



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΑΓΡΟΝΟΜΩΝ ΚΑΙ ΤΟΠΟΓΡΑΦΩΝ ΜΗΧΑΝΙΚΩΝ
ΤΟΜΕΑΣ ΤΟΠΟΓΡΑΦΙΑΣ

**Τεχνολογίες όρασης υπολογιστών για δυναμική
εκκένωση κτιρίων με χρήση οπτικών και
θερμικών δεδομένων.**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΤΟΥ

ΚΑΜΙΝΤΖΗ ΝΙΚΟΛΑ

Επιβλέπων : Νικόλαος Δουλάμης
Επ. Καθηγητής Ε.Μ.Π.

Αθήνα, Ιούλιος 2016



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΑΓΡΟΝΟΜΩΝ ΚΑΙ ΤΟΠΟΓΡΑΦΩΝ ΜΗΧΑΝΙΚΩΝ
ΤΟΜΕΑΣ ΤΟΠΟΓΡΑΦΙΑΣ

Τεχνολογίες όρασης υπολογιστών για δυναμική εκκένωση κτιρίων με χρήση οπτικών και θερμικών δεδομένων.

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΤΟΥ

ΚΑΜΙΝΤΖΗ ΝΙΚΟΛΑ

Επιβλέπων : Νικόλαος Δουλάμης
Επ. Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 20^η Ιουλίου 2016.

(Υπογραφή)

(Υπογραφή)

(Υπογραφή)

.....

.....

.....

Νικόλαος Δουλάμης

Κωνσταντίνος
Καράντζαλος

Αναστάσιος Δουλάμης

Επ. Καθηγητής Ε.Μ.Π

Επ. Καθηγητής Ε.Μ.Π.

Επ. Καθηγητής Ε.Μ.Π.

Αθήνα, Ιούλιος 2016

(Υπογραφή)

.....

ΚΑΜΙΝΤΖΗΣ ΝΙΚΟΛΑΣ

Διπλωματούχος Αγρονόμος και Τοπογράφος Μηχανικός Ε.Μ.Π.

© 2016 – All rights reserved

Περίληψη

Στην εποχή που κυριαρχεί η αυτοματοποίηση των διαδικασιών, η όραση υπολογιστών έρχεται να δώσει λύση σε πολλά προβλήματα που απασχολούν τον τομέα της ασφαλούς εκκένωσης κτιρίων. Στο πλαίσιο της παρούσας εργασίας, έχουν μελετηθεί τα δεδομένα που προέκυψαν από πειράματα, που πραγματοποιήθηκαν στο αεροδρόμιο "Ελευθέριος Βενιζέλος", σχετικά με τον τρόπο αντίδρασης των ανθρώπων σε διάφορα σενάρια έκτακτης αναγκαστικής εκκένωσης του χώρου. Με τη χρήση αισθητήρων (1 x Υπερφασματικοί, 2 x Θερμικός) έχει πραγματοποιηθεί λήψη δεδομένων από την διεξαγωγή των προαναφερθέντων σεναρίων, και τα αρχεία που προέκυψαν αποτελούν πρωτογενή δεδομένα για επεξεργασία.

Στην παρούσα διπλωματική εργασία παρουσιάζεται μια προσέγγιση για την επίλυση του προβλήματος εντοπισμού και παρακολούθησης ανθρώπινης δραστηριότητας σε αρχεία βίντεο. Αναπτύχθηκε ένας (motion-based) αλγόριθμος που χρησιμοποιεί ως μοντέλο παρασκήνιου για κάθε pixel, μια μίξη Γκαουσιανών μοντέλων, και μέσω ανάλυσης blob, εξαγει το προσκήνιο από το παρασκήνιο. Το προσκήνιο αποτελείται από τις ανιχνεύσεις τις οποίες ο αλγόριθμος αποθηκεύει ως tracks. Στη συνέχεια ο αλγόριθμος μέσω του φίλτρου Kalman παρακολουθεί την πορεία των tracks στο χώρο και στον χρόνο. Τέλος μέσω του αλγορίθμου συσταδοποίησης k-μέσων ομαδοποιούνται οι ανιχνεύσεις για την καλύτερη εποπτεία της ροής των ανθρώπων.

Παρατίθεται λεπτομερής περιγραφή των διαδικασιών που χρησιμοποιήθηκαν στον προτεινόμενο αλγόριθμο, καθώς και εκτίμηση της αξιοπιστίας του.

Λέξεις Κλειδιά: <<Όραση Υπολογιστών, αυτοματοποίηση, έκτακτη αναγκαστική εκκένωση, υπερφασματικοί, θερμικός, Γκαουσιανά μοντέλα, ανάλυση blob, παρασκήνιο, προσκήνιο, motion based αλγόριθμος, ανίχνευση, tracks, παρακολούθηση, ανθρώπινη δραστηριότητα, επεξεργασία εικόνας, αλγόριθμος συσταδοποίησης k-μέσων, φίλτρο Kalman>>

Abstract

In this age of automated processes, computer vision provides the solution to a great number of problems in the building evacuation planning section. This thesis regards the analysis and processing of data resulting from experiments conducted at the "Eleftherios Venizelos" Airport, concerning the way people react in the event of an emergency evacuation. Through the use of sensors (1 x Hyperspectral, 2 x Thermal) data illustrating such scenarios has been acquired, and it poses a challenge as a new type of data that requires processing.

This thesis works on the subject of detecting and tracking the human activity in digital images. The developed motion based algorithm uses a Gaussian mixture model and via blob analysis, extract the foreground from the background. The foreground is consists of detections that the algorithm stores as tracks. The algorithm uses a process of tracking objects in order to determine the position of these objects, before observing their movement route in a video(sequence of images), through the use of a Kalman filter. It then classifies the objects via k-means clustering algorithm in order to provide the observer a clearer picture of the crowd flow.

A detailed description of the processes used in the proposed algorithm is followed, as well as an estimation of its accuracy and liability.

Keywords: <<automated processes, computer vision, emergency evacuation, Hyperspectral, Thermal, detection, tracking, algorithm, Gaussian mixture model, background, foreground, clustering, k-means, Kalman Filter>>

Ευχαριστίες

Αρχικά θέλω να ευχαριστήσω τον επιβλέποντα καθηγητή μου κ. Νικόλαο Δουλάμη για την ανάθεση αυτής της Διπλωματικής και την βοήθεια που μου έδωσε κατά τη διάρκεια αυτής. Πάνω απ' όλα τον ευχαριστώ γιατί αποτέλεσε τον κύριο λόγο για να ασχοληθώ και πάλι με την πληροφορική.

Θέλω να ευχαριστήσω τους συμφοιτητές μου που σήμερα ονομάζω φίλους, για τη βοήθεια και συνεργασία τους στα πέντε χρόνια φοίτησης μου στο Πολυτεχνείο.

Τέλος, δεν μπορώ να βρω λέξεις για να εκφράσω την ευγνωμοσύνη μου προς τους γονείς μου. Η εμπιστοσύνη και η αλτρουιστική στήριξη τους είναι πραγματικά συγκινητική. Η διπλωματική αυτή είναι αφιερωμένη σε αυτούς.

Περιεχόμενα

Περίληψη.....	5
Abstract	7
Ευχαριστίες.....	8
Περιεχόμενα	9
Περιεχόμενα Εικόνων.....	11
1 Εισαγωγή	13
1.1 Δυναμικά συστήματα εκκένωσης κτιρίων.....	14
1.2 Γενικά για την όραση υπολογιστών	18
1.3 Ανίχνευση και παρακολούθηση πλήθους με τη βοήθεια τεχνικών όρασης υπολογιστών.....	19
1.4 Αντικείμενο διπλωματικής.....	20
1.5 Διάρθρωση της εργασίας	21
2 Σχετικές εργασίες	23
2.1 Παρακολούθηση πολλαπλών αντικειμένων μέσω εκτίμησης παρασκηνίου σε αλληλουχία υπερφασματικού βίντεο	23
2.2 Ανίχνευση ανθρώπων και αναγνώριση δραστηριοτήτων μέσω όρασης υπολογιστών	27
2.3 Μοντέλα μίξης προσαρμοστικών παρασκηνίων (adaptive backgrounds) για παρακολούθηση σε πραγματικό χρόνο.....	30
3 Θεωρητικό υπόβαθρο	33
3.1 Μέθοδος Otsu	33
3.1.1 Διατύπωση μεθόδου	34
3.2 Φίλτρο Kalman	38
3.2.1 Γενικά για το φίλτρο Kalman.....	38
3.2.2 Μαθηματική επεξήγηση	39
3.2.3 Αλγόριθμος Kalman.....	41
3.3 Αλγόριθμος συσταδοποίησης k-μέσων.....	43
3.3.1 Γενικά για τον αλγόριθμο k-μέσων.....	43
3.3.2 Βήματα αλγόριθμου.....	44
4 Πειραματική διάταξη.....	47
4.1 Γενικά.....	47
4.2 Δεδομένα από θερμικό αισθητήρα	48

4.2.1. Θεωρία λειτουργίας.....	48
4.2.2 Θερμικός αισθητήρας Tau 2.	49
4.2.3 Ψηφιακά αρχεία δεδομένων	50
4.3 Δεδομένα από υπερφασματικό αισθητήρα (οπτικά δεδομένα)	51
4.3.1: Υπερφασματικός αισθητήρας CMOSIS CMV2000	51
4.3.2 Ψηφιακά αρχεία δεδομένων	52
5 Ανάπτυξη αλγορίθμων για ανίχνευση και παρακολούθηση ανθρώπινης δραστηριότητας.....	55
5.1 Εισαγωγή.....	55
5.2 Αλγόριθμος 1.....	56
5.2.1 Γενικά	56
5.2.2 Αναλυτική επεξήγηση αλγορίθμου.....	57
5.2.3 Αποτελέσματα - Σχολιασμός.....	61
5.3 Αλγόριθμος 2.....	63
5.3.1 Γενικά	63
5.3.2 Αναλυτική επεξήγηση αλγορίθμου.....	64
5.3.3 Αποτελέσματα - Σχολιασμός.....	76
5.3.3.1 Σύγκριση αποτελεσμάτων μεταξύ θερμικών και οπτικών δεδομένων	86
6 Συσταδοποίηση ανιχνεύσεων και παρακολούθηση των συστάδων που δημιουργούνται.....	89
6.1 Εφαρμογή αλγορίθμου στα θερμικά δεδομένα	90
7 Συμπεράσματα και μελλοντικές κατευθύνσεις.....	95
7.1 Συμπεράσματα	95
7.2 Μελλοντικές κατευθύνσεις	98
Βιβλιογραφία.....	101

Περιεχόμενα Εικόνων

<i>Εικ 1.1.1 : Γενική λειτουργία συστήματος eVACUATE.....</i>	<i>17</i>
<i>Εικ. 2.1.1 : Διάγραμμα απεικόνισης αλγορίθμου για αυτόματη ανίχνευση και παρακολούθηση πολλαπλών αντικειμένων σε ακολουθία υπερφασματικών video [7]</i>	<i>26</i>
<i>Εικ. 2.1.2: Πειραματικά αποτελέσματα μετά την εφαρμογή του αλγορίθμου σε υπερφασματικό video dataset. [7].....</i>	<i>27</i>
<i>Εικ. 2.2.1: Σχηματική επεξήγηση διαδικασίας. [9].....</i>	<i>28</i>
<i>Εικ. 2.2.2 : Εικόνα πεζών που εντοπίστηκαν και χάρτης κίνησης πεζών. [9]</i>	<i>29</i>
<i>Εικ 2.3.1 : Εκτέλεση του προγράμματος. [10] α) παρούσα εικόνα, b) εικόνα που προέκυψε από τις πιο πιθανές Gaussians στο background model, c) εικόνα με τα foreground pixels, d) παρούσα εικόνα με το tracking</i>	<i>31</i>
<i>Εικ. 3.1.1 : Ιστόγραμμα επιλογής threshold Otsu Method</i>	<i>34</i>
<i>Εικ. 3.2.1: Αλγόριθμος KALMAN.....</i>	<i>42</i>
<i>Εικ 3.3.2.1 : Παρουσίαση διαδικασίας πρότυπου αλγορίθμου k-μέσων.....</i>	<i>45</i>
<i>Εικ. 4.2.2.1 : Χαρακτηριστικά θερμικού αισθητήρα Tau 2.....</i>	<i>49</i>
<i>Εικ. 4.2.2.2 : Θερμική κάμερα Tau 2 [19]</i>	<i>49</i>
<i>Εικ. 4.2.3.1 : Χαρακτηριστικά θερμικών βίντεο</i>	<i>50</i>
<i>Εικ. 4.2.3.2 : Στιγμιότυπο θερμικού βίντεο 1.....</i>	<i>50</i>
<i>Εικ. 4.2.3.3 : Στιγμιότυπο θερμικού βίντεο 2.....</i>	<i>51</i>
<i>Εικ. 4.3.1.1: Χαρακτηριστικά υπερφασματικού αισθητήρα CMOSIS CMV2000.....</i>	<i>52</i>
<i>Εικ. 4.3.1.2 : Υπερφασματική κάμερα [19].....</i>	<i>52</i>
<i>Εικ. 4.3.1.3 : Φίλτρο 16 φασματικών καναλιών [19].....</i>	<i>52</i>
<i>Εικ. 4.3.2.1 : Χαρακτηριστικά οπτικού βίντεο</i>	<i>53</i>
<i>Εικ 4.3.2.2 : Στιγμιότυπο οπτικού βίντεο.....</i>	<i>53</i>
<i>Εικ. 5.2.2.1 : Διάγραμμα γενικής περιγραφής του αλγορίθμου 1 (Otsu Thresholding Method)</i>	<i>60</i>
<i>Εικ.5.2.3.1: Αποτελέσματα αλγορίθμου σε περίπτωση ύπαρξης ανθρώπινης δραστηριότητας (θερμικό βίντεο)</i>	<i>61</i>
<i>Εικ 5.2.3.2: Αποτελέσματα αλγορίθμου στην περίπτωση μη ύπαρξης ανθρώπινης δραστηριότητας (θερμικό βίντεο)</i>	<i>62</i>

