPRS), 2004.

[105] *Wikipedia, The Free Encyclopedia* [Ηλεκτρονικό]. Διαθέσιμο: http://en.wikipedia.org/wiki/RANSAC. [Πρόσβαση: 4 Ιουνίου 2013].

[106] K. Derpanis. (13 Μαΐου 2010). «Overview of the RANSAC Algorithm», *YORK Univiersity, Department of Computer Science and Engineering* [Ηλεκτρονικό]. Διαθέσιμο: http://www.cse.yorku.ca/~kosta/CompVis_Notes/ransac.pdf. [Πρόσβαση: 4 Ιουνίου 2013].

[107] Λ. Γραμματικόπουλος, «Γεωμετρική Πληροφορία από Μεμονωμένες Εικόνες στην Φωτογραμμετρία και την Όραση Υπολογιστών», Διδακτορική Διατριβή, Εθνικό Μετσόβιο Πολυτεχνείο, Σχολή Αγρονόμων και Τοπογράφων Μηχανικών, Αθήνα, 2007.

[108] *Wikipedia, The Free Encyclopedia* [Ηλεκτρονικό]. Διαθέσιμο: http://en.wikipedia.org/wiki/Projective_geometry. [Πρόσβαση: 10 Ιουνίου 2013].

[109] *Wikipedia, The Free Encyclopedia* [Ηλεκτρονικό]. Διαθέσιμο: https://en.wikipedia.org/wiki/Homogeneous_coordinates. [Πρόσβαση: 4 Ιουνίου 2013].

[110] J. Bloomenthal and J. Rokne, «Homogeneous Coordinates», *The Visual Computer*, vol. 11, no. 1, pp. 15-26, 1994.

[111] S. Prince, *Computer Vision: Models, Learning and Inference*, New York: Cambridge University Press, 2012.

[112] «Homogeneous Coordinates», *Michigan Tech, Department of Computer Science* [Ηλεκτρονικό]. Διαθέσιμο: http://www.cs.mtu.edu/~shene/COURSES/cs3621/NOTES/geometry/homo-coor.html. [Πρόσβαση: 4 Ιουνίου 2013].

[113] Α. Γεωργόπουλος, *Ψηφιακή Φωτογραμμετρία*, Αθήνα: Εθνικό Μετσόβιο Πολυτεχνείο, Σχολή Αγρονόμων και Τοπογράφων Μηχανικών, 1998.

[114] Ε. Πέτσα, *Θεμελιώδεις Έννοιες και Θεμελιώδη Προβλήματα της Φωτογραμμετρίας*, Αθήνα: Τ.Ε.Ι. Αθήνας, Τμήμα Τοπογραφίας, 2000.

[115] Π. Πατιάς, *Εισαγωγή στη Φωτογραμμετρία*, Θεσσαλονίκη: Εκδόσεις Ζήτη, 1991.

[116] D. Forsyth and J. Ponce, *Computer Vision: A Modern Approach*, New Jersey: Pearson, 2012.

[117] Χ. Ιωαννίδης, «Διαφάνειες Διαλέξεων Φωτογραμμετρίας Ι», Εθνικό Μετσόβιο Πολυτεχνείο, Σχολή Αγρονόμων και Τοπογράφων Μηχανικών, Αθήνα, 2010.

[118] D. Brown, «Decentering Distortion of Lenses», *Photogrammetric Engineering*, vol. 32, no. 3, pp. 444-462, 1966.

[119] J. G. Fryer, «Lens Distortion for Close Range Photogrammetry», in *Proc. Symposium Real-Time Photogrammetry - A New Challenge*, Ottawa, Canada, 1986.

[120] Α. Γεωργόπουλος, Χ. Ιωαννίδης, Ε. Λάμπρου, Γ. Πανταζής και Κ. Νικολίτσας, «Πρόταση Μεθοδολογίας Βαθμονόμησης της Ψηφιακής Κάμερας Εικονογεωδαιτικού Σταθμού», σε *Πρακτικά 3ου Τακτικού Εθνικού Συνεδρίου Μετρολογίας*, Λάρνακα, Κύπρος, 2010.

[121] «The Pinhole Camera», *Donald Bren School University of Information and Computer Sciene, University of California* [Ηλεκτρονικό]. Διαθέσιμο: http://www.ics.uci.edu/~majumder/vispercep/cameracalib.pdf. [Πρόσβαση: 8 Ιουνίου 2013].

[122] R. Hartley and A. Zisserman. (1999). «Multiple View Geometry», *ANU College of Engineering & Computer Science* [Ηλεκτρονικό]. Διαθέσιμο: http://users.cecs.anu.edu.au/~hartley/Papers/CVPR99-tutorial/tut_4up.pdf. [Πρόσβαση: 8 Ιουνίου 2013].

[123] B. Fisher. (16 Απριλίου 1997). «3x4 Projection Matrix», *The University of Edinburgh - School of Informatics* [Ηλεκτρονικό]. Διαθέσιμο: http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/EPSRC_SSAZ/node3.html. [Πρόσβαση: 8 Ιουνίου 2013].

[124] R. Laganière, *OpenCV 2 Computer Vision Application Programming Cookbook*, Birmingham, UK: Packt Publishing Ltd, 2011.

[125] *OpenCV Documentation - Camera Calibration and 3d Reconstruction* [Ηλεκτρονικό]. Διαθέσιμο: http://opencv.willowgarage.com/documentation/cpp/camera_calibration_and_3d_reconstruction.html. [Πρόσβαση: 9 Ιουνίου 2013].

[126] Γ. Καρράς, *Γραμμικοί Μετασχηματισμοί Συντεταγμένων στην Φωτογραμμετρία*, Αθήνα: Εθνικό Μετσόβιο Πολυτεχνείο, Σχολή Αγρονόμων και Τοπογράφων Μηχανικών, 1998.

[127] C. Zhang and W. Yao, «The Comparisons of 3D Analysis Between Photogrammetry and Computer Vision», *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 37, part B3b, pp. 33-36, 2008.

[128] B. Horn. (1999). «Projective Geometry Considered Harmful», *MIT Computer Science and Artificial Intelligence Laboratory* [Ηλεκτρονικό]. Διαθέσιμο: http://people.csail.mit.edu/bkph/articles/Harmful.pdf. [Πρόσβαση: 7 Ιουνίου 2013].

[129] *Wikipedia, The Free Encyclopedia* [Ηλεκτρονικό]. Διαθέσιμο: http://en.wikipedia.org/wiki/Homography. [Πρόσβαση: 7 Ιουνίου 2013].

[130] «Introduction to Homography», *Slideshare* [Ηλεκτρονικό]. Διαθέσιμο: http://www.slideshare.net/mimimau/introduction-to-homography. [Πρόσβαση: 7 Ιουνίου 2013].

[131] E. Dubrofsky, «Homography Estimation», Master's Thesis, The University of British Columbia, Vancouver, Canada, 2009.

[132] D. Kriegman. (2007). «Homography Estimation», *UCSD Computer Science and Engineering* [Ηλεκτρονικό]. Διαθέσιμο: http://cseweb.ucsd.edu/classes/wi07/cse252a/homography_estimation/homography_estimation.pdf. [Πρόσβαση: 8 Ιουνίου 2013].

[133] Ι. Φιλιόπουλος, «Σχεδιασμός, Ίδρυση και Μέτρηση Πεδίου Ελέγχου για τη Βαθμονόμηση Μηχανών Αεροφωτογράφησης», Διπλωματική Εργασία, Εθνικό Μετσόβιο Πολυτεχνείο, Σχολή Αγρονόμων και Τοπογράφων Μηχανικών, Αθήνα, 2010.

[134] Π. Αγραφιώτης και Α. Γεωργόπουλος, «Αξιολόγηση και Σύγκριση Μεθόδων Βαθμονόμησης Ψηφιακών Μηχανών», σε *Πρακτικά 4ου Τακτικού Εθνικού Συνεδρίου Μετρολογίας*, Αθήνα, 2012.

[135] F. Remondino and C. Fraser, «Digital Camera Calibration Methods: Considerations and Comparisons», in *Proc. ISPRS Commission V Symposium Image Engineering and Vision Metrology*, Dresden, 2006.

[136] *OpenCV Documentation - Eroding and Dilating* [Ηλεκτρονικό]. Διαθέσιμο: http://docs.opencv.org/doc/tutorials/imgproc/erosion_dilatation/erosion_dilatation.html. [Πρόσβαση: 24 Ιουλίου 2013].

[137] S. Suzuki and K. Abe, «Topological Structural Analysis of Digitized Binary Images by Border Following», *Computer Vision, Graphics and Image Processing*, vol. 30, no. 1, pp. 32-46, 1985.

[138] «Defining Connectivity», *ImageProcessingPlace.com* [Ηλεκτρονικό]. Διαθέσιμο: http://www.imageprocessingplace.com/downloads_V3/root_downloads/tutorials/con

tour_tracing_Abeer_George_Ghuneim/connectivity.html. [Πρόσβαση: 29 Ιουλίου 2013].

[139] «Image Segmentation», *Computer Sciences* [Ηλεκτρονικό]. Διαθέσιμο: http://pages.cs.wisc.edu/~dyer/cs766/slides/segmentation/segment-2up.pdf. [Πρόσβαση: 29 Ιουλίου 2013].

[140] *OpenCV Source Code - checkchessboard.cpp* [Ηλεκτρονικό]. Διαθέσιμο: https://github.com/Itseez/opencv/blob/master/modules/calib3d/src/checkchessboard.cpp. [Πρόσβαση: 9 Ιουνίου 2013].

[141] Χ. Ζούζουλας και Β. Θεοδώρου, «Ανίχνευση Ίριδας και Εφαρμογές», Πτυχιακή εργασία, Τεχνολογικό Εκπαιδευτικό Ίδρυμα Κρήτης, Σχολή Τεχνολογικών Εφαρμογών, Τμήμα Εφαρμοσμένης Πληροφορικής & Πολυμέσων, Ηράκλειο, Κρήτη, 2011.

[142] *OpenCV Source Code - contours.cpp* [Ηλεκτρονικό]. Διαθέσιμο: https://github.com/Itseez/opencv/blob/master/modules/imgproc/src/contous.cpp. [Πρόσβαση: 25 Ιουλίου 2010].

[143] *OpenCV Documentation - Feauture Detection* [Ηλεκτρονικό]. Διαθέσιμο: http://docs.opencv.org/modules/imgproc/doc/feature_detection.html. [Πρόσβαση: 26 Ιουλίου 2013].

[144] *OpenCV source code - calibinit.cpp* [Ηλεκτρονικό]. Διαθέσιμο: https://github.com/Itseez/opencv/blob/master/modules/calib3d/src/calibinit.cpp. [Πρόσβαση: 9 Ιουνίου 2013].

[145] V. Douskos, I. Kalisperakis and G. Karras, «Automatic calibration of digital cameras using planar chess-board patterns», *Optical 3-D Measurement Techniques VIII, Wichman,* vol. 1, pp. 132-140, 2007.

[146] L. Grammatikopoulos, G. Karras and E. Petsa, «Camera Calibration Combining Images with Two Vanishing Points», *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences,* vol. 35, part B5, pp. 99-104, 2004.

[147] *OpenCV Source Code - calibration.cpp* [Ηλεκτρονικό]. Διαθέσιμο: https://github.com/Itseez/opencv/blob/master/modules/calib3d/src/calibration.cpp. [Πρόσβαση: 9 Ιουνίου 2013].