Εικ 5.2.3.3: Αποτελέσματα αλγορίθμου στο βίντεο που κατέγραψε η υπερφασματική κάμερα	62
Εικ 5.3.2.1 : Διάγραμμα γενικής περιγραφής του αλγόριθμου 2.	69
Εικ 5.3.2.2 : Διάγραμμα σταδίου δημιουργίας <i>System objects</i>	72
Εικ 5.3.3.1.1 : Αποτελέσματα <i>object detection</i> στο <i>frame #70</i>.	78
Εικ 5.3.3.1.2 : Αποτελέσματα <i>object detection</i> στο <i>frame #166</i>.	80
Εικ 5.3.3.1.3 : Αποτελέσματα <i>object detection</i> στο <i>frame #393</i>.	81
Εικ 5.3.3.2.1 : Αποτελέσματα <i>object detection</i> στο <i>frame #36</i>.	83
Εικ 5.3.3.2.2 : Αποτελέσματα <i>object detection</i> στο <i>frame #54</i>.	84
Εικ 5.3.3.2.3 : Αποτελέσματα <i>object detection</i> στο <i>frame #140</i>.	85
Εικ 5.3.3.1.1 : Παρακολούθηση (<i>tracking</i>) στο βίντεο θερμικών δεδομένων	87
Εικ 5.3.3.1.2 : Παρακολούθηση (<i>tracking</i>) στο βίντεο οπτικών δεδομένων	88
Εικ 6.1.1 : Εφαρμογή αλγορίθμου <i>k</i>-μέσων στα θερμικά δεδομένα για <i>k=2</i>	92
Εικ 6.1.2 : Εφαρμογή αλγορίθμου <i>k</i>-μέσων στα θερμικά δεδομένα για <i>k=3</i>	93
Εικ 6.1.3 : Παρουσίαση κέντρων συστάδων για <i>k=2</i>	94
Εικ 6.1.3 : Παρουσίαση κέντρων συστάδων για <i>k=3</i>	94

1

Εισαγωγή

Το θέμα της δημόσιας ασφάλειας είναι πιο επίκαιρο από ποτέ άλλοτε. Τα τελευταία χρόνια η ανθρωπότητα έρχεται αντιμέτωπη με αναρίθμητες τραγωδίες τρομοκρατικής κυρίως φύσεως, αλλά και τραγωδίες που οφείλονται σε άλλα ανθρώπινα ή μη αίτια. Το μέγεθος της ανθρωπιστικής τραγωδίας καθορίζεται από τον αριθμό των θυμάτων και των τραυματιών. Για αυτό τον λόγο χώροι στους οποίους παρατηρείται μαζική συγκέντρωση ανθρώπων είναι οι πιο εκτεθειμένοι στον κίνδυνο μιας τέτοιας τραγωδίας. Κτίρια όπως οι σταθμοί Μετρό και τραίνων, τα αεροδρόμια, τα στάδια, τα εμπορικά κέντρα, είναι χώροι που κατά διαστήματα φιλοξενούν τεράστια πλήθη ανθρώπων.

Μέχρι σήμερα έχει παρουσιαστεί σημαντική πρόοδος στο θέμα του σχεδιασμού ασφαλείας τέτοιων χώρων, ούτως ώστε να εξασφαλίζεται η ασφαλής διέλευση των ανθρώπων κατά την είσοδο, παραμονή, και έξοδο τους από αυτούς. Τα κυριότερα παράγωγα του παραπάνω σχεδιασμού είναι ο καθορισμός εξόδων ασφαλείας, τοποθέτηση συστημάτων πυροπροστασίας και

η κατάρτιση σχεδίων αναγκαστικής εκκένωσης σε περίπτωση έκτακτης ανάγκης.

Η παρούσα προσέγγιση στο θέμα του σχεδιασμού μπορεί να χαρακτηριστεί ως στατική υπό την έννοια του ότι δεν εκμεταλλεύεται δυναμικές πληροφορίες όπως η δυναμική του πλήθους κατά τη διάρκεια της εκκένωσης, η αποτελεσματικότητα των ενεργειών των υπευθύνων ασφαλείας, καθώς επίσης και τον παράγοντα ψυχολογία του πλήθους. Ο παράγοντας ψυχολογίας είναι πολύ σημαντικός για την ασφαλή εκκένωση κτηρίων σε περίπτωση έκτακτης ανάγκης, και η εκ των προτέρων πρόβλεψη του (στο στάδιο του καταρτισμού σχεδίου ασφαλείας) είναι σχεδόν αδύνατη.

Σε περίπτωση έκτακτης ανάγκης ή κρίσιμης κατάστασης, η αίσθηση του εγκλεισμού δημιουργεί μια κατάσταση ψυχικής δυσφορίας και πανικού στο πλήθος. Ο συνωστισμός που παρατηρείται όσο πλησιάζουμε προς τις εξόδους επιφέρει προβλήματα στην ελεύθερη και ομαλή κυκλοφορία με αποτέλεσμα να δημιουργείται μια κατάσταση φυσικής ανυπομονησίας. Τέτοιες καταστάσεις μπορεί να οδηγήσουν μερικά από τα μέλη του πλήθους σε πανικό, με αποτέλεσμα να μην ακολουθούν τα πρωτόκολλα εκκένωσης. Σε περίπτωση πανικού μπορεί να παρουσιαστούν προβλήματα σε χώρους με περιορισμένη χωρητικότητα όπως στις σκάλες και τις θύρες, δημιουργώντας έτσι πιθανώς μια ακόμη πιο επικίνδυνη κατάσταση μέσα σε μια κατάσταση έκτακτης ανάγκης.

Για την πιο αποτελεσματική αντιμετώπιση τέτοιων καταστάσεων πρέπει να καταρτιστεί ένα σχέδιο που θα διαχειρίζεται δυναμικά την κατάσταση. Με την βοήθεια της τεχνολογίας (αισθητήρες, εργαλεία επικοινωνίας και πρόβλεψης), το ανθρώπινο δυναμικό (που είναι υπεύθυνο για την εκτέλεση του σχεδίου) θα είναι σε θέση να κατευθύνει το πλήθος προς ασφαλείς εξόδους όσο πιο γρήγορα και όσο πιο ήρεμα γίνεται. Προς αυτή την κατεύθυνση κινούνται οι μελέτες ανάπτυξης δυναμικών συστημάτων εκκένωσης κτιρίων.

1.1 Δυναμικά συστήματα εκκένωσης κτιρίων

Η σημερινή πρακτική που εφαρμόζεται αναφορικά με την ασφαλή εκκένωση κτιρίων στην περίπτωση παρουσίας κινδύνου είναι η κατάρτιση σχεδίων εκκένωσης για διάφορα σενάρια. Ανάλογα με τη φύση του κινδύνου (φωτιά,

βόμβα, σεισμός) τα σχέδια αυτά προτείνουν διαδρόμους (routes) που πρέπει να ακολουθήσει το πλήθος έτσι ώστε να εκκενωθεί το κτίριο με ασφάλεια.

Σε μια κατάσταση έκτατου κινδύνου οι μεταβλητές του συστήματος - σχεδίου εκκένωσης είναι πάρα πολλές και η εκ των προτέρων πρόβλεψη τους είναι σχεδόν αδύνατη. Η βασική ιδέα της ανάπτυξης δυναμικών συστημάτων για ασφαλή εκκένωση κτιρίων είναι η εκμετάλλευση της πληροφορίας που μπορεί να εξαχθεί σε πραγματικό χρόνο από αισθητήρες που υπάρχουν στο κτίριο. Η αποτελεσματικότητα ενός συστήματος είναι ανάλογη με την ποσότητα των πληροφοριών που λαμβάνει.

Η ανάπτυξη της τεχνολογίας με την είσοδο των έξυπνων χώρων (smart spaces), των δικτύων αισθητήρων (sensor networks) και της υπολογιστικής δυνατότητας που παρουσιάζουν οι σύγχρονοι Η/Υ, παρέχουν την δυνατότητα ανάπτυξης δυναμικών συστημάτων με σκοπό την εκκένωση κτιρίων. Η χρήση των παραπάνω τεχνολογιών κάτω από το σωστό πλαίσιο διαδικασιών μπορεί να προσφέρει τεράστια βοήθεια στους φορείς λήψεως αποφάσεων (υπεύθυνους για την εκκένωση). Με την ροή πληροφοριών σε σχεδόν πραγματικό χρόνο, σε σύγκριση με την παρούσα κατάσταση κινδύνου, βοηθάει τους φορείς λήψεως αποφάσεων να αντιληφτούν καλύτερα την υπάρχουσα κατάσταση, γεγονός που συμβάλει στην επιλογή της βέλτιστης στρατηγικής που πρέπει να ακολουθηθεί για την ασφαλή εκκένωση του κτιρίου.

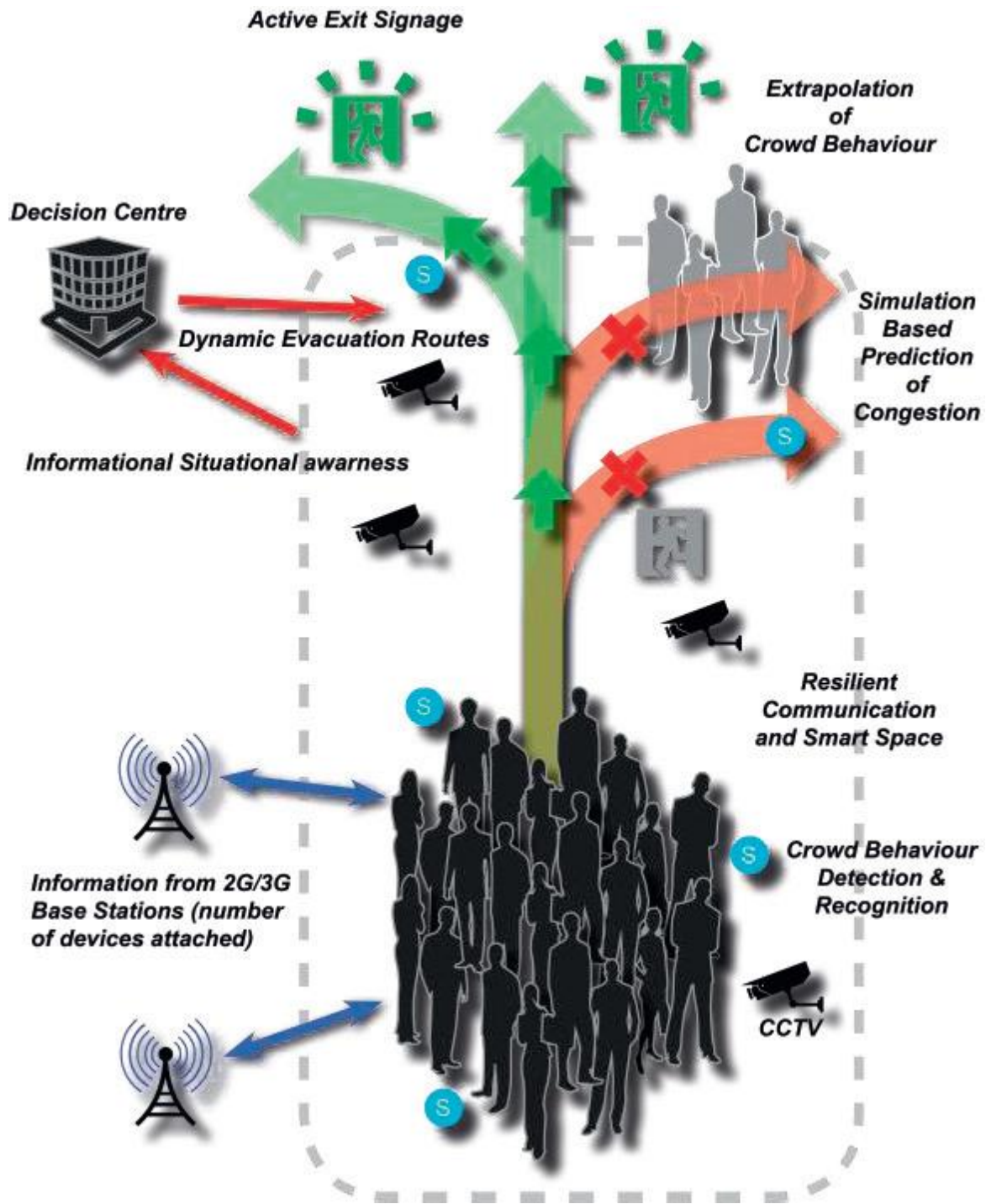
Προς την κατεύθυνση αυτή έχουν αναπτυχθεί διάφορα δυναμικά συστήματα εκκένωσης κτιρίων, ένα εξ αυτών είναι το eVACUATE. Η ομάδα του eVACUATE ανέπτυξε ένα σύστημα το οποίο είναι ολιστικό (holistic), έχει επίγνωση της κατάστασης και είναι ανεξάρτητο σεναρίου. Σκοπός του είναι η διατήρηση ενεργών διαδρόμων για την ασφαλή εκκένωση του κτιρίου από μεγάλα πλήθη. Προβλέπει την κίνηση του πλήθους κατά τη διάρκεια της εκκένωσης και παρέχει αυτή την πληροφορία στους φορείς λήψεως αποφάσεων. Τα δεδομένα που καταγράφονται - λαμβάνονται από τους διάφορους αισθητήρες που είναι τοποθετημένοι στο κτίριο, επεξεργάζονται μέσω του μοντέλου πλήθους (crowd model) που περιγράφει καλύτερα την κατάσταση, προβλέπει την τοποθεσία και τον χρόνο που θα παρατηρηθεί ενδεχόμενη κυκλοφοριακή συμφόρηση καθώς και τον χρόνο που απαιτείται για να ολοκληρωθεί η εκκένωση. Στη συνέχεια αναζητείται η χρονικά εξαρτημένη, βέλτιστη διαδρομή (route) η οποία ελαχιστοποιεί τις παραπάνω

μεταβλητές και αποτελεί τη βέλτιστη στρατηγική εκκένωσης για την εκάστοτε κατάσταση που πρέπει να αντιμετωπιστεί.

Οι απαιτήσεις για την αποτελεσματικότητα ενός δυναμικού συστήματος εκκένωσης κτιρίων μπορεί να συνοψιστούν στα παρακάτω σημεία:

- Επικοινωνία συστήματος με αισθητήρες και ροή δεδομένων σε πραγματικό χρόνο.
- Πρόβλεψη κυκλοφοριακής συμφόρησης βασισμένη στην παρούσα κατάσταση.
- Παροχή αναπληροφόρησης (feedback) σε σχεδόν πραγματικό χρόνο (close to real-time)
- Καθορισμό βέλτιστης διαδρομής εκκένωσης (βασισμένη στις παραμέτρους απόσταση, χρόνος, συμφόρηση)

Στη συνέχεια παρατίθεται μια εικόνα που περιγράφει την γενική λειτουργία του συστήματος eVACUATE.



Εικ 1.1.1 : Γενική λειτουργία συστήματος eVACUATE

1.2 Γενικά για την όραση υπολογιστών

Η Όραση Υπολογιστών είναι ο επιστημονικός τομέας που μελετά μεθόδους ανάλυσης και κατανόησης εικόνας ή ακολουθίας εικόνων με στόχο την εξαγωγή συμβολικής πληροφορίας. Οι αλγόριθμοι της Όρασης Υπολογιστών παίρνουν σαν είσοδο, αριθμητικά δεδομένα, π.χ. τιμές απο pixels εικόνων, και σαν έξοδο παράγουν συμβολική πληροφορία, π.χ. εντοπισμός αντικειμένων. Σε αντίθεση με τον τομέα των Γραφικών Υπολογιστών, που προσπαθεί να επιλύσει το πρόβλημα της μετατροπής συμβόλων σε εικόνα, η Όραση Υπολογιστών προσπαθεί να λύσει το αντίστροφο πρόβλημα, δηλαδή από εικόνες να εξάγει συμπεράσματα για το είδος των μεταβλητών ενδιαφέροντος που υπάρχουν σε μια σκηνή (σχήμα, μέγεθος, υφή, χρώμα, ταυτότητα), τη θέση τους (τοποθεσία, κίνηση) καθώς και τις σχέσεις και ομοιότητες μεταξύ τους. Παραδείγματα μεταβλητών ενδιαφέροντος είναι τα αντικείμενα, οι άνθρωποι, τα πρόσωπα, οι χειρονομίες ή οι ανθρώπινες δράσεις.

Οι απαρχές της Όρασης Υπολογιστών μπορούν να εντοπιστούν στη δεκαετία του 1960 και στις προσπάθειες ερευνητών από το χώρο της Τεχνητής Νοημοσύνης να κατασκευάσουν υπολογιστές που αντιλαμβάνονται τον ορατό κόσμο όπως οι άνθρωποι. Παρά την πρόοδο που έχει επιτευχθεί τα προηγούμενα χρόνια στον τομέα, απέχουμε πολύ από την κατασκευή ενός συστήματος που θα μπορεί να κατονομάσει και να σκιαγραφήσει όλα τα αντικείμενα που απεικονίζονται σε μια φωτογραφία με την άνεση που θα τα αναγνώριζε ακόμα και ένα μικρό παιδί.

Η Όραση Υπολογιστών είναι ένα διεπιστημονικό πεδίο που συνδυάζει μεθόδους και αρχές από επιστημονικές περιοχές όπως η Επεξεργασία Σημάτων, η Αναγνώριση Προτύπων, η Τεχνητή Νοημοσύνη, τα Εφαρμοσμένα Μαθηματικά και Φυσική, αλλά και η Νευροβιολογία και Ψυχολογία. Η Όραση Υπολογιστών βρίθει εφαρμογών. Ενδεικτικά βρίσκει εφαρμογή στην επεξεργασία εικόνων, τη ρομποτική, βιοϊατρική τεχνολογία και ιατρική απεικόνιση, την επικοινωνία ανθρώπου υπολογιστή, την τηλεπισκόπηση, τα ευφυή (έξυπνα) συστήματα, τον κινηματογράφο και την τεχνολογία βίντεο.

Κάποια ενδεικτικά προβλήματα που επιλύονται με μεθόδους της όρασης υπολογιστών είναι:

- η αναγνώριση αντικειμένων
- η ανίχνευση και εκτίμηση κίνησης
- η ανακατασκευή τρισδιάστατης δομής
- η κατάτμηση εικόνων
- η ανίχνευση προσώπου
- η ανίχνευση και κατηγοριοποίηση ανθρώπινων δράσεων
- η μοντελοποίηση και ανάλυση υφής, χρώματος και σχήματος.

1.3 Ανίχνευση και παρακολούθηση πλήθους με τη βοήθεια τεχνικών όρασης υπολογιστών

Τις τελευταίες δεκαετίες έχουν αναπτυχθεί πολλές εφαρμογές στον τομέα της όρασης υπολογιστών που εμπλέκουν ανάλυση εικόνων που περιέχουν ανθρώπους. Αυτός ο τομέας καλύπτει, μεταξύ άλλων, την αναγνώριση προσώπων, αναγνώριση χειρονομιών και την ανίχνευση ανθρώπινων φιγούρων. Το τεράστιο ενδιαφέρον που παρουσιάζουν αυτές οι εφαρμογές οφείλεται στην επιθυμία βελτίωσης της αλληλεπίδρασης ανθρώπου - μηχανής [1-6].

Μια απ' αυτές (τις εφαρμογές) είναι η αυτόματη καταμέτρηση ανθρώπων, η οποία έχει εφαρμογή σε πολλούς τομείς όπως η παρακολούθηση (surveillance), ασφάλεια και έλεγχος πλήθους (safety monitoring and crowd control), στη πρόσβαση σε κέντρα ασφαλείας (security access control), σε έρευνες που αφορούν την κίνηση πεζών κ.α. Ανάλογα βέβαια με τον σκοπό, διαφέρει και η ακρίβεια των αποτελεσμάτων του εκάστοτε συστήματος.

Αξίζει να σημειωθεί ότι πέρα από την ανάλυση βίντεο, προσφέρονται και άλλες τεχνολογίες για τη λύση του προβλήματος του εντοπισμού και της καταμέτρησης των ανθρώπων. Τέτοιες τεχνολογίες είναι αυτές που χρησιμοποιούν αισθητήρες που βασίζονται στις ακτίνες φωτός, τα μικροκύματα ή τους υπερήχους.

Οι αισθητήρες απεικόνισης βίντεο (κάμερες) έχουν αποδειχτεί οι πλέον κατάλληλοι, λόγω του εύλογου κόστους, της μεγάλης περιοχής κάλυψης και της ταχύτητας επεξεργασίας τους. Το αυξανόμενο ενδιαφέρον προς τις τεχνικές ψηφιακής επεξεργασίας εικόνας ως μέσο για επίλυση του προβλήματος του εντοπισμού της ανθρώπινης δραστηριότητας οφείλεται επίσης στο γεγονός ότι οι βιντεοκάμερες και οι τεχνικές ανάλυσης εικόνας, παρουσιάζουν μια βολική, μη προσβλητική και αποδοτική σε σχέση με το κόστος, λύση.

Ένα τεράστιο πρόβλημα που αντιμετωπίζει η παρούσα τεχνολογία είναι στο πώς να γίνεται η ανίχνευση και ο διαχωρισμός μεταξύ των ανθρώπων σε απεικονίσεις βίντεο. Το πρόβλημα οφείλεται στο γεγονός ότι οι πεζοί είναι μη άκαμπτα αντικείμενα που παρουσιάζουν μεγάλη ανομοιομορφία σε μέγεθος, σχήμα και δομή. Η λύση αυτού του προβλήματος αποτελεί πρόκληση για πολλούς ερευνητές, οι οποίοι αναζητούν καινούριες μεθόδους λύσης.

1.4 Αντικείμενο διπλωματικής.

Η παρούσα εργασία ασχολείται με το κομμάτι της ανίχνευσης και παρακολούθησης της ανθρώπινης δραστηριότητας, στο πλαίσιο του γενικού προβλήματος ανάπτυξης δυναμικών συστημάτων εκκένωσης κτιρίων. Δεδομένα της εργασίας υπήρξαν τρία αρχεία βίντεο από τη διενέργεια πειραματικών σεναρίων έκτακτης εκκένωσης στο πλαίσιο του έργου eVacuate. Τα δύο εξ αυτών των αρχείων λήφθηκαν από θερμική κάμερα, και το προϊόν τους είναι μια θερμική απεικόνιση σε grayscale χρωματική παλέτα. Το τρίτο αρχείο αν και λήφθηκε από υπερφασματική κάμερα, τα δεδομένα του θεωρούνται οπτικά καθώς η καταγραφή έγινε στα τρία κανάλια του οπτικού φάσματος (στα 465 nm 546 nm και 630 nm). Και στην περίπτωση του οπτικού αρχείου, η απεικόνιση είναι επίσης σε χρωματική παλέτα grayscale.

Επιχειρήθηκε η ανάπτυξη ενός αλγορίθμου σε περιβάλλον Matlab ο οποίος να εντοπίζει με αποτελεσματικότητα τα αντικείμενα (tracks) σε κάθε frame ενός βίντεο. Στη συνέχεια καταγράφει πληροφορίες (track id, centroid, bounding box, frame στο οποίο εμφανίζεται) για κάθε αντικείμενο σε μια δομή (structure) η οποία ανάλογα με κάποιους κανόνες εκσυγχρονίζει το περιεχόμενο της frame by frame.

Χρησιμοποιώντας πληροφορίες απο την δομή των tracks κατασκευάζεται ένα dataset το οποίο αποτελεί την χρονοσειρά του συνόλου των αντικειμένων που εμφανίστηκαν στη διάρκεια του βίντεο. Το dataset αποτελεί ένα 3D πίνακα που περιέχει πληροφορίες για το κάθε track (track id,centroid x, centroid y), δηλαδή την τοποθεσία που είχε το κάθε track σε κάθε frame.

Στον τομέα του σχεδιασμού βέλτιστων διαδρομών για εκκένωση κτιρίων η προσοχή επικεντρώνεται στην μελέτη της κίνησης του πλήθους σε ομάδες ατόμων και όχι την κίνηση των ανθρώπων σε μονάδες. Για το σκοπό αυτό εφαρμόστηκαν αλγόριθμοι ταξινόμησης στα tracks κάθε frame για την εξαγωγή εικόνων που παρουσιάζουν τον χώρο και τον τρόπο που διαρπάζεται το πλήθος σε ομάδες. Έτσι μπορεί να μελετηθεί η εξέλιξη αυτών των ομάδων στον χώρο και στον χρόνο, που είναι ένα απο τα ζητούμενα των δυναμικών συστημάτων εκκένωσης κτιρίων.

1.5 Διάρθρωση της εργασίας

Στην ενότητα αυτή, γίνεται η περιγραφή της οργάνωσης της παρούσας εργασίας, ενώ γίνεται και μια συνοπτική περιγραφή κάθε κεφαλαίου.

- Στο κεφάλαιο 2 πραγματοποιείται μια συνοπτική επισκόπηση ενός ενδεικτικού συνόλου προσεγγίσεων και εργασιών πάνω στο πρόβλημα της ανίχνευσης και παρακολούθησης ανθρώπινης δραστηριότητας μέσω τεχνικών όρασης υπολογιστών σε ακολουθίες εικόνων (βίντεο).
- Στο κεφάλαιο 3 γίνεται η εμπειριστατωμένη επεξήγηση των μαθηματικών εργαλείων / μεθόδων που χρησιμοποιήθηκαν.
- Το κεφάλαιο 4 αφορά την πειραματική διάταξη. Στο συγκεκριμένο κεφάλαιο αναφέρεται η προέλευση των δεδομένων καθώς και η περιγραφή των τεχνικών τους χαρακτηριστικών.
- Το κεφάλαιο 5 αποτελεί το σημαντικότερο κεφάλαιο της διπλωματικής. Στο κεφάλαιο αυτό γίνεται η παρουσίαση και η επεξήγηση των αλγορίθμων που χρησιμοποιήθηκαν για να λύσουν το πρόβλημα της ανίχνευσης και παρακολούθησης της ανθρώπινης δραστηριότητας. Στο

τελευταίο μέρος του κεφαλαίου παρουσιάζονται τα αποτελέσματα εφαρμογής του δεύτερου αλγορίθμου.

- Το κεφάλαιο 6 αφορά την εφαρμογή του αλγορίθμου συσταδοποίησης k-μέσων για τη μελέτη της συμπεριφοράς του πλήθους, όταν αυτό διαχωριστεί σε ομάδες.
- Στο κεφάλαιο 7 πραγματοποιείται μια γενική αξιολόγηση της εργασίας και προτείνεται ένα σύνολο κατευθύνσεων για πιθανές μελλοντικές επεκτάσεις και βελτιώσεις αυτής.