[148] Z. Zhang, «A flexible new technique for camera calibration», *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 22, no. 11, pp. 1330-1334, 2000.

[149] *OpenCV Source Code - fundam.cpp* [Ηλεκτρονικό]. Διαθέσιμο: https://github.com/Itseez/opencv/blob/master/modules/calib3d/src/fundam.cpp. [Πρόσβαση: 29 Ιουλίου 2013].

[150] *Wikipedia, The Free Encyclopedia* [Ηλεκτρονικό]. Διαθέσιμο: http://en.wikipedia.org/wiki/Levenberg%E2%80%93Marquardt_algorithm. [Πρόσβαση: 29 Ιουλίου 2013].

[151] G. Bradski and A. Kaehler, *Learning OpenCV*, USA: O'Reilly Media, 2008.

[152] Q. Wu, G. XU and L. Wang, «A Three-stage System for Camera Calibration», *SPIE Proceedings: Visualization and Optimization Techniques,* vol. 4553, no. 1, pp. 111-115, 2001.

[153] «Camera Calibration - Homography», *Brigham Young University - Robotic Vision Lab* [Ηλεκτρονικό]. Διαθέσιμο: http://ece631web.groups.et.byu.net/Lectures/ECEn631%2010%20-%20Homography%20Camera%20Calibration.pdf. [Πρόσβαση: 28 Ιουλίου 2013].

[154] Α. Σέχου, «Ρωμαλέες - Χαμηλής Πολυπλοκότητας Τεχνικές Εκτίμησης Θέσης και Προσανατολισμού Κάμερας», Μεταπτυχιακή Εργασία, Πανεπιστήμιο Πατρών, Τμήμα Μηχανικών Ηλεκτρονικών Υπολογιστών και Πληροφορικής, Πάτρα, 2012.

[155] *Wikipedia, The Free Encyclopedia* [Ηλεκτρονικό]. Διαθέσιμο: http://en.wikipedia.org/wiki/C%2B%2B. [Πρόσβαση: 14 Ιουνίου 2013].

[156] Θ. Πιερράτος. «Ανάπτυξη Εφαρμογών σε Προγραμματιστικό Περιβάλλον - Ανασκόπηση Θεωρίας - 6ο Κεφάλαιο - Εισαγωγή στον Προγραμματισμό», *Πανελλήνιο Σχολικό Δίκτυο* [Ηλεκτρονικό]. Διαθέσιμο: http://users.sch.gr/pierratos/anaptixi/theoria6.pdf. [Πρόσβαση: 14 Ιουνίου 2013].

[157] H. M. Deitel και P. Deitel, *C++ Προγραμματισμός*, 4η έκδ., Αθήνα: Μ. Γκιούρδας, 2003.

[158] B. Stroustrup. «A History of C++: 1979-1991», *Bjarne Stroustrup's homepage* [Ηλεκτρονικό]. Διαθέσιμο: http://www.stroustrup.com/hopl2.pdf. [Πρόσβαση: 14 Ιουνίου 2013].

[159] «History of C++», *cplusplus.com* [Ηλεκτρονικό]. Διαθέσιμο: http://www.cplusplus.com/info/history. [Πρόσβαση: 14 Ιουνίου 2013].

[160] «A Brief History of AR», *MathBits.com* [Ηλεκτρονικό]. Διαθέσιμο: http://mathbits.com/MathBits/CompSci/Introduction/history.htm. [Πρόσβαση: 14 Ιουνίου 2013].

[161] Κ. Μαργαρίτης και Γ. Τσιομπίκας. (2004). *Εισαγωγή στη C++* [Ηλεκτρονικό]. Διαθέσιμο: http://www.ebooks4greeks.gr/forum/viewtopic.php?p=260. [Πρόσβαση: 14 Ιουνίου 2013].

[162] Σ. Βερυκόκου, «Ανάπτυξη Εφαρμογής σε Λειτουργικό Σύστημα Android», Εργασία Εξαμήνου στο Μάθημα - Θέμα: Εφαρμογές Διασυνδεδεμένων Ψηφιακών Συστημάτων, Εθνικό Μετσόβιο Πολυτεχνείο, Σχολή Αγρονόμων και Τοπογράφων Μηχανικών, Αθήνα, 2012.

[163] «Αντικειμενοστρεφής Προγραμματισμός», *University of Cyprus - Department of Computer Science* [Ηλεκτρονικό]. Διαθέσιμο: http://www.cs.ucy.ac.cy/~mourlas/epl132/notes/notes19-21.pdf. [Πρόσβαση: 14 Ιουνίου 2013].

[164] *Βικιπαίδεια, Η Ελεύθερη Εγκυκλοπαίδεια* [Ηλεκτρονικό]. Διαθέσιμο: http://el.wikipedia.org/wiki/Αντικειμενοστρεφής_προγραμματισμός. [Πρόσβαση: 14 Ιουνίου 2013].

[165] «Compiler, Assembler, Linker and Loader: A Brief History», *Tenouk* [Ηλεκτρονικό]. Διαθέσιμο: http://www.tenouk.com/ModuleW.html. [Πρόσβαση: 14 Ιουνίου 2013].

[166] «Headers and Includes: Why and How», *cplusplus.com* [Ηλεκτρονικό]. Διαθέσιμο: http://www.cplusplus.com/forum/articles/10627. [Πρόσβαση: 14 Ιουνίου 2013].

[167] *Wikipedia, The Free Encyclopedia* [Ηλεκτρονικό]. Διαθέσιμο: http://en.wikipedia.org/wiki/Microsoft_Visual_Studio_Express. [Πρόσβαση: 14 Ιουνίου 2013].

[168] *Wikipedia, The Free Encyclopedia* [Ηλεκτρονικό]. Διαθέσιμο: http://en.wikipedia.org/wiki/Visual_C%2B%2B. [Πρόσβαση: 14 Ιουνίου 2013].

[169] Ι. Γεωργουδάκης και Π. Μαστοροκώστας. (2011). «Συνοπτικό Εγχειρίδιο Χρήσης του Microsoft Visual Studio 2010», *Τ.Ε.Ι. Κεντρικής Μακεδονίας - Σέρρες* [Ηλεκτρονικό]. Διαθέσιμο: http://www.teiser.gr/icd/staff/mastorokostas/official/COURSES/PROGRAMMING_I/default.htm. [Πρόσβαση: 14 Ιουνίου 2013].

[170] «OpenCV», *Willow Garage* [Ηλεκτρονικό]. Διαθέσιμο: http://www.willowgarage.com/pages/software/opencv. [Πρόσβαση: 15 Ιουνίου 2013].

[171] *Wikipedia, The Free Encyclopedia* [Ηλεκτρονικό]. Διαθέσιμο: http://en.wikipedia.org/wiki/OpenCV. [Πρόσβαση: 15 Ιουνίου 2013].

[172] *OpenCV* [Ηλεκτρονικό]. Διαθέσιμο: http://opencv.org. [Πρόσβαση: 15 Ιουνίου 2013].

[173] *Wikipedia, The Free Encyclopedia* [Ηλεκτρονικό]. Διαθέσιμο: http://en.wikipedia.org/wiki/OpenGL. [Πρόσβαση: 15 Ιουνίου 2013].

[174] «Εισαγωγή στην OpenGL: Σημειώσεις Μαθήματος», *Multimedia Technology Laboratory, National Technical University of Athens* [Ηλεκτρονικό]. Διαθέσιμο: http://www.medialab.ntua.gr/education/ComputerGraphics/OpenGL_Lectures/01-Introduction.pdf. [Πρόσβαση: 15 Ιουνίου 2013].

[175] *Wikipedia, The Free Encyclopedia* [Ηλεκτρονικό]. Διαθέσιμο: http://en.wikipedia.org/wiki/OpenGL_Utility_Library. [Πρόσβαση: 15 Ιουνίου 2013].

[176] *Wikipedia, The Free Encyclopedia*, [Ηλεκτρονικό]. Διαθέσιμο: http://en.wikipedia.org/wiki/OpenGL_Utility_Toolkit. [Πρόσβαση: 15 Ιουνίου 2013].

[177] OpenGL Architecture Review Board, M. Woo, J. Neider and T. Davis, *OpenGL Programming Guide,* 2nd ed., Addison-Wesley, 1997.

[178] *freeglut - The Free OpenGL Utility Toolkit* [Ηλεκτρονικό]. Διαθέσιμο: http://freeglut.sourceforge.net. [Πρόσβαση: 15 Ιουνίου 2013].

[179] *GLEW - The OpenGL Extension Wrangler Library* [Ηλεκτρονικό]. Διαθέσιμο: http://glew.sourceforge.net. [Πρόσβαση: 15 Ιουνίου 2013].

[180] D. Shreiner, G. Sellers, J. Kessenich και B. Licea-Kane, *OpenGL Programming Guide*, 8th ed., United States: Addison-Wesley, 2013.

[181] «Φωτορεαλισμός», *Multimedia Technology Laboratory, National Technical University of Athens* [Ηλεκτρονικό]. Διαθέσιμο: http://www.medialab.ntua.gr/education/ComputerGraphics/OpenGL_Slides/08-Chapter6.pdf. [Πρόσβαση: 18 Ιουνίου 2013].

[182] H. A. Song. «OpenGL Normal Vector Transformation», *songho.ca* [Ηλεκτρονικό]. Διαθέσιμο: http://www.songho.ca/opengl/gl_normaltransform.html. [Πρόσβαση: 18 Ιουνίου 2013].

[183] Π. Ράπτης, «Γραφικά Υπολογιστών: Αποκοπή στις 3D Διαστάσεις», Διαφάνειες Μαθήματος, Τ.Ε.Ι. Θεσσαλονίκης, Τμήμα Πληροφορικής, Θεσσαλονίκη.

[184] H. A. Song. «OpenGL Projection Matrix», *songho.ca* [Ηλεκτρονικό]. Διαθέσιμο: http://www.songho.ca/opengl/gl_projectionmatrix.html. [Πρόσβαση: 21 Ιουνίου 2013].

[185] «Είσοδος / Έξοδος - Διαχείριση γεγονότων - Αποκοπή στις δύο διαστάσεις», *Multimedia Technology Laboratory, National Technical University of Athens* [Ηλεκτρονικό]. Διαθέσιμο: http://www.medialab.ntua.gr/education/ComputerGraphics/OpenGL_Lectures/03-Chapter2.pdf. [Πρόσβαση: 21 Ιουνίου 2013].

[186] E.R. Bachmann and P.L. McDowell «Viewport Transformations», *MOVES Institute* [Ηλεκτρονικό]. Διαθέσιμο: http://www.movesinstitute.org/~mcdowell/mv4202/notes/lect8.pdf. [Πρόσβαση: 21 Ιουνίου 2013].

[187] H. A. Song. «OpenGL Transformation», *songho.ca* [Ηλεκτρονικό]. Διαθέσιμο: http://www.songho.ca/opengl/gl_transform.html. [Πρόσβαση: 18 Ιουνίου 2013].

[188] *Wikipedia, The Free Encyclopedia* [Ηλεκτρονικό]. Διαθέσιμο: https://en.wikipedia.org/wiki/Wavefront_.obj_file. [Πρόσβαση: 29 Ιουνίου 2013].

[189] Ε. Τριγώνη, «Κατασκευή Τρισδιάστατων Χώρων σε Μορφότυπο OBJ», Πτυχιακή Εργασία, Τεχνικό Εκπαιδευτικό Ίδρυμα Κρήτης, Σχολή Τεχνολογικών Εφαρμογών, Τμήμα Εφαρμοσμένης Πληροφορικής και Πολυμέσων, Ηράκλειο, 2010.

[190] *FileFormat.Info - Wavefront OBJ File Format Summary* [Ηλεκτρονικό]. Διαθέσιμο: www.fileformat.info/format/wavefrontobj/egff.htm. [Πρόσβαση: 29 Ιουνίου 2013].

[191] «Object Files (.obj)», *Département Informatique* [Ηλεκτρονικό]. Διαθέσιμο: www710.univ-lyon1.fr/~jciehl/Public/utiles/maya_obj_spec.pdf. [Πρόσβαση: 29 Ιουνίου 2013].

[192] «Εργαστήρια OpenGL - Μάθημα 1ο», *Computer Graphics Group* [Ηλεκτρονικό]. Διαθέσιμο: http://graphics.di.uoa.gr/Courses/opengl_course.html. [Πρόσβαση: 1 Ιουλίου 2013].

[193] Γ. Βουγιατζής. «Γραφικά Υπολογιστών, Μάθημα 1», Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης [Ηλεκτρονικό]. Διαθέσιμο:

http://users.auth.gr/voyatzis/CG/COURSE1/gLesson_1.pdf. [Πρόσβαση: 1 Ιουλίου 2013].

[194] *OpenCV Source code - surf.cpp* [Ηλεκτρονικό]. Διαθέσιμο: https://github.com/Itseez/opencv/blob/master/modules/nonfree/src/surf.cpp. [Πρόσβαση: 2 Ιουλίου 2013].

[195] «The OpenCV Reference Manual, Release 2.4.6.0», *OpenCV Documentation* [Ηλεκτρονικό]. Διαθέσιμο: http://docs.opencv.org/opencv2refman.pdf. [Πρόσβαση: 2 Ιουλίου 2013].

[196] *OpenCV Documentation - Common Interfaces of Descriptor Extractors* [Ηλεκτρονικό]. Διαθέσιμο: http://opencv.willowgarage.com/documentation/cpp/features2d_common_interfaces_of_descriptor_extractors.html. [Πρόσβαση: 2 Ιουλίου 2013].

[197] «Απόδοση Υφής», *Multimedia Technology Laboratory, National Technical University of Athens* [Ηλεκτρονικό]. Διαθέσιμο: http://www.medialab.ntua.gr/education/ComputerGraphics/OpenGL_Lectures/08-Chapter7.pdf. [Πρόσβαση: 9 Ιουλίου 2013].

[198] J. Barbic. (15 Φεβρουαρίου 2012). «Texture Mapping», *University of Southern California* [Ηλεκτρονικό]. Διαθέσιμο: http://run.usc.edu/cs480-s12/lec11-textures/11-textures.pdf. [Πρόσβαση: 9 Ιουλίου 2013].

[199] *Wikipedia, The Free Encyclopedia* [Ηλεκτρονικό]. Διαθέσιμο: http://en.wikipedia.org/wiki/C%2B%2B/CLI. [Πρόσβαση: 14 Ιουνίου 2013].

[200] *Wikipedia, The Free Encyclopedia* [Ηλεκτρονικό]. Διαθέσιμο: http://en.wikipedia.org/wiki/Managed_Extensions_for_C%2B%2B. [Πρόσβαση: 14 Ιουνίου 2013].

[201] *Wikipedia, The Free Encyclopedia* [Ηλεκτρονικό]. Διαθέσιμο: http://en.wikipedia.org/wiki/.NET_Framework. [Πρόσβαση: 14 Ιουνίου 2013].

[202] Β. Κόλιας. «Εισαγωγή στη C# και το .NET 4.0», *Multimedia Technology Laboratory, National Technical University of Athens* [Ηλεκτρονικό]. Διαθέσιμο: http://believe.medialab.ntua.gr/guest2/pluginfile.php/564/mod_resource/content/0/UNIT03/unit03.pdf. [Πρόσβαση: 14 Ιουνίου 2013].

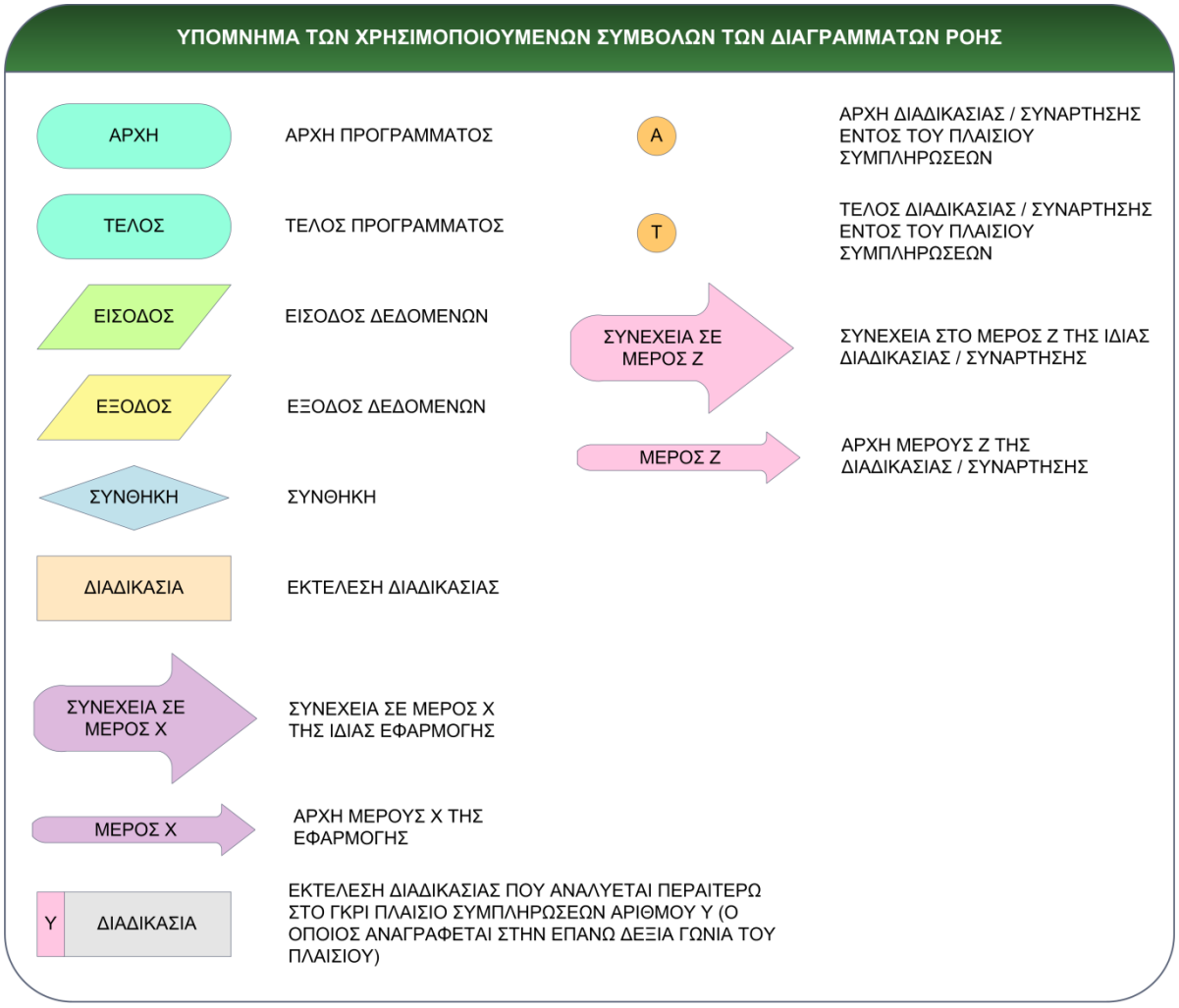
[203] *Wikipedia, The Free Encyclopedia* [Ηλεκτρονικό]. Διαθέσιμο: http://en.wikipedia.org/wiki/Application_programming_interface. [Πρόσβαση: 15 Ιουνίου 2013].

[204] *Βικιπαίδεια, Η Ελεύθερη Εγκυκλοπαίδεια* [Ηλεκτρονικό]. Διαθέσιμο: http://el.wikipedia.org/wiki/Διεπαφή_προγραμματισμού_εφαρμογών. [Πρόσβαση: 15 Ιουνίου 2013].

ΠΑΡΑΡΤΗΜΑ Α

ΧΡΗΣΙΜΟΠΟΙΟΥΜΕΝΑ ΣΥΜΒΟΛΑ ΤΩΝ ΔΙΑΓΡΑΜΜΑΤΩΝ ΡΟΗΣ

Στην Εικόνα Α-1 παρουσιάζεται το υπόμνημα των χρησιμοποιούμενων συμβόλων των διαγραμμάτων ροής της παρούσας διπλωματικής εργασίας.



Εικόνα Α-1. Υπόμνημα των χρησιμοποιούμενων συμβόλων των διαγραμμάτων ροής.

ΠΑΡΑΡΤΗΜΑ Β

ΑΠΟΔΕΙΞΗ ΤΗΣ ΣΧΕΣΗΣ (3.39)

Έστω **K** ο πίνακας της μηχανής στην περίπτωση κατά την οποία ο συντελεστής skewness θεωρείται μηδενικός (s=0):

$$\mathbf{K} = \begin{bmatrix} c_x & 0 & x_0 \\ 0 & c_y & y_0 \\ 0 & 0 & 1 \end{bmatrix}$$

Ο αντίστροφος του πίνακα της μηχανής, $\mathbf{K}^{-1}$, είναι ο εξής:

$$\mathbf{K}^{-1} = \frac{1}{c_x \cdot c_y} \begin{bmatrix} c_y & 0 & -x_0 \cdot c_y \\ 0 & c_x & -y_0 \cdot c_x \\ 0 & 0 & c_x \cdot c_y \end{bmatrix}$$

Η σχέση (3.38) γράφεται ισοδύναμα:

$$\lambda \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} c_x & 0 & x_0 \\ 0 & c_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad \rightarrow$$

$$\lambda \cdot \begin{bmatrix} x & y & 1 \end{bmatrix}^T = \mathbf{K} \cdot \begin{bmatrix} \mathbf{R} \mid \mathbf{t} \end{bmatrix} \cdot \begin{bmatrix} X & Y & Z & 1 \end{bmatrix}^T \quad \rightarrow$$

$$\lambda \cdot \mathbf{K}^{-1} \cdot \begin{bmatrix} x & y & 1 \end{bmatrix}^T = \mathbf{K}^{-1} \cdot \mathbf{K} \cdot \begin{bmatrix} \mathbf{R} \mid \mathbf{t} \end{bmatrix} \cdot \begin{bmatrix} X & Y & Z & 1 \end{bmatrix}^T \quad \rightarrow$$

$$\lambda \cdot \mathbf{K}^{-1} \cdot \begin{bmatrix} x & y & 1 \end{bmatrix}^T = \begin{bmatrix} \mathbf{R} \mid \mathbf{t} \end{bmatrix} \cdot \begin{bmatrix} X & Y & Z & 1 \end{bmatrix}^T \quad \rightarrow$$

$$\frac{\lambda}{c_x \cdot c_y} \cdot \begin{bmatrix} c_y & 0 & -x_0 \cdot c_y \\ 0 & c_x & -y_0 \cdot c_x \\ 0 & 0 & c_x \cdot c_y \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad \rightarrow$$

$$\frac{\lambda}{c_x \cdot c_y} \cdot \begin{bmatrix} c_y & 0 & -x_0 \cdot c_y \\ 0 & c_x & -y_0 \cdot c_x \\ 0 & 0 & c_x \cdot c_y \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix} \quad \rightarrow$$

$$\frac{\lambda}{c_x \cdot c_y} \cdot \begin{bmatrix} x \cdot c_y - x_0 \cdot c_y \\ y \cdot c_x - y_0 \cdot c_x \\ c_x \cdot c_y \end{bmatrix} = \mathbf{R} \cdot \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \mathbf{t} \quad \rightarrow \quad \begin{bmatrix} \lambda \cdot (x - x_0)/c_x \\ \lambda \cdot (y - y_0)/c_y \\ \lambda \end{bmatrix} = \mathbf{R} \cdot \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \mathbf{t} \quad \rightarrow$$

$$\begin{bmatrix} x' \\ y' \\ w' \end{bmatrix} = \mathbf{R} \cdot \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \mathbf{t} \quad \text{όπου} \quad \left\{ \begin{bmatrix} x' = \dfrac{(x - x_0) \cdot \lambda}{c_x} \\ y' = \dfrac{(y - y_0) \cdot \lambda}{c_y} \\ w' = \lambda \end{bmatrix} \right\} \quad \rightarrow \quad \left\{ \begin{bmatrix} x' = \dfrac{(x - x_0) \cdot w'}{c_x} \\ y' = \dfrac{(y - y_0) \cdot w'}{c_y} \end{bmatrix} \right\}$$