2

Σχετικές εργασίες

2.1 Παρακολούθηση πολλαπλών αντικειμένων μέσω εκτίμησης παρασκηνίου σε αλληλουχία υπερφασματικού βίντεο

Ο Ζ. Κανδυλάκης και οι συνεργάτες του [7] σχεδίασαν, ανέπτυξαν και αξιολόγησαν ένα γενικό πλαίσιο εργασιών με σκοπό την παρακολούθηση πολλαπλών αντικειμένων σε υπερφασματική αλληλουχία βίντεο. Η εκτίμηση του παρασκηνίου (background) αντιμετωπίστηκε αποτελεσματικά μέσω προηγμένου φιλτραρίσματος κλίμακας χώρου και μείωσης διαστάσεων (advanced scale space filtering and dimensionality reduction). Για σκοπούς μείωσης της υπολογιστικής πολυπλοκότητας, η ανίχνευση των αντικειμένων γίνεται στην μειωμένων πλέον διαστάσεων αναπαράσταση του εκάστοτε στιγμιότυπου. Η ανίχνευση των αντικειμένων επιτυγχάνεται μέσω κανόνων ταξινόμησης που βασίζονται σε συγκεκριμένα γεωμετρικά και φασματικά χαρακτηριστικά των αντικειμένων.

Η παραπάνω προσέγγιση θέλει να εκμεταλλευτεί τα πλεονεκτήματα που προσφέρουν οι υπερφασματικοί αισθητήρες απεικόνισης. Η αλληλουχία υπερφασματικών βίντεο με υψηλή χρονική ανάλυση (high temporal resolution), συνδυάζει τα πλεονεκτήματα τόσο του βίντεο, όσο και της υπερφασματικής απεικόνισης. Παρά την υψηλή χωρική ανάλυση που έχουν οι σύγχρονες RGB κάμερες ή οι μονοχρωματικές βιντεοκάμερες, οι δυνατότητες τους σχετικά με την ανίχνευση και την κατηγοριοποίηση αντικειμένων είναι περιορισμένες λόγω του στενού εύρους φασματικής κάλυψης που έχουν. Οι υπερφασματικοί αισθητήρες προσφέρουν επαναλαμβανόμενα, συνεπή και περιεκτικά datasets με αυξημένη δυνατότητα διάκρισης των διαφόρων χαρακτηριστικών των αντικειμένων.

Στο διάγραμμα που ακολουθεί (Εικ 2.1.1) παρουσιάζεται ο αλγόριθμος που αναπτύχθηκε. Τα διάφορα μέρη του αλγορίθμου συνεισφέρουν στην επεξεργασία του αρχικού υπερκύβου (hypercube) και παραδίδουν τα αποτελέσματα της ανίχνευσης των αντικειμένων σε κάθε στιγμιότυπο της αλληλουχίας βίντεο.

Η μεθοδολογία που ακολουθήθηκε χωρίζεται σε τρεις επιμέρους ενότητες : i) της συνεγράφης φασματικών καναλιών, ii) του υπερφασματικού φιλτραρίσματος κλίμακας χώρου και μείωσης διαστάσεων, και iii) της αφαίρεσης παρασκηνίου - ανίχνευση πολλαπλών κινούμενων αντικειμένων.

I. Συν - εγγραφή φασματικών καναλιών (Co - registration of spectral bands):

Σε πρώτο στάδιο πρέπει να αποκατασταθεί η γεωμετρία στα υπερφασματικά κανάλια σε κάθε frame. Αυτό επιτυγχάνεται μέσω μιας αυτόματης διαδικασίας εγράφης βασισμένης σε περιγραφικά χαρακτηριστικά [8]. Το φάσμα μήκους κύματος στο οποίο γίνεται η επεξεργασία, χωρίζεται σε ένα κατάλληλο αριθμό φασματικών ομάδων. Στη συνέχεια γίνεται η συν-εγγραφή όλων των φασματικών καναλιών για κάθε φασματική ομάδα. Μ' αυτό τον τρόπο η αναζήτηση σε ένα ευρύ χώρο λύσεων, περιορίζεται λόγω της παραπάνω επιλογής φασματικών καναλιών, γεγονός που κάνει τη διαδικασία αποκατάστασης της γεωμετρίας πολύ πιο γρήγορη.

II. Υπερφασματικό φιλτράρισμα κλίμακας χώρου και μείωση διαστάσεων (Hyperspectral scale space filtering and dimensionality reduction)

Σε κάθε frame τα φασματικά κανάλια καταγράφουν ένα υπερκύβο (I) ο οποίος απλοποιείται μέσω μιας ανισοτροπικής μορφολογικής επιπεδοποίησης (anisotropic morphological leveling). Εν ολίγης, με τη χρήση μη γραμμικού φιλτραρίσματος κλίμακας χώρου (non-linear scale space filtering) επιτυγχάνεται η αφαίρεση του θορύβου, ενώ τα δεδομένα είναι αποτελεσματικά απλοποιημένα.

Με τη σύνδεση του αρχικού υπερκύβου I με μια σειρά από δείκτριες συναρτήσεις $\{h_1, h_2, \dots, h_n\}$ μπορούμε να πετύχουμε πολυκλιματική επιπεδοποίηση (multiscale leveling) - όλες οι h είναι αυξανόμενα ομαλότεροι υπερκύβοι στον R^3 . Η επιπεδοποίηση που δημιουργείται είναι μια σειρά με μορφή : $g_n = \Lambda(g_{n-1}, h_{n-1})$.

Η σειρά g_n με όλο και πιο απλούς υπερκύβους, με όλο και λιγότερες ομαλές ζώνες, δημιουργείται σχηματίζοντας μια 4D κλίμακα χώρου με $g: \Omega \subset R^4$ και $g(x,y,z,n) = g_n(x,y,z)$. Με αυτές τις μειωμένες ζώνες, οι απλοποιημένοι υπερκύβοι περιγράφουν, με ένα πιο σαφή τρόπο, την χωρική και φασματική υπογραφή των παρατηρήσεων, σχηματίζοντας κατάλληλες εκδώσεις για την εκτίμηση της εγγενούς διαστασιοποίησης των δεδομένων.

Με την παραπάνω διαδικασία απλοποιούνται οι υπολογιστικές διαδικασίες, και έχοντας εκτιμήσει την εγγενή διαστασιοποίηση των datasets, η μετατροπή τους αποκτά μια ουσιαστική αναπαράσταση (στη μειωμένη πλέον διάσταση).

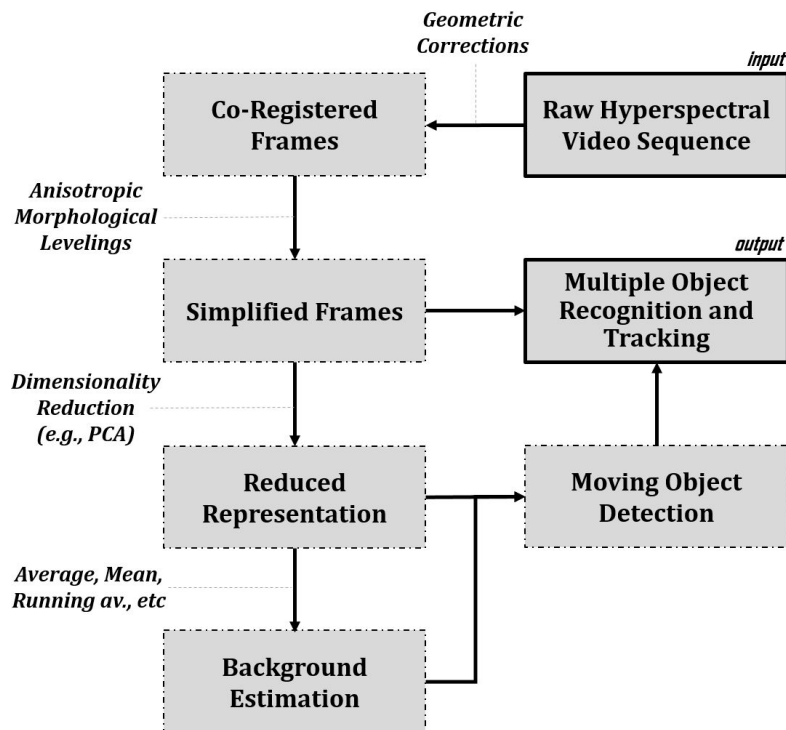
Ακολουθεί η ταξινόμηση των φασματικών καναλιών ανάλογα με τη διακύμανση και την εγγύτητα τους σε φασματικές περιοχές κλειδιά. Η συν-εγγραφή στις φασματικές ομάδες πραγματοποιείται εντός και ανάμεσα σ' αυτές τις φασματικές περιοχές κλειδιά.

III. Αφαίρεση παρασκηνίου - Ανίχνευση πολλαπλών κινούμενων αντικειμένων.

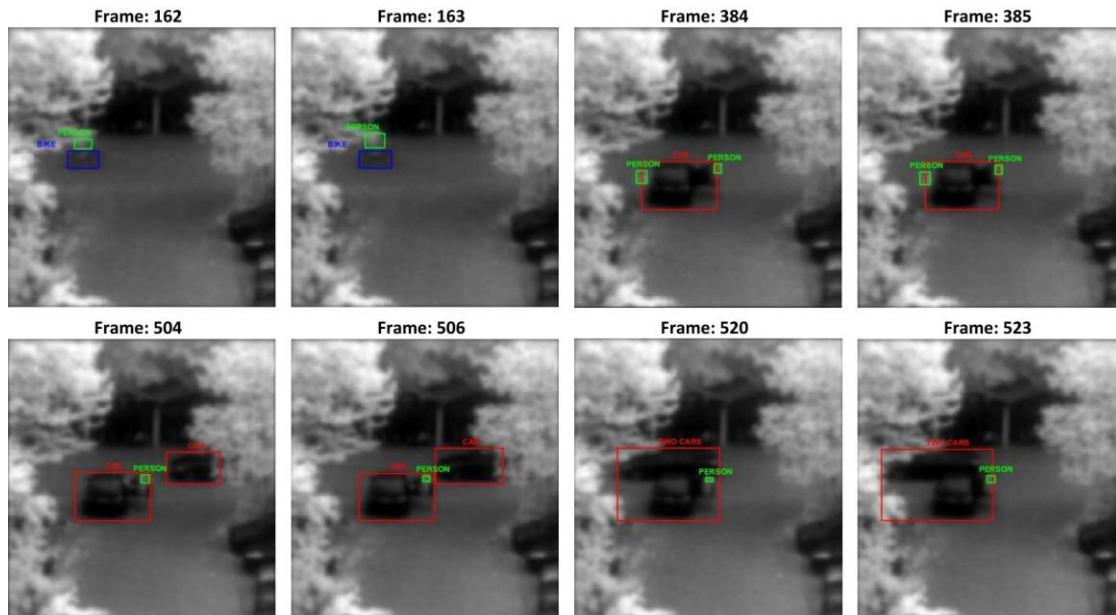
Η αρχική ανίχνευση των αντικειμένων έγινε με χρήση της μεθόδου αφαίρεσης παρασκηνίου (background subtraction) από την οποία προκύπτουν τα αρχικά seeds/blobs που τροφοδοτούν την ενότητα του αλγορίθμου της πολλαπλής ανίχνευσης και αναγνώρισης αντικειμένων.

Σε κάθε frame υπολογίζονται οι γεωμετρικές και φασματικές ιδιότητες του κάθε αντικειμένου που εντοπίζεται. Στη συνέχεια μέσω ενός ιεραρχικού γνωσιακού σχεδίου με συγκεκριμένους κανόνες, γίνεται η κατηγοριοποίηση και ο χαρακτηρισμός κάθε αντικειμένου.

Τα αποτελέσματα που προέκυψαν από την εφαρμογή του παραπάνω αλγορίθμου σε υπερφασματικό βίντεο είναι πολύ υποσχόμενα και αποδεικνύουν την αποτελεσματικότητα της συγκεκριμένης προσέγγισης. Στην εικόνα (2.1.2) φαίνονται τα αποτελέσματα της εφαρμογής του αλγορίθμου σε πειραματικά δεδομένα.