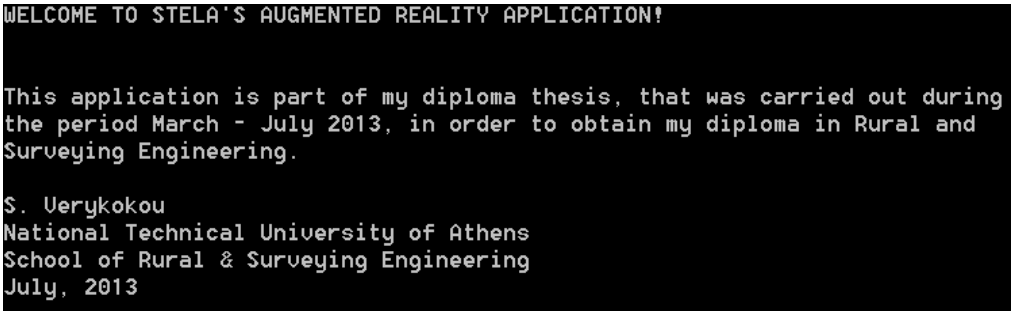
Συνεπώς, οι σχέσεις: $\lambda \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \mathbf{K} \cdot [\mathbf{R} \mid \mathbf{t}] \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$ και $\begin{bmatrix} x' \\ y' \\ w' \end{bmatrix} = \mathbf{R} \cdot \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \mathbf{t}$ είναι ισοδύναμες.

ΠΑΡΑΡΤΗΜΑ Γ

ΕΓΧΕΙΡΙΔΙΟ ΧΡΗΣΗΣ ΤΗΣ ΕΦΑΡΜΟΓΗΣ 1

Το παρόν παράρτημα αποτελεί εγχειρίδιο χρήσης της Εφαρμογής 1, στο οποίο παρουσιάζεται η λειτουργία της για όλες τις πιθανές επιλογές του χρήστη.

Με την έναρξη της Εφαρμογής 1, εμφανίζεται ένα εισαγωγικό μήνυμα σχετικά με το δημιουργό της εφαρμογής, καθώς επίσης και το σκοπό και την περίοδο ανάπτυξης της εφαρμογής, όπως φαίνεται στην Εικόνα Γ-1.

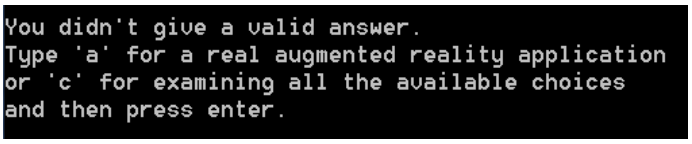


Εικόνα Γ-1. Εισαγωγικό μήνυμα των Εφαρμογών 1 και 2.

Στη συνέχεια, ο χρήστης ερωτάται αν επιθυμεί μία εφαρμογή επαυξημένης πραγματικότητας σύμφωνη με τον ορισμό που δόθηκε από τον Azuma, κατά την οποία η επαύξηση της πραγματικότητας γίνεται σε πραγματικό χρόνο μέσω λήψης στιγμιότυπων από την κάμερα του ηλεκτρονικού υπολογιστή, ή αν θέλει να εξετάσει όλες τις πιθανές επιλογές του. Ο χρήστης καλείται να πληκτρολογήσει το χαρακτήρα «a» στην πρώτη περίπτωση, ή το χαρακτήρα «c» στη δεύτερη και έπειτα να πατήσει το πλήκτρο «Enter» (βλ. Εικόνα Γ-2). Η πληκτρολόγηση του χαρακτήρα «a» οδηγεί σε μετέπειτα εμφάνιση ενός μόνο παραθύρου που θα απεικονίζει την επαυξημένη σκηνή.

Εικόνα Γ-2. Μήνυμα που εμφανίζεται στο χρήστη της Εφαρμογής 1 για την πληκτρολόγηση της επιλογής του σχετικά με το αν επιθυμεί μία πραγματική εφαρμογή επαυξημένης πραγματικότητας ή αν επιδιώκει να εξετάσει τις δυνατές επιλογές του.

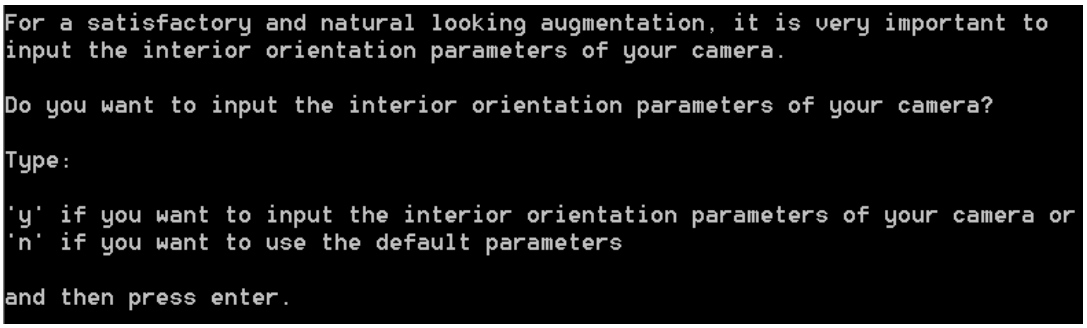
Για την εξασφάλιση της ορθότητας της απάντησης του χρήστη, δηλαδή της πληκτρολόγησης ενός εκ των χαρακτήρων «a» ή «c», πραγματοποιείται ένας έλεγχος της εισόδου του χρήστη. Έτσι, αν η είσοδος δεν συμπίπτει με κανέναν από τους δύο χαρακτήρες, εμφανίζεται ένα μήνυμα που τον ενημερώνει για τη λανθασμένη του απάντηση και τον προτρέπει να πληκτρολογήσει εκ νέου τη σωστή απάντηση, όπως φαίνεται στην Εικόνα Γ-3. Η διαδικασία αυτή επαναλαμβάνεται μέχρι ο χρήστης να πληκτρολογήσει έναν από τους χαρακτήρες «a» ή «c».



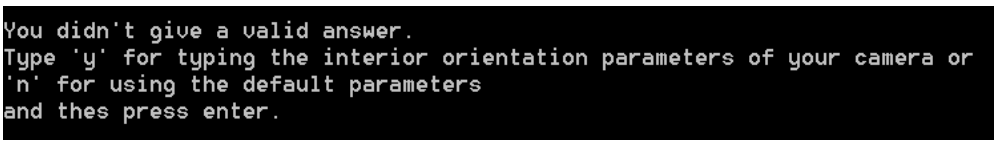
Εικόνα Γ-3. Επισήμανση της λανθασμένης απάντησης του χρήστη της Εφαρμογής 1 και προτροπή να πληκτρολογήσει τη σωστή απάντηση, ανάμεσα στους χαρακτήρες «a» ή «c».

Γ.1 Επιλογή της Κανονικής Εφαρμογής Επαυξημένης Πραγματικότητας

Όταν ο χρήστης, μετά από το μήνυμα της Εικόνας Γ-2, πληκτρολογήσει το χαρακτήρα «a», που δηλώνει επιλογή της κανονικής εφαρμογής επαυξημένης πραγματικότητας, εμφανίζεται ένα νέο μήνυμα που τον ενημερώνει για την αναγκαιότητα της εισαγωγής των στοιχείων του εσωτερικού προσανατολισμού της κάμερας του υπολογιστή του ώστε να επιτευχθεί μία ικανοποιητική και ρεαλιστική επαύξηση της σκηνής (βλ. Εικόνα Γ-4).



Εικόνα Γ-4. Μήνυμα που εμφανίζεται στο χρήστη των εφαρμογών με το ερώτημα αν θέλει να πληκτρολογήσει τα στοιχεία του εσωτερικού προσανατολισμού της κάμερας.



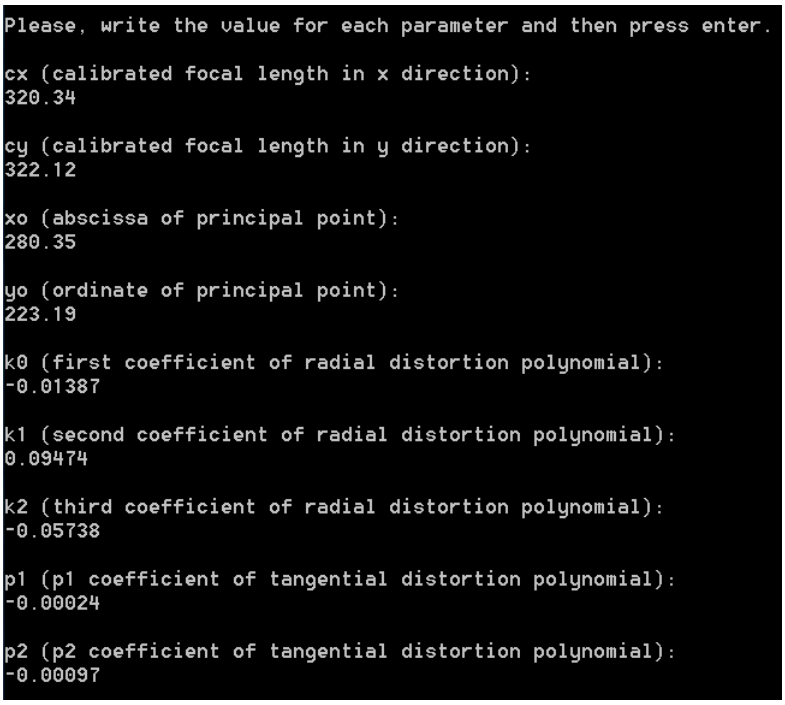
Εικόνα Γ-5. Επισήμανση της λανθασμένης απάντησης του χρήστη των εφαρμογών και προτροπή να πληκτρολογήσει τη σωστή απάντηση, ανάμεσα στους χαρακτήρες «y» ή «n».

Ομοίως ελέγχεται η απάντηση του χρήστη (βλ. Εικόνα Γ-5), η οποία πρέπει να είναι ένας εκ των χαρακτήρων «y» ή «n». Αν δοθεί λανθασμένη απάντηση, ο χρήστης ενημερώνεται για

το λάθος του και καλείται να πληκτρολογήσει τη σωστή απάντηση, διαδικασία η οποία επαναλαμβάνεται μέχρι να πληκτρολογήσει το χαρακτήρα «y» ή «n».

Όταν ο χρήστης πληκτρολογήσει τη σωστή απάντηση, και εφόσον αυτή είναι ο χαρακτήρας «n», δεν εμφανίζεται άλλο μήνυμα που να απαιτεί από αυτόν να δηλώσει κάποια επιλογή του και στη συνέχεια εμφανίζονται τα αποτελέσματα της εφαρμογής επαυξημένης πραγματικότητας.

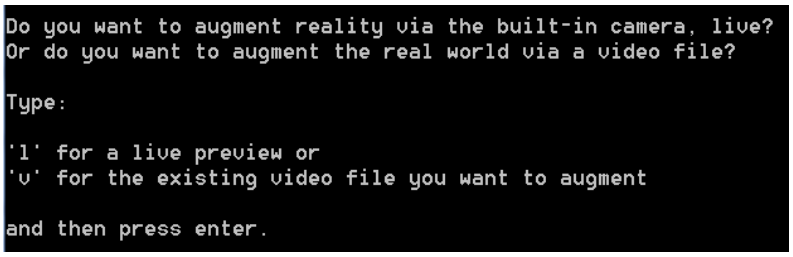
Διαφορετικά, αν ο χρήστης επέλεξε να εισαγάγει τα στοιχεία του εσωτερικού προσανατολισμού, πληκτρολογώντας το χαρακτήρα «y», εμφανίζεται ένα μήνυμα που τον ενημερώνει να πληκτρολογήσει τα στοιχεία του εσωτερικού προσανατολισμού. Στην Εικόνα Γ-6 φαίνεται η αλληλεπίδραση του χρήστη με την εφαρμογή, κατά την οποία κάθε φορά που πληκτρολογεί ένα αριθμό που αντιστοιχεί σε ένα στοιχείο του εσωτερικού προσανατολισμού και πατά το πλήκτρο «Enter», του ζητείται να πληκτρολογήσει το επόμενο στοιχείο, μέχρις ότου αυτός πληκτρολογήσει όλα τα στοιχεία του εσωτερικού προσανατολισμού. Αυτά είναι οι δύο τιμές για τη σταθερά της μηχανής (c_x και c_y), οι συντεταγμένες του πρωτεύοντος σημείου (x_0 και y_0), οι τρεις συντελεστές του πολυωνύμου ακτινικής διαστροφής και οι δύο συντελεστές του πολυωνύμου της ασύμμετρης διαστροφής εκκεντρότητας. Μετά από την πληκτρολόγηση των στοιχείων του εσωτερικού προσανατολισμού, εμφανίζονται τα αποτελέσματα της εφαρμογής επαυξημένης πραγματικότητας.



Εικόνα Γ-6. Αλληλεπίδραση του χρήστη με την εφαρμογή, κατά την οποία πληκτρολογεί τα στοιχεία του εσωτερικού προσανατολισμού.

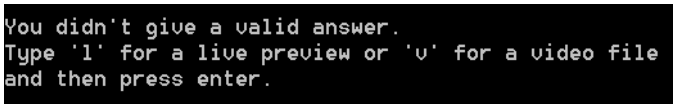
Γ.2 Εξέταση των Εναλλακτικών Επιλογών που Παρέχονται από την Εφαρμογή

Αν ο χρήστης δεν επιλέξει την κανονική εφαρμογή επαυξημένης πραγματικότητας αλλά δηλώσει – μέσω της πληκτρολόγησης του χαρακτήρα «c», ύστερα από την εμφάνιση του κειμένου που παρουσιάζεται στην Εικόνα Γ-2 – ότι θέλει να εξετάσει τις δυνατότητες που του προσφέρει η εφαρμογή, ερωτάται αν επιθυμεί την επαύξηση της πραγματικότητας μέσω της κάμερας του υπολογιστή του, σε πραγματικό χρόνο, ή μέσω ενός αρχείου βίντεο, όπως φαίνεται στην Εικόνα Γ-7.



Εικόνα Γ-7. Μήνυμα που εμφανίζεται στο χρήστη της Εφαρμογής 1 με το ερώτημα αν θέλει να επαυξήσει την πραγματικότητα ζωντανά μέσω της κάμερας του υπολογιστή του ή μέσω ενός αρχείου βίντεο.

Ομοίως, η απάντηση που δίνει ο χρήστης ελέγχεται και εφόσον δεν είναι σωστή, δηλαδή είναι διάφορη του χαρακτήρα «l», που δηλώνει επαύξηση μέσω της κάμερας του υπολογιστή σε πραγματικό χρόνου, και του χαρακτήρα «v», που δηλώνει επαύξηση των σκηνών ενός αρχείου βίντεο, εμφανίζεται κατάλληλο μήνυμα (βλ. Εικόνα Γ-8). Το τελευταίο ενημερώνει το χρήστη για τη λανθασμένη απάντηση, τον προτρέπει να πληκτρολογήσει εκ νέου τη σωστή απάντηση και εμφανίζεται στο χρήστη μέχρι να πληκτρολογήσει το σωστό χαρακτήρα.



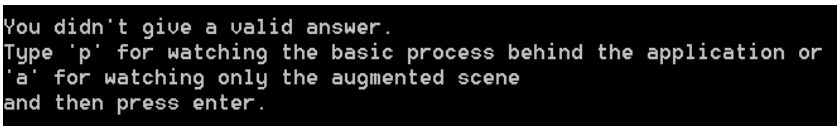
Εικόνα Γ-8. Επισήμανση της λανθασμένης απάντησης του χρήστη της Εφαρμογής 1 και προτροπή να πληκτρολογήσει τη σωστή απάντηση, ανάμεσα στους χαρακτήρες «l» ή «v».

Γ.2.1 Χρήση της κάμερας του υπολογιστή

Αν ο χρήστης, μετά από το μήνυμα της Εικόνας Γ-7, πληκτρολογήσει το χαρακτήρα «l», που υποδηλώνει χρήση της κάμερας του υπολογιστή σε πραγματικό χρόνο, ερωτάται αν εκτός από το παράθυρο θέασης της επαυξημένης σκηνής επιθυμεί να εμφανιστούν και άλλα παράθυρα, που παρουσιάζουν τη βασική διαδικασία που πραγματοποιείται με χρήση της βιβλιοθήκης OpenCV, προκειμένου να επιτευχθεί ο σκοπός της εφαρμογής. Αν ο χρήστης το επιθυμεί, καλείται να πληκτρολογήσει το χαρακτήρα «p». Σε διαφορετική περίπτωση, πληκτρολογεί το χαρακτήρα «a» που υποδηλώνει ότι αυτός θα δει μόνο την επαυξημένη σκηνή (βλ. Εικόνα Γ-9).

Εικόνα Γ-9. Μήνυμα με το οποίο ερωτάται ο χρήστης των εφαρμογών αν θέλει να δει τη βασική διαδικασία που πραγματοποιείται για την επαύξηση της πραγματικότητας.

Ομοίως, πραγματοποιείται έλεγχος της απάντησης του χρήστη, ο οποίος σε περίπτωση λανθασμένης πληκτρολόγησης καλείται να πληκτρολογήσει εκ νέου τη σωστή απάντηση (βλ. Εικόνα Γ-10), διαδικασία που επαναλαμβάνεται μέχρι να πληκτρολογήσει έναν εκ των χαρακτήρων «a» ή «p».



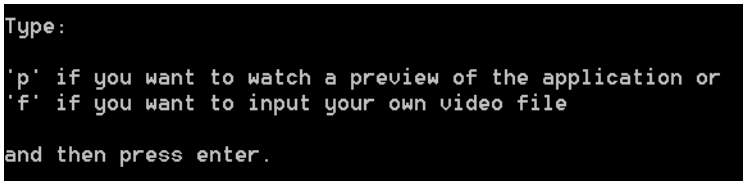
Εικόνα Γ-10. Επισήμανση της λανθασμένης απάντησης του χρήστη και προτροπή να πληκτρολογήσει τη σωστή απάντηση, ανάμεσα στους χαρακτήρες «p» και «a».

Όταν ο χρήστης πληκτρολογήσει τη σωστή απάντηση, ερωτάται αν θα εισαγάγει τα στοιχεία του εσωτερικού προσανατολισμού της κάμερας του υπολογιστή του, δηλαδή εμφανίζεται το μήνυμα που φαίνεται στην Εικόνα Γ-4 και εν συνεχεία, αν κάνει λάθος, εμφανίζεται το μήνυμα που απεικονίζεται στην Εικόνα Γ-5. Στην περίπτωση που επιλέξει να εισαγάγει τα στοιχεία του εσωτερικού προσανατολισμού καλείται να τα πληκτρολογήσει ένα-ένα, όπως φαίνεται στην Εικόνα Γ-6. Κατόπιν, ο χρήστης δεν καλείται να δηλώσει κάποια άλλη επιλογή του και ακολούθως εμφανίζονται τα αποτελέσματα της εφαρμογής επαυξημένης πραγματικότητας, ανάλογα με τις προηγούμενες επιλογές του.

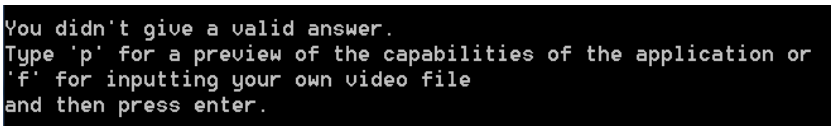
Γ.2.2 Επαύξηση ενός αρχείου βίντεο

Αν ο χρήστης, μετά από το μήνυμα της Εικόνας Γ-7, πληκτρολογήσει το χαρακτήρα «v» αντί του χαρακτήρα «l», γεγονός το οποίο σημαίνει ότι για την επαύξηση της πραγματικότητας θα χρησιμοποιηθεί κάποιο αρχείο βίντεο, ερωτάται αν επιθυμεί να παρακολουθήσει μία προεπισκόπηση της εφαρμογής με ένα προεπιλεγμένο αρχείο βίντεο, ή αν θέλει να εισαγάγει το δικό του βίντεο, προκειμένου να επαυξηθούν οι σκηνές του. Στην πρώτη περίπτωση ο χρήστης πρέπει να πληκτρολογήσει το χαρακτήρα «p», ενώ στη δεύτερη περίπτωση το χαρακτήρα «f» (βλ. Εικόνα Γ-11).

Ομοίως εμφανίζεται ένα μήνυμα που τον καλεί να πληκτρολογήσει εκ νέου την απάντηση, στην περίπτωση που είναι διάφορη του χαρακτήρα «p» και του χαρακτήρα «f» και εξακολουθεί να εμφανίζεται μέχρι ο χρήστης να δώσει τη σωστή απάντηση (βλ. Εικόνα Γ-12).



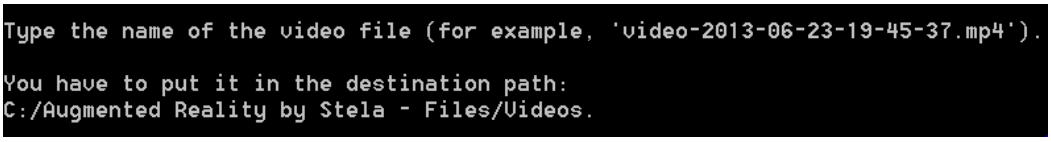
Εικόνα Γ-11. Μήνυμα που εμφανίζεται στο χρήστη της Εφαρμογής 1 προκειμένου να επιλέξει να δει μία προεπισκόπηση της εφαρμογής ή να εισαγάγει το δικό του αρχείο βίντεο.