Εικ. 2.1.1 : Διάγραμμα απεικόνισης αλγορίθμου για αυτόματη ανίχνευση και παρακολούθηση πολλαπλών αντικειμένων σε ακολουθία υπερφασματικών video [7]



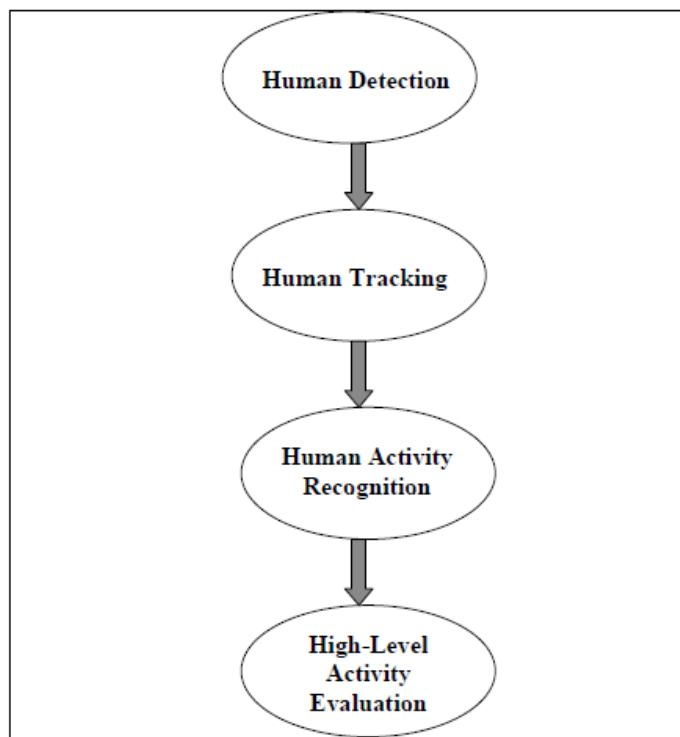
Εικ. 2.1.2: Πειραματικά αποτελέσματα μετά την εφαρμογή του αλγορίθμου σε υπερφασματικό video dataset. [7]

2.2 Ανίχνευση ανθρώπων και αναγνώριση δραστηριοτήτων μέσω όρασης υπολογιστών

Η έρευνα του R. Bodor και των συνεργατών του [9] αφορά την ανάπτυξη κάποιων κομματιών ενός αυτόματου συστήματος "smart video" για την παρακολούθηση πεζών και εντοπισμό επικίνδυνων καταστάσεων, αλλά και ύποπτων κινήσεων σε περιοχές όπου στεγάζονται μεγάλοι σταθμοί Μ.Μ.Μ. Το λογισμικό παρακολουθεί κάθε πεζό ξεχωριστά καθώς διέρχεται από το οπτικό πεδίο της κάμερας, και χρησιμοποιεί αλγορίθμους όρασης υπολογιστών για να κατηγοριοποιήσει την κίνηση και τη δραστηριότητα του κάθε ατόμου. Η παρακολούθηση των πεζών επιτυγχάνεται με τη δημιουργία διαδρομής (path) χαρακτηριστικών θέσης και ταχύτητας κάθε πεζού ξεχωριστά, μέσω της χρήσης του φίλτρου Kalman.

Η παραπάνω εργασία υλοποιήθηκε σε δύο στάδια : i) ανίχνευση πεζών και εξαγωγή εικόνας για κάθε πεζό, ii) αναγνώριση δραστηριότητας πεζού, μέσω ανάλυσης των παραμέτρων της θέσης και ταχύτητας του. Στο πρώτο μέρος της έρευνας ανήκουν, η ανίχνευση και η παρακολούθηση των ανθρώπων (human detection, human tracking) ενώ στο δεύτερο μέρος ανήκει η υψηλού

επιπέδου αξιολόγηση της ανθρώπινης δραστηριότητας. Ως ανθρώπινη δραστηριότητα ορίζεται ένα σύνολο ενεργειών.



Εικ. 2.2.1: Σχηματική επεξήγηση διαδικασίας. [9]

I. Μέρος Α : Παρακολούθηση πεζών και εξαγωγή εικόνας.

Ο στόχος του μέρους Α είναι η εξαγωγή μίας εικόνας για κάθε πεζό ξεχωριστά, που σε αυτή να απεικονίζεται μόνο η φιγούρα του ανθρώπου. Η ίδια διαδικασία επαναλαμβάνεται για κάθε πεζό ξεχωριστά, στο σύνολο της ακολουθίας των εικόνων (video). Με αυτό τον τρόπο δημιουργείται μια χρονοσειρά από εικόνες που επιτρέπει την παρακολούθηση της δραστηριότητας του κάθε ανθρώπου ξεχωριστά.

II. Μέρος Β : Αναγνώριση δραστηριότητας πεζού, μέσω ανάλυσης των παραμέτρων θέσης και ταχύτητας.

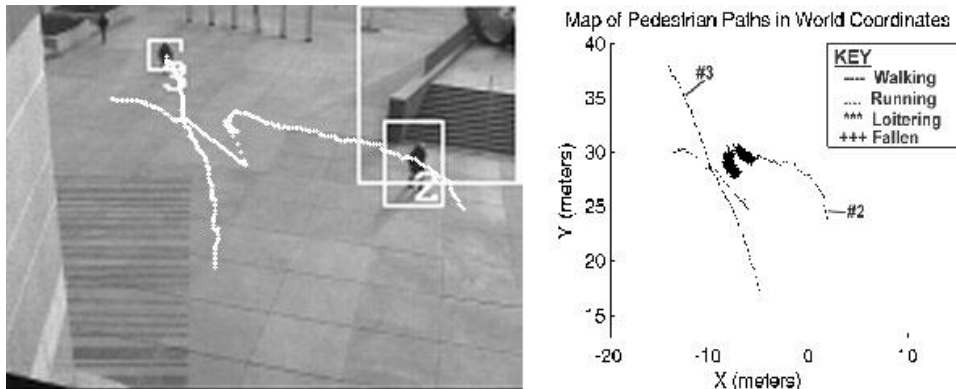
Σε αυτό το μέρος εκτιμάται η κίνηση του πεζού βάση της θέσης και της ταχύτητας του. Η βασική υπόθεση είναι ότι με τη χρήση αυτών των απλών παραγόντων μπορούμε να εξακριβώσουμε την πλειονότητα των ανθρώπινων δραστηριοτήτων. Η μέτρηση αυτών των ποσοτήτων

προσφέρει πολλαπλά πλεονεκτήματα στην αρθρωτή ανάλυση κίνησης (articulated motion analysis) : δηλαδή αυτές οι μετρήσεις μπορούν να γίνουν σε πραγματικό χρόνο και είναι πολύ πιο εύρωστες (robust) σε εικόνες με θόρυβο ή σε χαμηλής ποιότητας εικόνες.

Η διαδικασία σε στάδια έχει ως εξής:

1. Παρακολούθηση κάθε πεζού ξεχωριστά καθ όλη τη διάρκεια που βρίσκεται εντός του οπτικού πεδίου της κάμερας χρησιμοποιώντας το φίλτρο Kalman.
2. Καταγραφή της κατάστασης θέσης και ταχύτητας.
3. Δημιουργία διαδρομής (path) χαρακτηριστικών θέσης και ταχύτητας για κάθε πεζό.
4. Ορισμός ενός σήματος κινδύνου ανάλογα με την κάθε περίπτωση (π.χ. είσοδος σε περιοχή ασφαλείας , πεζός κινείται με ταχύτητα, πεζός έπεσε κάτω κλπ)

Στην εικόνα (2.2.2) παρουσιάζεται μια ακολουθία όπου φαίνεται η κίνηση δύο πεζών σε μια πλατεία. Ο πεζός #2 σε κάποια στιγμή παρουσιάζεται ότι έπεσε κάτω.



Εικ. 2.2.2 : Εικόνα πεζών που εντοπίστηκαν και χάρτης κίνησης πεζών. [9]

2.3 Μοντέλα μίξης προσαρμοστικών παρασκηνίων (adaptive backgrounds) για παρακολούθηση σε πραγματικό χρόνο

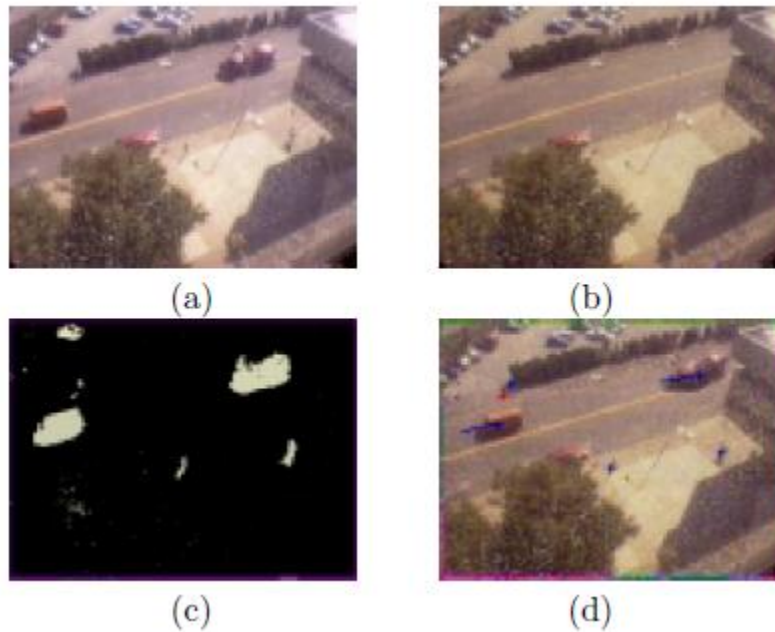
Οι πιο συνηθισμένες μέθοδοι για την επίτευξη κατάτμησης μιας εικόνας σε κινητές και ακίνητες περιοχές είναι μέσω διαδικασιών αφαίρεσης παρασκηνίου (background subtraction) ή μέσω οριοθέτησης (thresholding) του σφάλματος μεταξύ μιας εκτίμησης της στατικής εικόνας σε σχέση με τη παρούσα εικόνα. Οι διάφορες προσεγγίσεις στο πιο πάνω πρόβλημα διαφέρουν στον τύπο του μοντέλου παρασκηνίου (background model) που χρησιμοποιείται, και στη διαδικασία εκσυγχρονισμού (update) αυτού του μοντέλου.

Στην έρευνα των C. Staufferr και W.E.L Grimson [10] προτείνεται η μοντελοποίηση κάθε pixel σαν μίξη από Γκαουσιανές (Κατανομές του Gauss) και τη χρήση σε απευθείας σύνδεση εκτιμήσεων που θα εκσυγχρονίζουν το μοντέλο. Στη συνέχεια οι Γκαουσιανές του μοντέλου θα αξιολογούνται και θα καθορίζεται ποιές είναι πιθανότερο να προέκυψαν από τη διαδικασία παρασκηνίου. Κάθε pixel κατηγοριοποιείται ανάλογα με το αν η Γκαουσιανή που το αντιπροσωπεύει πιο αποτελεσματικά, θεωρείται ως μέρος του μοντέλου παρασκηνίου (background model). Pixels με τιμές που δεν προσαρμόζονται στην κατανομή των παρασκηνίων τους, θεωρούνται ότι ανήκουν στο προσκήνιο μέχρι μια Γκαουσιανή να τις περιέχει επαρκώς. Τότε αυτά τα pixels θεωρούνται και πάλι ότι ανήκουν στο παρασκήνιο.

Μεθοδολογία:

Η ανάγκη για χρήση πολλαπλών προσαρμοστικών Γκαουσιανών είναι απαραίτητη γιατί στην πράξη, το σκηνικό που καταγράφει μια κάμερα αποτελείται από ποικιλόμορφες επιφάνειες που δέχονται διαφορετική επίδραση φωτισμού. Αν κάθε pixel κατέγραφε μια μόνο συγκεκριμένη επιφάνεια, κάτω από συγκεκριμένο φωτισμό, τότε μια μόνο Γκαουσιανή θα ήταν αρκετή για την μοντελοποίηση του pixel. Αν η μόνη μεταβλητή ήταν η αλλαγή του φωτισμού στο χρόνο, τότε το πρόβλημα θα λυνόταν με τη χρήση μιας προσαρμοστικής Γκαουσιανής ανά pixel. Έτσι η προσέγγιση των ερευνητών είναι μια διαδικασία που χρησιμοποιεί μια μίξη από προσαρμοστικές Γκαουσιανές.

Κάθε φορά που οι παράμετροι των Γκαουσιανών εκσυγχρονίζονται, οι Γκαουσιανές αξιολογούνται και γίνεται η υπόθεση για το ποιές από αυτές είναι πιο πιθανό να είναι μέρος της διαδικασίας παρασκηνίου. Τα pixels με τιμές που δεν ταιριάζουν με αυτή των pixel των Γκαουσιανών παρασκηνίων, ομαδοποιούνται με τη χρήση συνδεδεμένων συνιστωσών (connected components). Τέλος, οι συνδεδεμένες συνιστώσες παρακολουθούνται από frame σε frame με τη χρήση ενός tracker πολλαπλών υποθέσεων (multiple hypothesis tracker). Η διαδικασία παρουσιάζεται στην εικόνα (3.3.1).



Εικ 2.3.1 : Εκτέλεση του προγράμματος. [10]

α) παρούσα εικόνα, b) εικόνα που προέκυψε από τις πιο πιθανές Gaussians στο background model, c) εικόνα με τα foreground pixels, d) παρούσα εικόνα με το tracking

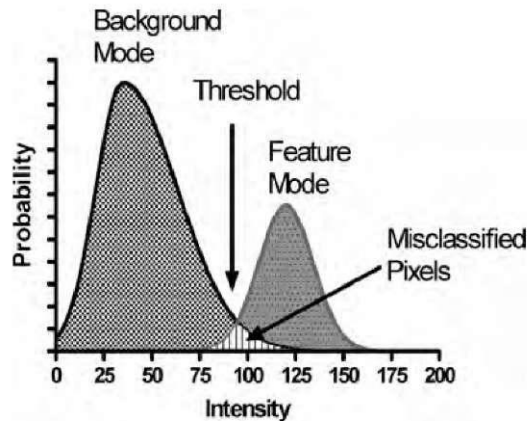
3

Θεωρητικό υπόβαθρο

3.1 Μέθοδος Otsu

Ο Nobuyuki Otsu ο επιστήμονας από τον οποίο πήρε το όνομα του η μέθοδος, προτείνει μια μη παραμετρική και μη επιβλεπόμενη μέθοδο αυτόματης επιλογής κατωφλίου (threshold) για κατάτμηση εικόνας (picture segmentation). Για να μεγιστοποιηθεί η ικανότητα διαχωρισμού των τάξεων των γκρι επιπέδων, επιλέγεται ένα βέλτιστο κατώφλι που προκύπτει από το διακρίνων κριτήριο (discriminant criterion).