Εικόνα Γ-12. Επισήμανση της λανθασμένης απάντησης του χρήστη της Εφαρμογής 1 και προτροπή να πληκτρολογήσει τη σωστή απάντηση, ανάμεσα στους χαρακτήρες «p» και «f».

Γ.2.2.1 Εισαγωγή ενός αρχείου βίντεο από το χρήστη

Αν ο χρήστης πληκτρολογήσει το χαρακτήρα «f» μετά από το μήνυμα της Εικόνας Γ-11, καλείται να εισαγάγει το όνομα του αρχείου βίντεο που περιλαμβάνει τη σκηνή την οποία θέλει να επαυξήσει (βλ. Εικόνα Γ-13). Το αρχείο αυτό για να εντοπιστεί από την εφαρμογή θα πρέπει να βρίσκεται στον υποφάκελο «Videos» που βρίσκεται εντός του φακέλου «Augmented Reality by Stela - Files» που διανέμεται μαζί με την εφαρμογή, και ο τελευταίος να τοποθετηθεί στο δίσκο C του υπολογιστή.



Εικόνα Γ-13. Μήνυμα που καλεί το χρήστη της Εφαρμογής 1 να πληκτρολογήσει το όνομα του αρχείου βίντεο, το περιεχόμενο του οποίου επιθυμεί να επαυξήσει.

Όταν ο χρήστης πληκτρολογήσει το όνομα του αρχείου, εμφανίζεται το μήνυμα που φαίνεται στην Εικόνα Γ-9, για να διαπιστωθεί αν επιθυμεί να δει τη βασική διαδικασία που οδηγεί στην επαύξηση της πραγματικότητας. Εφόσον πληκτρολογήσει λάθος απάντηση, εμφανίζεται το μήνυμα της Εικόνας Γ-10. Στη συνέχεια, εμφανίζεται το μήνυμα που φαίνεται στην Εικόνα Γ-4, για εισαγωγή των στοιχείων του εσωτερικού προσανατολισμού, και ανάλογα με την απάντησή του μπορεί να εμφανιστούν τα μηνύματα των Εικόνων Γ-5 και Γ-6. Ακολούθως, εμφανίζονται τα αποτελέσματα της εφαρμογής επαυξημένης πραγματικότητας, ανάλογα με τις επιλογές του χρήστη.

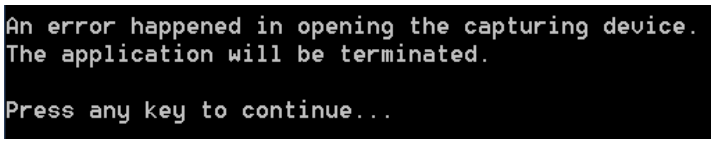
Γ.2.2.2 Χρήση του προεπιλεγμένου αρχείου βίντεο

Αν ο χρήστης, μετά από το μήνυμα της Εικόνας Γ-11, πληκτρολογήσει το χαρακτήρα «p», που σημαίνει ότι επιθυμεί να δει μία προεπισκόπηση της εφαρμογής, παραλείπεται το μήνυμα της Εικόνας Γ-13 και ακολουθεί η εμφάνιση του μηνύματος της Εικόνας Γ-9 και

ανάλογα με την ορθότητα της απάντησης που θα δώσει, μπορεί να ακολουθήσει η εμφάνιση του μηνύματος της Εικόνας Γ-10. Καθώς ο χρήστης έχει επιλέξει να δει μία προεπισκόπηση της εφαρμογής με κάποιο προεπιλεγμένο αρχείο βίντεο, δεν υφίσταται η ανάγκη να εισαγάγει τα στοιχεία του εσωτερικού προσανατολισμού, αφού αυτά έχουν υπολογιστεί μέσω της διαδικασίας της βαθμονόμησης και έχουν εισαχθεί στο πρόγραμμα. Κατόπιν, εμφανίζονται τα αποτελέσματα της εφαρμογής επαυξημένης πραγματικότητας.

Γ.3 Επιπρόσθετες Διευκρινίσεις

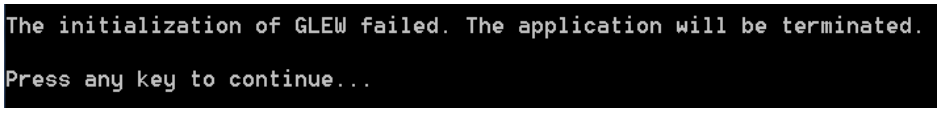
Αν η κάμερα του υπολογιστή ή το αρχείο βίντεο που εισήχθη από το χρήστη ή το προεπιλεγμένο αρχείο βίντεο – ανάλογα με τις προηγούμενες επιλογές του – δεν άνοιξε επιτυχώς, εμφανίζεται το κατάλληλο προειδοποιητικό μήνυμα για τον τερματισμό της εφαρμογής και ο χρήστης καλείται να πληκτρολογήσει κάποιο χαρακτήρα, ακολουθούμενο από το πλήκτρο «Enter», προκειμένου να εξασφαλιστεί ότι είδε το αντίστοιχο μήνυμα. Μετά από την είσοδο οποιουδήποτε χαρακτήρα από το χρήστη, η εφαρμογή τερματίζεται (βλ. Εικόνες Γ-14 και Γ-15). Το ανεπιτυχές άνοιγμα της κάμερας του υπολογιστή μπορεί να οφείλεται στο γεγονός ότι δεν έχει ενεργοποιηθεί από το χρήστη, ενώ η μη επιτυχία κατά το άνοιγμα του αρχείου βίντεο μπορεί να οφείλεται σε λανθασμένη πληκτρολόγηση του ονόματός του από το χρήστη. Στην τελευταία περίπτωση, εμφανίζεται κατάλληλη προειδοποίηση και από τη βιβλιοθήκη OpenCV (πρώτη γραμμή που εμφανίζεται στην Εικόνα Γ-15).



Εικόνα Γ-14. Προειδοποιητικό μήνυμα προς το χρήστη της Εφαρμογής 1 σχετικά με την αποτυχία κατά το άνοιγμα της κάμερας του υπολογιστή.

Εικόνα Γ-15. Προειδοποιητικό μήνυμα προς το χρήστη σχετικά με την αποτυχία κατά το άνοιγμα του αρχείου βίντεο.

Επίσης, σε περίπτωση ανεπιτυχούς αρχικοποίησης της βιβλιοθήκης GLEW, εμφανίζεται κατάλληλο μήνυμα (βλ. Εικόνα Γ-16) που ειδοποιεί για τον επερχόμενο τερματισμό της εφαρμογής, μετά από την πληκτρολόγηση κάποιου χαρακτήρα από το χρήστη, ακολουθούμενο από το πλήκτρο «Enter».

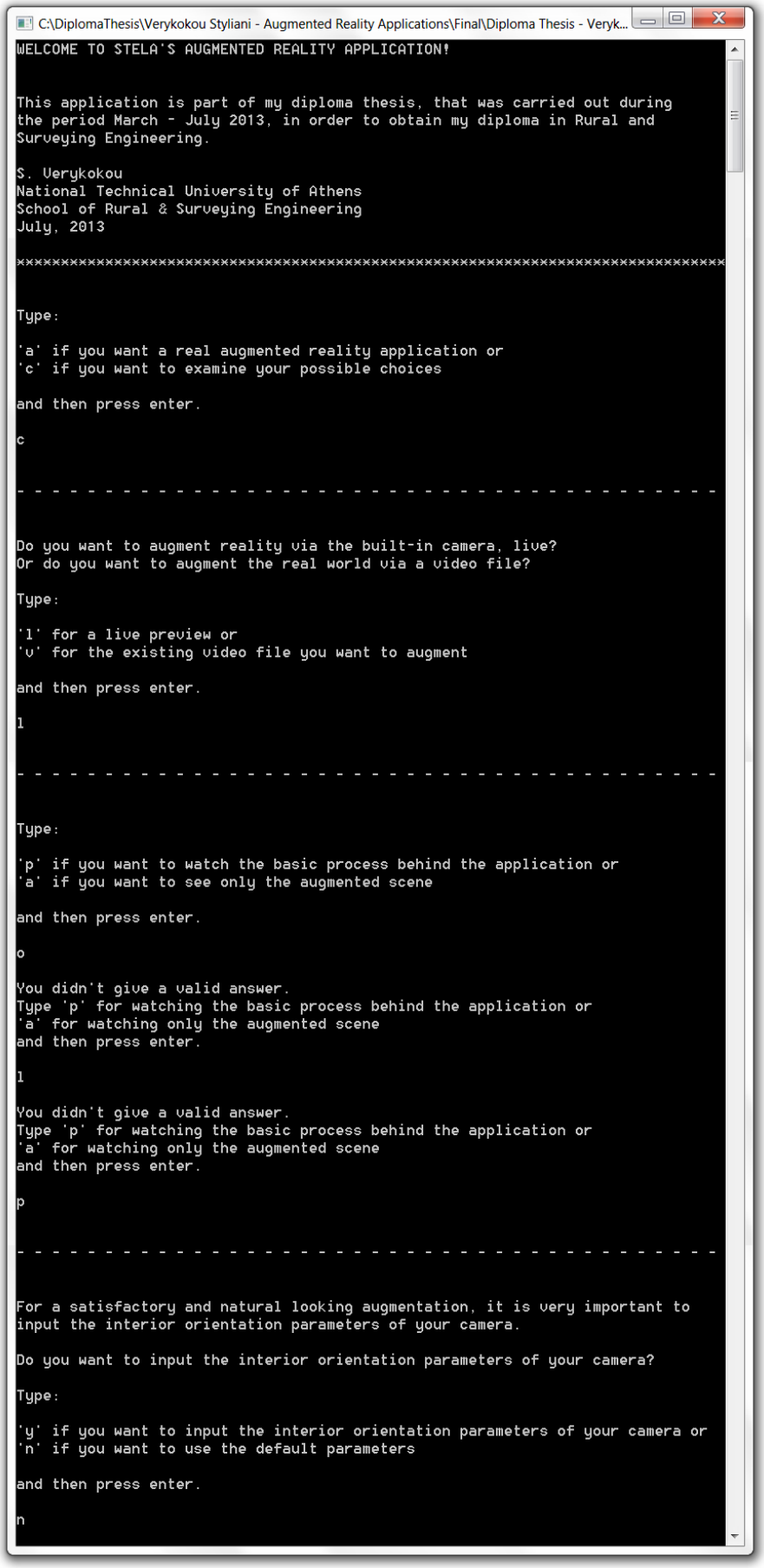


Εικόνα Γ-16. Μήνυμα σφάλματος κατά την αρχικοποίηση της βιβλιοθήκης GLEW.

Αν ο χρήστης πατήσει το πλήκτρο «Escape» και το παράθυρο «AR Window» είναι ενεργό, η εφαρμογή τερματίζεται. Αν, ωστόσο, πατηθεί το πλήκτρο «Escape» ενώ στο προσκήνιο είναι ένα άλλο παράθυρο, το πάτημα του πλήκτρου αυτού δεν θα φέρει καμία αλλαγή στην εφαρμογή. Έτσι, για να τερματιστεί η τελευταία με το πάτημα του «Escape» απαιτείται προηγουμένως να έχει επισημανθεί με το ποντίκι το παράθυρο «AR Window» (μέσω πατήματος του κουμπιού του ποντικιού στο συγκεκριμένο παράθυρο). Επιπλέον, η εφαρμογή τερματίζεται αν σταματήσει το βίντεο, σε περίπτωση που δεν έχει επιλεγεί η ζωντανή επαύξηση της πραγματικότητας.

Αν ο χρήστης έχει επιλέξει να παρακολουθήσει τη βασική διαδικασία που διεξάγεται μέσω της OpenCV, μετά τον καθορισμό των επιλογών του εμφανίζεται ένα παράθυρο στο οποίο παρουσιάζεται η εικόνα - πρότυπο. Για να κλείσει το παράθυρο αυτό, ο χρήστης πρέπει να πατήσει οποιοδήποτε πλήκτρο και το εν λόγω παράθυρο να είναι ενεργό. Στη συνέχεια, εμφανίζεται ένα παράθυρο που απεικονίζει τα σημεία ενδιαφέροντος που εντοπίστηκαν στην εικόνα - πρότυπο. Ομοίως για να κλείσει το παράθυρο αυτό και να ξεκινήσει η διαδικασία επαύξησης της πραγματικότητας, ο χρήστης πρέπει να πατήσει οποιοδήποτε πλήκτρο και το εν λόγω παράθυρο να είναι ενεργό.

Στην Εικόνα Γ-17 παρουσιάζεται ένα παράδειγμα αλληλεπίδρασης του χρήστη με την Εφαρμογή 1.



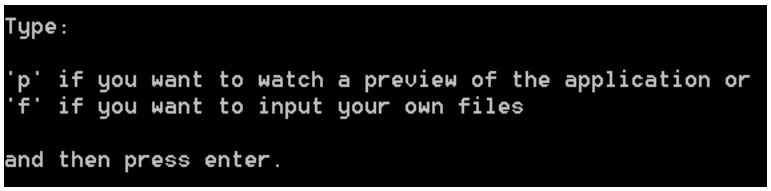
Εικόνα Γ-17. Παράδειγμα αλληλεπίδρασης του χρήστη με την Εφαρμογή 1.

ΠΑΡΑΡΤΗΜΑ Δ

ΕΓΧΕΙΡΙΔΙΟ ΧΡΗΣΗΣ ΤΗΣ ΕΦΑΡΜΟΓΗΣ 2

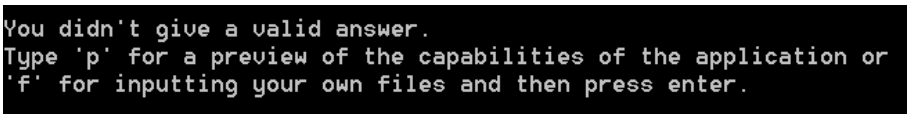
Το παρόν παράρτημα αποτελεί εγχειρίδιο χρήσης της Εφαρμογής 2, στο οποίο παρουσιάζεται η λειτουργία της για όλες τις πιθανές επιλογές του χρήστη.

Με την έναρξη της Εφαρμογής 2, εμφανίζεται ένα εισαγωγικό μήνυμα, ίδιο με εκείνο που εμφανίζεται στην Εφαρμογή 1 και φαίνεται στην Εικόνα Γ-1 (βλ. Παράρτημα Γ). Ακολουθεί η εμφάνιση ενός μηνύματος που προτρέπει το χρήστη να πληκτρολογήσει το χαρακτήρα «p», αν επιδιώκει να δει μία προεπισκόπηση της εφαρμογής, ή το χαρακτήρα «f» αν επιθυμεί να εισαγάγει τα δικά του αρχεία (δηλαδή το αρχείο βίντεο και το αρχείο με τις τιμές του GPS για τη θέση του κινητού κατά τη διάρκεια λήψης του βίντεο), όπως φαίνεται στην Εικόνα Δ-1.



Εικόνα Δ-1. Μήνυμα που εμφανίζεται στο χρήστη της Εφαρμογής 2 για το αν επιθυμεί να δει μία προεπισκόπηση της εφαρμογής ή να εισαγάγει τα δικά του αρχεία.

Η απάντηση του χρήστη ελέγχεται και εφόσον δεν είναι ο χαρακτήρας «p» ή ο χαρακτήρας «f» εμφανίζεται κατάλληλο μήνυμα που τον ενημερώνει για τη λανθασμένη απάντησή του και τον παρακινεί να εισαγάγει τη σωστή, όπως φαίνεται στην Εικόνα Δ-2. Η διαδικασία επαναλαμβάνεται μέχρι ο χρήστης να πληκτρολογήσει τη σωστή απάντηση.

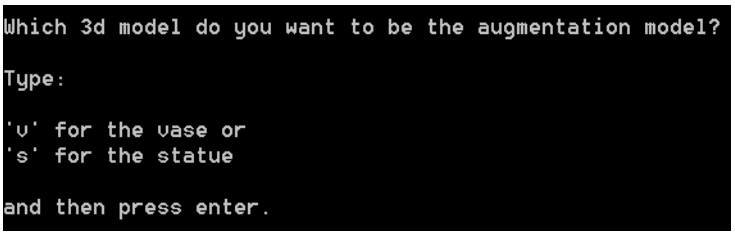


Εικόνα Δ-2. Επισήμανση της λανθασμένης απάντησης του χρήστη της Εφαρμογής 2 και προτροπή να πληκτρολογήσει τη σωστή απάντηση, ανάμεσα στους χαρακτήρες «p» και «f».

Δ.1 Παρακολούθηση Προεπισκόπησης της Εφαρμογής

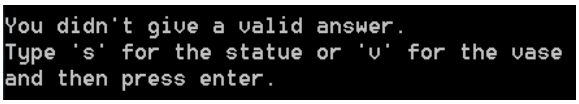
Αν ο χρήστης επιλέξει να δει μία προεπισκόπηση της εφαρμογής, μέσω πληκτρολόγησης του χαρακτήρα «p» μετά από το μήνυμα της Εικόνας Δ-1, ερωτάται ποιο τρισδιάστατο μοντέλο επιθυμεί να αποτελέσει το μοντέλο επαύξησης. Συγκεκριμένα, καλείται να

πληκτρολογήσει το χαρακτήρα «v» για το αγγείο ή «s» για το άγαλμα και στη συνέχεια να πατήσει το πλήκτρο «Enter» (βλ. Εικόνα Δ-3).



Εικόνα Δ-3. Ερώτηση που απευθύνεται στο χρήστη της Εφαρμογής 2 σχετικά με την επιλογή του τρισδιάστατου μοντέλου που θα αποτελέσει το μοντέλο επαύξησης της πραγματικότητας.

Ομοίως, η είσοδος του χρήστη ελέγχεται. Στην περίπτωση που είναι λανθασμένη εμφανίζεται κατάλληλο μήνυμα που τον πληροφορεί για τη μη έγκυρη απάντηση και τον καλεί να πληκτρολογήσει εκ νέου το σωστό χαρακτήρα, μεταξύ των «v» και «s» (βλ. Εικόνα Δ-4). Το μήνυμα αυτό εμφανίζεται κάθε φορά που εισάγει μία λανθασμένη απάντηση, μέχρι να πληκτρολογήσει το σωστό χαρακτήρα.



Εικόνα Δ-4. Επισήμανση της λανθασμένης απάντησης του χρήστη της Εφαρμογής 2 και προτροπή να πληκτρολογήσει τη σωστή απάντηση, ανάμεσα στους χαρακτήρες «s» και «v».

Στη συνέχεια, ο χρήστης ερωτάται αν επιθυμεί να παρακολουθήσει – εκτός από την επαυξημένη σκηνή – τη βασική διαδικασία που εκτελείται μέσω της βιβλιοθήκης OpenCV, προκειμένου να επιτευχθεί το τελικό αποτέλεσμα. Το αντίστοιχο μήνυμα είναι αυτό που εμφανίζεται στην Εικόνα Γ-9 (βλ. Παράρτημα Γ), το οποίο προτρέπει το χρήστη να πληκτρολογήσει το χαρακτήρα «p» ή το χαρακτήρα «a». Ο επαναληπτικός έλεγχος που πραγματοποιείται σε περίπτωση λανθασμένης απάντησης του χρήστη φαίνεται στην Εικόνα Γ-10 (βλ. Παράρτημα Γ). Όταν ο χρήστης πληκτρολογήσει μία εκ των δύο επιτρεπτών απαντήσεων, εμφανίζονται τα αποτελέσματα της εφαρμογής επαυξημένης πραγματικότητας. Καθώς εξετάζεται η περίπτωση κατά την οποία ο χρήστης είχε επιλέξει να δει μία προεπισκόπηση της εφαρμογής, δεν υφίσταται η ανάγκη να εισαγάγει στοιχεία του εσωτερικού προσανατολισμού.

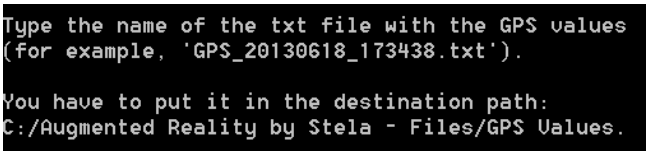
Δ.2 Εισαγωγή Αρχείων από το Χρήστη

Αν ο χρήστης δεν επιλέξει να παρακολουθήσει μία προεπισκόπηση της εφαρμογής αλλά δηλώσει την επιθυμία – μέσω πληκτρολόγησης του χαρακτήρα «f» μετά από το μήνυμα που απεικονίζεται στην Εικόνα Δ-1 – να εισαγάγει τα δικά του αρχεία (βίντεο και GPS), εμφανίζεται αρχικά ένα μήνυμα που τον καλεί να πληκτρολογήσει το όνομα του αρχείου βίντεο (βλ. Εικόνα Δ-5). Επίσης, ενημερώνεται ότι το αρχείο πρέπει να είναι τοποθετημένο

στον υποφάκελο «Videos» του φακέλου «Augmented Reality by Stela - Files», που διανέμεται μαζί με την εφαρμογή, και ότι ολόκληρος ο φάκελος πρέπει να βρίσκεται στο δίσκο C του υπολογιστή, προκειμένου να εντοπιστεί το αρχείο από την εφαρμογή. Το εν λόγω αρχείο βίντεο θα πρέπει να έχει ληφθεί μέσω της εφαρμογής «Camera by Stela», προκειμένου να συνοδεύεται και με ανάλογο αρχείο κειμένου που περιέχει τις τιμές του GPS που αναφέρονται στη θέση της κινητής συσκευής, κατά τη διάρκεια λήψης του βίντεο.

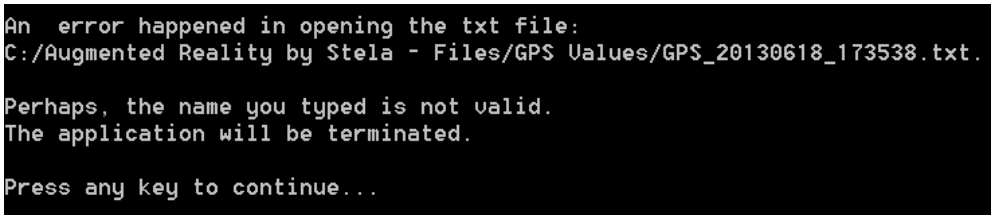
Εικόνα Δ-5. Μήνυμα που καλεί το χρήστη της Εφαρμογής 2 να πληκτρολογήσει το όνομα του αρχείου βίντεο, του οποίου το περιεχόμενο επιθυμεί να επαυξήσει.