Κατά την επεξεργασία της εικόνας είναι πολύ σημαντική η επιλογή του κατάλληλου ορίου (threshold) γκρι επιπέδου για την εξαγωγή αντικειμένων από το παρασκήνιο τους. Σε μια ιδανική περίπτωση το ιστόγραμμα της εικόνας παρουσιάζει μια βαθιά και απότομη κοιλάδα ανάμεσα σε δύο κορυφές (οι δύο κορυφές αυτές αντιπροσωπεύουν τα αντικείμενα και το background). Στην περίπτωση αυτή ως όριο (threshold) επιλέγεται μια τιμή κοντά στο κάτω μέρος αυτής της κοιλάδας. [11]



Εικ. 3.1.1 : Ιστόγραμμα επιλογής threshold Otsu Method

Στις περισσότερες πραγματικές εικόνες, η ανίχνευση (με ικανοποιητική ακρίβεια) της κάτω κοιλάδας είναι πολύ δύσκολη. Περιπτώσεις που καθιστούν δύσκολη την ανίχνευση της κάτω κοιλάδας είναι, όταν η κοιλάδα είναι ευρεία και επίπεδη, όταν στην κοιλάδα παρουσιάζεται θόρυβος, ή όταν οι δυο κορυφές είναι εξαιρετικά άνισες σε ύψος. Σε αυτές τις περιπτώσεις μπορεί να μην ανιχνευτεί καμία κοιλάδα από τον αλγόριθμο.

Η επιλογή του ορίου (threshold) πρέπει να προκύπτει αποκλειστικά και μόνο από το γκρι επίπεδο του ιστογράμματος, χωρίς κάποια άλλη a-priori γνώση. Η μεθοδολογία εκτός από το ότι αποτελεί μια πρότυπη τεχνική επεξεργασίας εικόνας, είναι απαραίτητη σε προβλήματα που αναφέρονται σε μη επιβλεπόμενη αναγνώριση μοτίβων (pattern recognition).

3.1.1 Διατύπωση μεθόδου

Έστω ότι τα εικονοστοιχεία (pixels) μιας δεδομένης εικόνας αντιπροσωπεύουν L επίπεδα του γκρι. $[1,2,..,L]$. Ο αριθμός των εικονοστοιχείων στο επίπεδο i συμβολίζεται με n_i και το σύνολο των εικονοστοιχείων με $N = n_1 + n_2 + .. + n_L$. Για σκοπούς απλοποίησης το ιστόγραμμα των γκρι επιπέδων είναι κανονικοποιημένο και θεωρείται ως μια κατανομή πιθανοτήτων.

$$p_i = n_i/N, p_i \geq 0, \sum_{i=1}^L p_i = 1 \quad (1)$$

Τώρα ας υποθέσουμε ότι έχουμε διχοτομήσει τα pixels σε δύο κατηγορίες C_0 και C_1 (υπόβαθρο και αντικείμενα, ή το αντίστροφο) βάση ενός κατώτατου ορίου (threshold) σε ένα επίπεδο k . Το C_0 υποδηλώνει pixels με επίπεδα $[1,2,\dots,k]$, και το C_1 υποδηλώνει pixels με επίπεδα $[k+1,\dots,L]$. Στη συνέχεια η πιθανότητα εμφάνισης του επιπέδου, και η τάξη του μέσου επιπέδου αντίστοιχα, δίνονται από:

$$\omega_0 = Pr(C_0) = \sum_{i=1}^k p_i = \omega(k) \quad (2)$$

$$\omega_1 = Pr(C_1) = \sum_{i=k+1}^L p_i = 1 - \omega(k) \quad (3)$$

και

$$\mu_0 = \sum_{i=1}^k i Pr(i|C_0) = \sum_{i=1}^k ip_i / \omega_0 = \mu(k) / \omega(k) \quad (4)$$

$$\mu_1 = \sum_{i=k+1}^L i Pr(i|C_1) = \sum_{i=k+1}^L ip_i / \omega_1 = \frac{\mu_T - \mu(k)}{1 - \omega(k)} \quad (5)$$

όπου

$$\omega(k) = \sum_{i=1}^k p_i \quad (6)$$

και

$$\mu(k) = \sum_{i=1}^k ip_i \quad (7)$$

είναι τα μηδενικά και τα πρώτης τάξης αθροιστικά στιγμιότυπα του ιστογράμματος μέχρι και το k - επίπεδο αντίστοιχα, και

$$\mu_T = \mu(L) = \sum_{i=1}^L ip_i \quad (8)$$

είναι το σύνολο των μέσων επιπέδων της εικόνας. Μπορεί να αποδειχτεί ότι οι ακόλουθες σχέσεις ισχύουν για οποιαδήποτε τιμή της τιμής k .

$$\omega_0 \mu_0 + \omega_1 \mu_1 = \mu_T, \quad \omega_0 + \omega_1 = 1 \quad (9)$$

Οι διακυμάνσεις της κλάσης προκύπτουν από :

$$\sigma_0^2 = \sum_{i=1}^k (i - \mu_0)^2 Pr(i|C_0) \sum_{i=1}^k (i - \mu_0)^2 p_i / \omega_0 \quad (10)$$

$$\sigma_1^2 = \sum_{i=k+1}^L (i - \mu_1)^2 Pr(i|C_1) \sum_{i=k+1}^L (i - \mu_1)^2 p_i / \omega_1 \quad (11)$$

Αυτό απαιτεί δεύτερης τάξης στατιστικά σωρευτικά στιγμιότυπα (cumulative moments).

Προκειμένου να αξιολογηθεί η "αποτελεσματικότητα" του ορίου (threshold) στο επίπεδο k , παρουσιάζονται οι ακόλουθες μεταβλητές που αποτελούν διακρίνων κριτήρια [12].

$$\lambda = \sigma_B^2 / \sigma_W^2, \quad \kappa = \sigma_T^2 / \sigma_W^2, \quad \eta = \sigma_B^2 / \sigma_T^2 \quad (12)$$

όπου

$$\sigma_W^2 = \omega_0 \sigma_0^2 + \omega_1 \sigma_1^2 \quad (13)$$

$$\sigma_B^2 = \omega_0 (\mu_0 - \mu_T)^2 + \omega_1 (\mu_1 - \mu_T)^2 = \omega_0 \omega_1 (\mu_1 - \mu_0)^2 \quad (14)$$

(λόγο της (9)) και

$$\sigma_T^2 = \sum_{i=1}^L (i - \mu_T)^2 p_i \quad (15)$$

είναι η διακύμανση εντός-τάξης (within-class variance), η διακύμανση μεταξύ των τάξεων (between-class variance), και η συνολική διακύμανση των επιπέδων (total variance of levels) αντίστοιχα. Στη συνέχεια το πρόβλημα μας περιορίζεται σε ένα πρόβλημα βελτιστοποίησης για την αναζήτηση ενός κατώτατου ορίου k (threshold k) το οποίο μεγιστοποιεί μια μεταβλητή διακρίνων κριτηρίου στη σχέση (12).

Από τα παραπάνω προκύπτει η εικασία ότι καλά ορισμένες κλάσεις (wellthresholded) θα χωρίζονται σε επίπεδα του γκρι, και αντιστρόφως, ένα όριο που δίνει το καλύτερο διαχωρισμό γκρι επιπέδων, θα είναι το καλύτερο όριο.

Η μεγιστοποίηση των κριτηρίων λ , κ , και η , αντίστοιχα για κάποιο k , οδηγεί στο συμπέρασμα ότι υπάρχει ισοδυναμία μεταξύ των κριτηρίων; π.χ. : $\kappa = \lambda + 1$ και $\eta = \lambda / (\lambda + 1)$ με όρους λ , εφόσον ισχύει πάντα η σχέση :

$$\sigma_W^2 + \sigma_B^2 = \sigma_T^2 \quad (16)$$

Παρατηρείται ότι οι ποσότητες σ_w^2 και σ_B^2 είναι συναρτήσεις του ορίου του επιπέδου k , αλλά το σ_T^2 είναι ανεξάρτητο του k . Σημειώνεται επίσης ότι το σ_w^2 βασίζεται σε δεύτερης τάξης στατιστική (class variance), ενώ το σ_B^2 βασίζεται σε πρώτης τάξης στατιστική (class means). Έτσι παίρνουμε ως διακρίνων κριτήριο το " η " για την αξιολόγηση της αποτελεσματικότητας του ορίου στο επίπεδο K .

Το ιδανικό όριο k^* (optimal threshold k^*) το οποίο μεγιστοποιεί το " η ", ή αντίστοιχα μεγιστοποιεί το σ_B^2 , επιλέγεται με την ακόλουθη διαδοχική αναζήτηση των απλών σωρευτικών ποσοτήτων από τις (6) και (7), ή από την (6) χρησιμοποιώντας πρώτα τις (2)-(5):

$$\eta(k) = \sigma_B^2(k) / \sigma_T^2 \quad (17)$$

$$\sigma_B^2(k) = \frac{[\mu_T \omega(k) - \mu(k)]^2}{\omega(k)[1 - \omega(k)]} \quad (18)$$

και το ιδανικό όριο k^* είναι:

$$\sigma_B^2(k^*) = \max_{1 \leq k < L} \sigma_B^2(k) \quad (19)$$

Από το πρόβλημα προκύπτει ότι το φάσμα των τιμών του k του οποίου αναζητούμε την μεγιστοποίηση, μπορεί να περιοριστεί σε:

$$S^* = \{k; \omega_0 \omega_1 = \omega(k)[1 - \omega(k)] > 0, \quad or \quad 0 < \omega(k) < 1\}$$

Θα το ονομάσουμε ως το effective range του gray-level ιστογράμματος.

Από τη (14) , προκύπτει ότι το κριτήριο σ_B^2 (ή " η ") παίρνει ελάχιστες τιμές μηδενικών για τιμές k όταν $k \in S - S^* = \{k; \omega(k) = 0 \text{ ή } 1\}$ (δηλαδή κάνοντας όλα τα εικονοστοιχεία είτε C_0 είτε C_1) και παίρνει θετική και φραγμένη τιμή για $k \in S^*$. Ως εκ τούτου είναι προφανές ότι υπάρχει πάντοτε το μέγιστο.

3.2 Φίλτρο Kalman

3.2.1 Γενικά για το φίλτρο Kalman

Το φίλτρο Kalman [14-16] πήρε το όνομα του από τον Rudolf E. Κάλμάν που ήταν ο κύριος εμπνευστής αυτής της θεωρίας. Αποτελεί ένα ιδιαίτερα διαδεδομένο και εύχρηστο φίλτρο που χρησιμοποιείται για απαλλαγή από τις θορυβώδεις μετρήσεις. Ο αλγόριθμος αυτός, πραγματοποιεί πρόβλεψη της θέσης του στόχου και χρησιμοποιεί τις μετρήσεις για να διορθώσει την πρόβλεψη αυτή. Επομένως έχει την δυνατότητα να παρέχει εκτιμήσεις με σφάλμα μικρότερο του σφάλματος μέτρησης. Είναι μια επέκταση της μεθόδου ελαχίστων τετραγώνων και πρόκειται ουσιαστικά για ένα σύνολο μαθηματικών εξισώσεων, οι οποίες εισαγάγουν έναν εκτιμητή πρόβλεψης – διόρθωσης. Αυτός ο εκτιμητής είναι βέλτιστος από την άποψη, ότι ελαχιστοποιεί την συνδιακύμανση σφάλματος εκτίμησης, όταν η δυναμική κατάσταση του στόχου και ο θόρυβος μέτρησης έχουν μοντελοποιηθεί με ακρίβεια. Σε γραμμικά περιβάλλοντα ο αλγόριθμος του Kalman αποτελεί τη βέλτιστη λύση.

Η εκτεταμένη χρήση του φίλτρου Kalman οφείλεται στα πολλά πλεονεκτήματα που προσφέρει στο πρόβλημα της καλύτερης δυνατής εκτίμησης της δυναμικής κατάστασης ενός συστήματος. Σε συστήματα με χρονομεταβλητές δυναμικές καταστάσεις ή μετρήσεις, οι διαδικασίες των οποίων μεταβάλλονται με τον χρόνο, το γεγονός ότι γίνεται ανανέωση του κέρδους (gain factor) με κάθε νέα μέτρηση διευκολύνει την ανίχνευση. Επιπλέον, η συνδιακύμανση του θορύβου μέτρησης μπορεί να αυξομειωθεί, έτσι ώστε η κάθε μέτρηση να αποκτήσει διαφορετικό βάρος ανάλογα με την απόσταση από τον στόχο, την ισχύ του σήματος ή οποιαδήποτε άλλη ενδεχόμενη ένδειξη της ποιότητας της μέτρησης. Αυτό παρέχει τη δυνατότητα χρήσης διαφορετικού τύπου αισθητήρων όπου ο χρήστης μπορεί να κάνει μία αρχικοποίηση ανάλογα με την κατά περίπτωση αναμενόμενη αξιοπιστία του εκάστοτε αισθητήρα. Το σημαντικότερο πλεονέκτημα του φίλτρου είναι η αποδοτικότητά του σε υπολογιστικά συστήματα περιορισμένης μνήμης, γεγονός που το κάνει ιδιαίτερα δημοφιλή σε προγράμματα που δουλεύουν σε πραγματικό χρόνο. Όλη η πληροφορία για τις αρχικές καταστάσεις, τις μετρήσεις και την κίνηση περιέχεται στον πίνακα συνδιακύμανσης σφάλματος P_k .

3.2.2 Μαθηματική επεξήγηση

Για την εκτίμηση ενός δυναμικού συστήματος γενικά χρησιμοποιούνται 2 μοντέλα :

- 1) Το μοντέλο που περιγράφει την εξέλιξη της κατάστασης στο χρόνο (κινηματικό μοντέλο).
- 2) Το μοντέλο που σχετίζεται με τις μετρήσεις θορύβου (μοντέλο μέτρησης).