Επόμενο βήμα, μετά την πληκτρολόγηση του ονόματος του αρχείου βίντεο από το χρήστη, αποτελεί η εμφάνιση μηνύματος για την εισαγωγή του ονόματος του αρχείου κειμένου με τις τιμές του GPS που συνοδεύει το βίντεο, του οποίου το όνομα μόλις εισήγαγε (βλ. Εικόνα Δ-6). Το αρχείο με τις τιμές του GPS πρέπει να τοποθετηθεί στον υποφάκελο «GPS Values» του φακέλου «Augmented Reality by Stela - Files», ο οποίος πρέπει να είναι τοποθετημένος στο δίσκο C του υπολογιστή.



Εικόνα Δ-6. Μήνυμα που καλεί το χρήστη της Εφαρμογής 2 να πληκτρολογήσει το όνομα του αρχείου κειμένου που περιλαμβάνει τις τιμές του GPS.

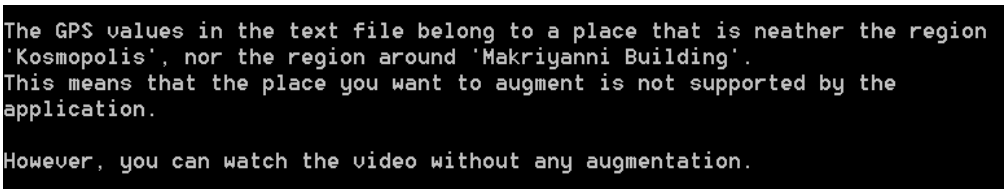
Όταν ο χρήστης πληκτρολογήσει το όνομα του αρχείου κειμένου, γίνεται προσπάθεια από το πρόγραμμα να ανοίξει το αρχείο αυτό. Αν το τελευταίο δεν ανοίξει επιτυχώς, εμφανίζεται κατάλληλο μήνυμα που ενημερώνει το χρήστη ότι το αρχείο δεν άνοιξε, λόγω πιθανότατα κάποιου λάθους κατά την πληκτρολόγηση του ονόματός του, και τον προτρέπει να πληκτρολογήσει κάποιο χαρακτήρα προκειμένου να εξασφαλιστεί ότι θα δει το μήνυμα, πριν τον τερματισμό της εφαρμογής. Όταν ο χρήστης πληκτρολογήσει κάποιο χαρακτήρα – ακολουθούμενο από το πάτημα του πλήκτρου «Enter» – η εφαρμογή τερματίζεται. Στην Εικόνα Δ-7 φαίνεται το μήνυμα που εμφανίζεται σε ανάλογη περίπτωση, κατά την οποία ο χρήστης πληκτρολόγησε το όνομα του αρχείου GPS_20130618_173538.txt το οποίο δεν υπάρχει στη διαδρομή «C:/Augmented Reality by Stela - Files/GPS Values» του υπολογιστή.



Εικόνα Δ-7. Μήνυμα σφάλματος της Εφαρμογής 2 στο άνοιγμα του αρχείου με τις τιμές του GPS.

Εφόσον το αρχείο με τις τιμές του GPS ανοίξει επιτυχώς, καθορίζεται βάσει αυτού αν το μοντέλο επαύξησης είναι το αγγείο ή το άγαλμα και εμφανίζεται στο χρήστη μήνυμα για να εξακριβωθεί αν επιθυμεί να δει τη βασική διαδικασία που διεξάγεται μέσω της OpenCV για την επαύξηση της πραγματικότητας (βλ. Εικόνα Γ-9 του Παραρτήματος Γ). Ο χρήστης καλείται να πληκτρολογήσει την απάντηση, η οποία αν είναι λανθασμένη εμφανίζεται το μήνυμα της Εικόνας Γ-10 (βλ. Παράρτημα Γ), που τον προτρέπει να πληκτρολογήσει το σωστό χαρακτήρα και εξακολουθεί να εμφανίζεται κάθε φορά που δίνει λάθος απάντηση. Στη συνέχεια, ο χρήστης καλείται να εισαγάγει τα στοιχεία του εσωτερικού προσανατολισμού της κάμερας της κινητής συσκευής μέσω της οποίας έγινε η λήψη του βίντεο, αν το επιθυμεί (βλ. Εικόνα Γ-4 του Παραρτήματος Γ). Σε περίπτωση λανθασμένης απάντησης εμφανίζεται το μήνυμα της Εικόνας Γ-5 (βλ. Παράρτημα Γ) και εφόσον ο χρήστης δήλωσε ότι θα εισαγάγει τα στοιχεία του εσωτερικού προσανατολισμού, μπορεί να αλληλεπιδράσει με την εφαρμογή, όπως φαίνεται στην Εικόνα Γ-6 (βλ. Παράρτημα Γ). Στη συνέχεια, εμφανίζονται τα αποτελέσματα της εφαρμογής επαυξημένης πραγματικότητας.

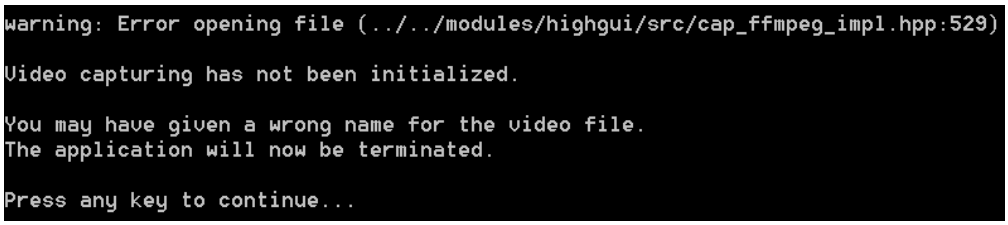
Σε περίπτωση που οι τιμές του GPS που βρίσκονται στο αρχείο που εισήγαγε ο χρήστης δεν αντιστοιχούν ούτε στην περιοχή του κτηρίου Μακρυγιάννη ούτε στην περιοχή του κτηρίου Καθαρτήρια, ο χρήστης έχει εισαγάγει βίντεο από περιοχή που δεν υποστηρίζεται από την εφαρμογή. Για το λόγο αυτό, εμφανίζεται κατάλληλο μήνυμα που τον ενημερώνει ότι με βάση τις πληροφορίες του αρχείου GPS που εισήγαγε, η περιοχή που θέλει να επαυξήσει δεν υποστηρίζεται από την εφαρμογή. Ωστόσο, ενημερώνεται ότι μπορεί να δει το βίντεο που εισήγαγε, χωρίς κάποιου είδους επαύξηση (βλ. Εικόνα Δ-8).



Εικόνα Δ-8. Μήνυμα προειδοποίησης προς το χρήστη της Εφαρμογής 2 ότι η περιοχή που επιθυμεί να επαυξήσει δεν υποστηρίζεται από την εφαρμογή.

Όταν εμφανιστεί το μήνυμα της Εικόνας Δ-8, ξεκινά η αναπαραγωγή του βίντεο που εισήγαγε ο χρήστης, μέσω της βιβλιοθήκης OpenCV. Για τη διαδικασία αυτή, απαιτείται να ανοίξει το αρχείο. Στην περίπτωση που αυτό δεν είναι εφικτό (συνήθως λόγω λανθασμένης πληκτρολόγησης του ονόματος του αρχείου) εμφανίζεται κατάλληλη προειδοποίηση τόσο

από τη βιβλιοθήκη OpenCV (πρώτη γραμμή που εμφανίζεται στην Εικόνα Δ-9), όσο και από το πρόγραμμα, που ενημερώνει το χρήστη ότι η εφαρμογή θα τερματιστεί και τον προτρέπει να πληκτρολογήσει οποιοδήποτε χαρακτήρα για το κλείσιμο της εφαρμογής, προκειμένου να εξασφαλιστεί ότι αυτός θα δει το μήνυμα.



Εικόνα Δ-9. Μήνυμα σφάλματος της Εφαρμογής 2 κατά το άνοιγμα του αρχείου βίντεο.

Αν το αρχείο βίντεο ανοίξει επιτυχώς, αυτό αναπαράγεται σε παράθυρο με τίτλο «Input Video» και η εφαρμογή τερματίζεται όταν το βίντεο τελειώσει ή αν ο χρήστης πατήσει το πλήκτρο Escape ενώ το παράθυρο «Input Video» είναι ενεργό. Για παράδειγμα, στην Εικόνα Δ-10 φαίνεται το εν λόγω παράθυρο στο οποίο αναπαράγεται ένα βίντεο που εικονίζει περιοχή η οποία – με βάση τις τιμές του αρχείου GPS που εισήχθει μαζί με το βίντεο – δεν υποστηρίζεται από την εφαρμογή.

Εικόνα Δ-10. Παράθυρο «Input Video» της Εφαρμογής 2 στο οποίο αναπαράγεται ένα αρχείο βίντεο που δεν προορίζεται για επαύξηση από την Εφαρμογή.

Δ.3 Επιπρόσθετες Διευκρινίσεις

Αν το αρχείο βίντεο – είτε αυτό που εισήγαγε ο χρήστης, μέσω πληκτρολόγησης του ονόματός του, είτε ένα από τα δύο προεπιλεγμένα, σε περίπτωση που αυτός επέλεξε να παρακολουθήσει μία προεπισκόπηση της εφαρμογής – δεν ανοίξει επιτυχώς, εμφανίζεται κατάλληλο προειδοποιητικό μήνυμα για τον επερχόμενο τερματισμό της εφαρμογής και ο χρήστης καλείται να πληκτρολογήσει ένα χαρακτήρα, ακολουθούμενο από το πλήκτρο «Enter» προκειμένου η εφαρμογή να τερματιστεί (βλ. Εικόνα Γ-15 του Παραρτήματος Γ).

Επίσης, στην περίπτωση ανεπιτυχούς αρχικοποίησης της βιβλιοθήκης GLEW, εμφανίζεται κατάλληλο μήνυμα που ειδοποιεί για τον τερματισμό της εφαρμογής, μετά από την

πληκτρολόγηση κάποιου χαρακτήρα από το χρήστη, ακολουθούμενο από το πλήκτρο «Enter» (βλ. Εικόνα Γ-16 του Παραρτήματος Γ).

Αν ο χρήστης έχει επιλέξει να παρακολουθήσει τη βασική διαδικασία που διεξάγεται μέσω της OpenCV, μετά τον καθορισμό των επιλογών του εμφανίζεται ένα παράθυρο στο οποίο παρουσιάζεται η εικόνα - πρότυπο. Για να κλείσει το παράθυρο αυτό, ο χρήστης πρέπει να πατήσει οποιοδήποτε πλήκτρο και το εν λόγω παράθυρο να είναι ενεργό. Στη συνέχεια, εμφανίζεται ένα παράθυρο που απεικονίζει τα σημεία ενδιαφέροντος που εντοπίστηκαν στην εικόνα - πρότυπο. Ομοίως για να κλείσει το παράθυρο αυτό και να ξεκινήσει η διαδικασία επαύξησης των σκηνών του βίντεο, ο χρήστης πρέπει να πατήσει οποιοδήποτε πλήκτρο και το εν λόγω παράθυρο να είναι ενεργό.

Η εφαρμογή τερματίζεται αν ο χρήστης πατήσει το πλήκτρο «Escape» ενώ το παράθυρο «AR Window» είναι ενεργό, καθώς επίσης και αν σταματήσει το βίντεο.