Έστω $x \in \mathcal{R}^n$ το διάνυσμα κατάστασης ενός στόχου, του οποίου η δυναμική κατάσταση περιγράφεται από την γραμμική εξίσωση :

$$x_k = Ax_{k-1} + Du_k + w_{k-1} \quad (1)$$

όπου

- A : $(n \times n)$ πίνακας που αποτελεί τον πίνακα μετάβασης της κατάστασης από την $k-1$ μέτρηση στην k , όπου το χρονικό διάστημα μέτρησης είναι T .
- D : $(n \times l)$ πίνακας ο οποίος συσχετίζει την κατάσταση με το διάνυσμα εισόδου $u \in \mathcal{R}^l$.
- w : είναι η τυχαία μεταβλητή που μοντελοποιεί τον λευκό, γκαουσιανό θόρυβο διαδικασίας μηδενικής μέσης τιμής και με συνδιακύμανση που δίνεται από τον πίνακα $(n \times n)$ Q . Ο θόρυβος διαδικασίας αναπαριστά τις τυχαίες επιδράσεις, τις οποίες υφίσταται το σύστημα και την ατελή μοντελοποίηση της δυναμικής κατάστασης του στόχου.

Από την (1) είναι σαφές ότι η στατιστική αναπαράσταση της διαδικασίας την παρούσα στιγμή (k -οστή μέτρηση) καθορίζεται αποκλειστικά από την προηγούμενη (μέτρηση $k-1$). Η δυναμική κατάσταση του στόχου οδηγείται από την ντετερμινιστική είσοδο και τον τυχαίο θόρυβο διαδικασίας. Η ντετερμινιστική είσοδος του συστήματος είναι δυνατόν να περιέχει κάποιες γνωστές μεταβλητές εισόδου, οι οποίες δεν περιλαμβάνονται στις καταστάσεις του συστήματος.

Έστω ακόμη ότι η μέτρηση $z \in \mathcal{R}^m$ περιγράφεται από το μοντέλο:

$$z_k = Hx_k + v_k \quad (2)$$

όπου,

- H : $(m \times n)$ πίνακας που συνδέει την κατάσταση με το διάνυσμα μετρήσεων.
- v : Αντίστοιχα με το w η τυχαία μεταβλητή v παριστάνει τον θόρυβο μέτρησης μηδενικής μέσης τιμής και συνδιακύμανσης R $(m \times m)$.

Οι πίνακες A , H , Q και R είναι γενικά χρονικά μεταβαλλόμενοι, αλλά εφεξής θεωρούνται σταθεροί για λόγους απλότητας

Έστω τώρα $x_k^- \in \mathcal{R}^n$ το *a priori* εκτιμώμενο διάνυσμα κατάστασης της σάρωσης k , γνωρίζοντας την διαδικασία μέχρι την σάρωση k και $x_k \in \mathcal{R}^n$ το *a posteriori* διάνυσμα εκτίμησης στο βήμα k , γνωρίζοντας την μέτρηση z_k . Στην συνέχεια ορίζονται αντίστοιχα το *a priori* και *a posteriori* σφάλμα εκτίμησης :

$$e_k^- \equiv x_k^- - x_k \quad (3)$$

$$e_k \equiv x_k - x_k \quad (4)$$

Επίσης αντίστοιχα ορίζονται και οι *a priori* και *a posteriori* πίνακες συνδιακύμανσης σφάλματος εκτίμησης :

$$P_k^- = E[e_k^- e_k^{-T}] \quad (5)$$

$$P_k = E[e_k e_k^T] \quad (6)$$

Προκειμένου να παραχθούν οι εξισώσεις του φίλτρου χρειάζεται μια σχέση, η οποία να υπολογίζει την *a posteriori* εκτίμηση της κατάστασης ως γραμμικό συνδυασμό της *a priori* εκτίμησης και μιας διαφοράς της πραγματικής μέτρησης και της πρόβλεψης μέτρησης $H x_k^-$:

$$x_k = x_k^- + K(z_k - H x_k^-) \quad (7)$$

Η διαφορά $(z_k - H x_k^-)$ συναντάται στη βιβλιογραφία ως innovation ή residual vector και αναπαριστά την διαφορά ανάμεσα στην πρόβλεψη μέτρησης και την πραγματική μέτρηση και μηδενική τιμή σημαίνει, ότι η πρόβλεψη συμπίπτει με την πραγματική μέτρηση.

K : $(n \times m)$ πίνακας αντιπροσωπεύει το κέρδος (gain factor) του φίλτρου.

Το κέρδος K ελαχιστοποιεί τον *a posteriori* πίνακα συνδιακύμανσης σφάλματος εκτίμησης $P_k = E[e_k e_k^T]$. Η ελαχιστοποίηση επιτυγχάνεται ως εξής :
Με αντικατάσταση της (7) στην (4) :

$$e_k = x_k - x_k^- (I + KH) - Kz_k \quad (8)$$

Στην συνέχεια η (8) αντικαθίσταται στην (6) και το ίχνος του αποτελέσματος ως προς K τίθεται ίσο προς μηδέν.

Το K_k , το οποίο ελαχιστοποιεί την (6) μπορεί να γραφεί και στην μορφή :

$$K_k = P_k^- H^T (H P_k^- H^T + R)^{-1} = \frac{P_k^- H^T}{H P_k^- H^T + R} \quad (9)$$

Παρατηρούμε ότι καθώς ο πίνακας συνδιακύμανσης σφάλματος μέτρησης R τείνει στο μηδέν, το κέρδος K καθορίζει το υπόλοιπο περισσότερο :

$$\lim_{R_k \rightarrow 0} K_k = H^{-1}$$

Παράλληλα, καθώς ο *a priori* πίνακας συνδιακύμανσης σφάλματος εκτίμησης P_k τείνει στο μηδέν, το κέρδος καθορίζει το υπόλοιπο λιγότερο :

$$\lim_{P_k^- \rightarrow 0} K_k = 0$$

Με άλλα λόγια, καθώς το σφάλμα μέτρησης τείνει στο μηδέν, η πραγματική μέτρηση γίνεται πιο αξιόπιστη, ενώ η πρόβλεψη μέτρησης γίνεται λιγότερο αξιόπιστη. Αντίστοιχα, καθώς ο *a priori* πίνακας συνδιακύμανσης σφάλματος εκτίμησης τείνει στο μηδέν, η πραγματική μέτρηση γίνεται λιγότερο αξιόπιστη και η πρόβλεψη μέτρησης περισσότερο.

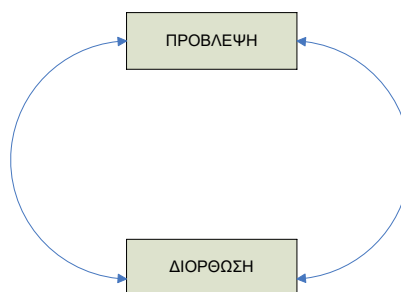
3.2.3 Αλγόριθμος Kalman

Το Φίλτρο Kalman αποτελεί έναν αναδρομικό αλγόριθμο εκτίμησης της δυναμικής κατάστασης ενός συστήματος (recursive state-estimation) [17]. Ο αλγόριθμος του φίλτρου πραγματοποιεί την εκτίμηση, χρησιμοποιώντας μια μορφή ελέγχου ανάδρασης (feedback control). Αρχικά το φίλτρο εκτιμά την κατάσταση του συστήματος για κάποια στιγμή και στην συνέχεια αποκτά ανάδραση από τις μετρήσεις. Επομένως οι εξισώσεις του φίλτρου Kalman

διακρίνονται σε δύο ομάδες: i) στις εξισώσεις ενημέρωσης χρόνου (time update) και, ii) στις εξισώσεις ενημέρωσης μέτρησης (measurement update).

Οι εξισώσεις της πρώτης ομάδας είναι υπεύθυνες για την χρονική προώθηση της εκτίμησης της παρούσας κατάστασης και την *a priori* εκτίμηση της συνδιακύμανσης σφάλματος για το επόμενο βήμα. Οι εξισώσεις της δεύτερης είναι υπεύθυνες για την ανάδραση, όπως για παράδειγμα την εισαγωγή νέων μετρήσεων στην *a priori* εκτίμηση, ώστε να παραχθεί βελτιωμένη *a posteriori* εκτίμηση.

Συνοπτικά μπορεί να ειπωθεί ότι οι εξισώσεις ενημέρωσης χρόνου είναι οι εξισώσεις πρόβλεψης (predictor equations), ενώ οι εξισώσεις ενημέρωσης μέτρησης είναι οι εξισώσεις διόρθωσης (corrector equations). Επομένως το φίλτρο Kalman μπορεί να θεωρηθεί ως ένας αλγόριθμος πρόβλεψης – διόρθωσης.



Εικ. 3.2.1: Αλγόριθμος KALMAN

Οι εξισώσεις ενημέρωσης χρόνου του φίλτρου είναι οι εξής :

$$x_k^- = Ax_{k-1} + Du_k \quad (10)$$

$$P_k^- = AP_{k-1}A^T + Q \quad (11)$$

και οι εξισώσεις ενημέρωσης μέτρησης είναι :

$$K_k = P_k^- H^T (H P_k^- H^T + R)^{-1} \quad (12)$$

$$x_k = x_k^- + K_k (z_k - H x_k^-) \quad (13)$$

$$P_k = (I - K_k H) P_k^- \quad (14)$$

Αρχικά υπολογίζεται η *a priori* εκτίμηση της κατάστασης, με βάση την *a posteriori* εκτίμηση της προηγούμενης σάρωσης, όπως φαίνεται στην σχέση (10).

Αντίστοιχα υπολογίζεται και η *a priori* εκτίμηση του πίνακα συνδιακύμανσης σφάλματος από την *a posteriori* εκτίμηση του προηγούμενου βήματος (σχέση 11). Αυτό είναι το στάδιο της πρόβλεψης.

Στη συνέχεια ακολουθεί η φάση της διόρθωσης των μεγεθών αυτών, με χρήση των νέων μετρήσεων. Υπολογίζεται η νέα τιμή του κέρδους K και πραγματοποιείται η *a posteriori* εκτίμηση της κατάστασης με βάση την προβλεπόμενη τιμή, το κέρδος και το υπόλοιπο μέτρησης (σχέση 13).

Τέλος λαμβάνει χώρα η εκτίμηση της *a posteriori* τιμής του πίνακα συνδιακύμανσης σφάλματος (σχέση 14).

Οι *a posteriori* τιμές χρησιμοποιούνται στο επόμενο βήμα εκ νέου για ενημέρωση των προβλέψεων.

3.3 Αλγόριθμος συσταδοποίησης k-μέσων

3.3.1 Γενικά για τον αλγόριθμο k-μέσων

Ο αλγόριθμος k-μέσων [21] είναι ένας από τους πιο δημοφιλείς αλγορίθμους μη επιτηρούμενης μάθησης (unsupervised learning algorithm), ο οποίος δίνει λύση στο πρόβλημα της ομαδοποίησης των δεδομένων. Είναι ιδιαίτερα χρήσιμος στον τομέα της συσταδικής ανάλυσης (cluster analysis) σε εφαρμογές εξόρυξης δεδομένων (data mining). Η διαδικασία που ακολουθεί ο

αλγόριθμος για την κατηγοριοποίηση ενός συνόλου δεδομένων σε ένα συγκεκριμένο αριθμό συστάδων (ο οποίος έχει καθοριστεί a priori), είναι εξαιρετικά απλή και εύκολη.

Υπάρχουν διάφορες παραλλάξεις του αρχικού αλγορίθμου. Στο πλαίσιο της εργασίας έγινε η χρήση του πρότυπου αλγορίθμου (standard k-means algorithm) ο οποίος συναντάται στη βιβλιογραφία και ως Lloyd's algorithm [22].

Πριν την έναρξη των διαδικασιών του αλγορίθμου, πρέπει να επιλεχτεί η μέθοδος αρχικοποίησης του αλγορίθμου, από την οποία προκύπτουν τα αρχικά κέντρα (centroids) των συστάδων k που επιλέγηκαν. Οι συνηθέστερες μέθοδοι αρχικοποίησης είναι η μέθοδος Forgy και η μέθοδος τυχαίου διαχωρισμού (random partition method) [23]. Η μέθοδος Forgy επιλέγει ως αρχικό centroid κάθε συστάδας το centroid μιας τυχαίας παρατήρησης από το σύνολο των δεδομένων. Η μέθοδος τυχαίου διαχωρισμού αρχικά αναθέτει με τυχαίο τρόπο σε συστάδα (cluster) σε κάθε παρατήρηση για το σύνολο των παρατηρήσεων του dataset. Στη συνέχεια μέσω ενός σταδίου εκσυγχρονισμού (update step), υπολογίζεται το κέντρο κάθε συστάδας ως το κέντρο βάρους των παρατηρήσεων που ανήκουν στην εκάστοτε συστάδα.

Η μέθοδος Forgy τείνει να διασκορπίζει τα αρχικά κέντρα των συστάδων, σε αντίθεση με την μέθοδο τυχαίου διαχωρισμού όπου τα κέντρα των συστάδων τοποθετούνται κοντά στο κέντρο του συνόλου των παρατηρήσεων. Στην περίπτωση εφαρμογής του πρότυπου αλγορίθμου προτείνεται η χρήση της μεθόδου Forgy, χωρίς να συνεπάγεται ότι η χρήση της μεθόδου τυχαίας επιλογής θα παρουσιάσει μη ικανοποιητικά αποτελέσματα.

3.3.2 Βήματα αλγόριθμου

Μετά τον ορισμό της μεθόδου αρχικοποίησης που περιγράφηκε πιο πάνω, ο αλγόριθμος συνεχίζει με την εναλλαγή ανάμεσα σε δυο βήματα. Δηλαδή τα δυο αυτά βήματα αποτελούν μια επαναληπτική διαδικασία.

- Βήμα Ανάθεσης:

Εδώ γίνεται η ανάθεση κάθε παρατήρησης στην κλάση της οποίας το κέντρο αποδίδει το πιο μικρό άθροισμα τετραγώνων εντός συστάδας

(within-cluster sum of squares). Το άθροισμα τετραγώνων εντός συστάδας στην ουσία είναι το τετράγωνο της Ευκλείδειας απόστασης, άρα η ανάθεση κάθε παρατήρησης γίνεται στο "κοντινότερο" κέντρο μιας συστάδας (nearest mean). Από μαθηματικής πλευράς, αυτό σημαίνει ότι ο διαχωρισμός των παρατηρήσεων γίνεται σύμφωνα με το διάγραμμα Voronoi που παράγεται από τα centroids των συστάδων.

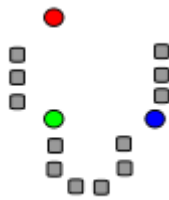
$$S_i^{(t)} = \{x_p : \|x_p - m_i^{(t)}\|^2 \leq \|x_p - m_j^{(t)}\|^2 \forall j, 1 \leq j \leq k\}$$

όπου κάθε x_p έχει ανατεθεί σε μία μόνο $S_i^{(t)}$, ακόμα και αν θα μπορούσε να ανατεθεί σε δύο ή και περισσότερες.

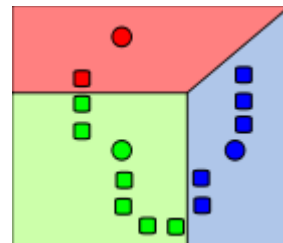
- Βήμα εκσυγχρονισμού:

Υπολογίζει τα καινούρια κέντρα των συστάδων ούτως ώστε να είναι τα centroids των παρατηρήσεων σε κάθε καινούρια συστάδα.

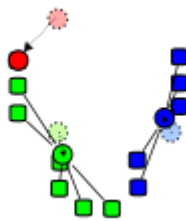
$$m_i^{(t+1)} = \frac{1}{|S_i^{(t)}|} \sum_{x_j \in S_i^{(t)}} x_j$$



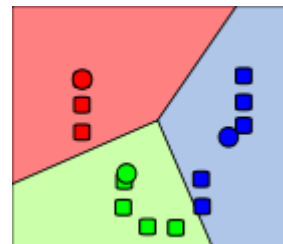
1. Τα αρχικά centroids ($k=3$, εμφανίζονται με χρώμα) δημιουργούνται τυχαία μέσα στο χώρο των δεδομένων.



2. Δημιουργούνται k συστάδες από την ανάθεση μιας κλάσης σε κάθε παρατήρηση με το κριτήριο nearest mean. Ο διαχωρισμός που παρουσιάζεται στην εικόνα αντιπροσωπεύει το διάγραμμα Voronoi που παράγεται από τα κέντρα των συστάδων.



3. Το centroid κάθε k συστάδας γίνεται το καινούριο κέντρο.



4. Επαναλαμβάνονται τα στάδια 2 και 3 μέχρι την επίτευξη σύγκλισης στα αποτελέσματα της επανάληψης.

Εικ 3.3.2.1 : Παρουσίαση διαδικασίας πρότυπου αλγορίθμου k -μέσων

4

Πειραματική διάταξη

4.1 Γενικά

Στο πλαίσιο του έργου eVacuate διεξήχθησαν κάποια πειράματα σχετικά με την δυναμική κίνηση των ανθρώπων σε περίπτωση ανάγκης έκτακτης εκκένωσης κτηρίων. Τα πειράματα αυτά διεξήχθησαν σε τρεις διαφορετικές χώρες, σε τέσσερα κτίρια υψηλής συγκέντρωσης πλήθους. Η ομάδα των ελλήνων επιστημόνων επέλεξε την διεξαγωγή των πειραμάτων της στο Αεροδρόμιο "Ελευθέριος Βενιζέλος" στην Αθήνα. Συγκεκριμένα τα πειράματα έλαβαν χώρα σε μία τερματική πύλη (terminal gate), με εθελοντές πολίτες να παίζουν το ρόλο των επιβατών. Διεξήχθησαν διάφορα σενάρια έκτακτης ανάγκης εκκένωσης του χώρου (π.χ. φωτιά, βόμβα κλπ), στα οποία καταγράφηκε η κίνηση των ανθρώπων από 3 στατικές κάμερες - αισθητήρες (1x υπερφασματικός και 2x θερμικοί).

Η παρούσα εργασία ασχολείται με την επεξεργασία και ανάλυση των βίντεο που καταγράφηκαν από τις κάμερες που αναφέρθηκαν πιο πάνω.

Συγκεκριμένα σαν δεδομένα χρησιμοποιήθηκαν τρία αρχεία βίντεο (.avi), δυο εκ το οποίων λήφθηκαν με τη θερμική κάμερα και ένα με την υπερφασματική. Λεπτομέρειες για τα χαρακτηριστικά αυτών των αρχείων δίνονται παρακάτω.

4.2 Δεδομένα από θερμικό αισθητήρα

Οι θερμικές κάμερες (thermal imaging cameras) είναι συσκευές οι οποίες καταγράφουν υπέρυθρη ακτινοβολία, με παρόμοιο τρόπο λειτουργίας με τις παραδοσιακές κάμερες που δημιουργούν μια εικόνα χρησιμοποιώντας το φάσμα του ορατού φωτός. Σε αντίθεση με τις παραδοσιακές κάμερες που έχουν εύρος φασματικής κάλυψης τα 400 - 700 nm (φάσμα ορατού φωτός), οι θερμικές κάμερες καταγράφουν μήκη κύματος που φτάνουν μέχρι και τα 14000 nm (14μm).

Οι εικόνες που προέρχονται από θερμικές κάμερες τείνουν να είναι μονόχρωμες, για το λόγο ότι ο αισθητήρας απεικόνισης που χρησιμοποιείται δεν διαχωρίζει την υπέρυθρη ακτινοβολία σε περιοχές (δηλαδή δε λειτουργεί σαν παραδοσιακή κάμερα που χωρίζει το φάσμα του ορατού φωτός σε περιοχές RGB). Αν και υπάρχουν έγχρωμοι αισθητήρες για θερμικές κάμερες, η σημασία τους είναι περιορισμένη γιατί η έγχρωμη απεικόνιση σε περιοχές εκτός του ορατού φάσματος είναι λανθασμένη από την άποψη ότι δεν μπορεί να γίνει αντιληπτή από τον άνθρωπο. Η συνήθης πρακτική στην περίπτωση που δεν επιθυμείται μονόχρωμη απεικόνιση είναι η χρήση της χρωματικής παλέτας pseudocolor.

4.2.1. Θεωρία λειτουργίας

Η υπέρυθρη ενέργεια είναι ένα μέρος του ηλεκτρομαγνητικού φάσματος, η οποία περιλαμβάνει την ακτινοβολία από ακτίνες γάμμα, ακτίνες Χ, υπεριώδη, καθώς επίσης και μια μικρή φασματική ζώνη από τις περιοχές του ορατού φωτός, υπέρυθρου, των μικροκυμάτων και των ραδιοκυμάτων. Οι παραπάνω φασματικές περιοχές συσχετίζονται και διαφοροποιούνται ως προς τα μήκη κύματος τους.

Όλα τα αντικείμενα εκπέμπουν μια συγκεκριμένη ποσότητα ακτινοβολίας μελανού σώματος (black body radiation) ως συνάρτηση της θερμοκρασίας

τους. Γενικά, όσο υψηλότερη είναι η θερμοκρασία ενός αντικειμένου, τόσο περισσότερη είναι η υπέρυθη ακτινοβολία που εκπέμπεται ως black body radiation. Έτσι σε μια απεικόνιση που λήφθηκε από θερμική κάμερα, με πιο ανοικτό τόνο του γκρι, απεικονίζονται τα πιο θερμά αντικείμενα. Η θερμοκρασία του σώματος του ανθρώπου είναι πιο ψηλή από την θερμοκρασία του περιβάλλοντος που τον περιβάλλει, και για αυτό το λόγο σε θερμικές απεικονίσεις μπορεί εύκολα να διακριθεί αφού απεικονίζεται με πιο ανοικτό τόνο από το παρασκήνιο του.

4.2.2 Θερμικός αισθητήρας Tau 2.

Για της ανάγκες του έργου (eVACUATE), επιλέχτηκε ο θερμικός αισθητήρας Tau 2 LWIR της εταιρίας Flir [18]. Τα κυριότερα τεχνικά χαρακτηριστικά του οποίου παρουσιάζονται συνοπτικά στον πίνακα που ακολουθεί:

Χαρακτηριστικά θερμικού αισθητήρα Tau 2	
Spectral range	8-14 μm
Sensitivity	<60 mK
Spatial resolution	620 x 500
Frame rates	9 Hz
Lens	19mm, HFOV from 26 ⁰ to 32 ⁰

Εικ. 4.2.2.1 : Χαρακτηριστικά θερμικού αισθητήρα Tau 2



Εικ. 4.2.2.2 : Θερμική κάμερα Tau 2 [19]

4.2.3 Ψηφιακά αρχεία δεδομένων

Κατά τη διεξαγωγή των πειραμάτων του έργου eVacuate καταγράφηκαν πολλά διαφορετικά σενάρια. Η παρούσα εργασία όσον αφορά τα θερμικά δεδομένα, ασχολείται με την επεξεργασία των αρχείων "AIA_TH_04_1.avi" και "AIA_TH_04_2.avi" που κατέγραψε η θερμική κάμερα Ταυ 2. Και τα δύο αρχεία βίντεο είναι σε μονοχρωματική απεικόνιση και τα χαρακτηριστικά τους παρουσιάζονται στον επόμενο πίνακα. Στη συνέχεια παρουσιάζονται δυο στιγμιότυπα των θερμικών βίντεο (ένα από κάθε αρχείο).

Χαρακτηριστικά θερμικών βίντεο	
Frame width	640 pixels
Frame height	512 pixels
Frame rate	7 fps (frames per second)
Frame captured	420 frames
Video length	58 seconds
Video size	393,0 MB

Εικ. 4.2.3.1 : Χαρακτηριστικά θερμικών βίντεο



Εικ. 4.2.3.2 : Στιγμιότυπο θερμικού βίντεο 1



Εικ. 4.2.3.3 : Στιγμιότυπο θερμικού βίντεο 2

4.3 Δεδομένα από υπερφασματικό αισθητήρα (οπτικά δεδομένα)

Γενικά οι υπερφασματικές κάμερες, σε σχέση με τις παραδοσιακές κάμερες (RGB), χωρίζουν το φάσμα του φωτός σε πολλά μικρά κανάλια μήκους κύματος (wavelength bands). Επομένως μια υπερφασματική κάμερα καταγράφει το φασματικό αποτύπωμα, μία μοναδική φασματική καμπύλη των αντικειμένων που καταγράφει. Αυτό το φασματικό αποτύπωμα ή φασματική υπογραφή δίνει πολύ λεπτομερή πληροφορίες σχετικά με τα υλικά που συνιστούν τα αντικείμενα που απεικονίζονται σε μια εικόνα. Η υπερφασματική απεικόνιση αποτελεί αξιόπιστη λύση στο πρόβλημα της ταυτοποίησης και ταξινόμησης των αντικειμένων και σήμερα θεωρείται ως η βασική τεχνολογία γενικής εφαρμογής για την ανάπτυξη νέας γενιάς τεχνολογιών - εφαρμογών ασφάλειας.

Τα δεδομένα αν και λήφθηκαν από υπερφασματικό αισθητήρα, χαρακτηρίζονται οπτικά, για τον λόγο ότι η καταγραφή τους έγινε στα τρία κανάλια του ορατού φάσματος (κανάλια RGB).

4.3.1: Υπερφασματικός αισθητήρας CMOSIS CMV2000

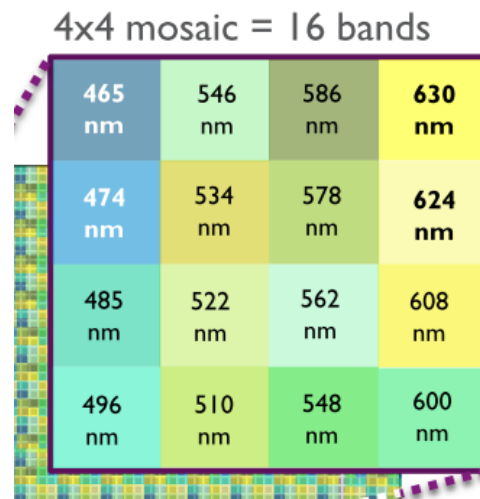
Ο υπερφασματικός αισθητήρας που επιλέχθηκε να γίνει η καταγραφή των δεδομένων είναι ο CMOSIS CMV2000 της εταιρίας Imec [19]. Λεπτομερή χαρακτηριστικά του οποίου παρατίθενται στον πιο κάτω πίνακα.

Χαρακτηριστικά υπερφασματικού αισθητήρα CMOSIS CMV2000	
Φασματική περιοχή	περίπου 400nm – 700nm
Φασματικά κανάλια	16 φασματικά κανάλια
Χωρική ανάλυση	>1,3MP (αισθητήρα)
Ταχύτητα καταγραφής frame	≥ 30 fps (frames per sec)
Φακός	35mm

Εικ. 4.3.1.1: Χαρακτηριστικά υπερφασματικού αισθητήρα CMOSIS CMV2000



Εικ. 4.3.1.2 : Υπερφασματική κάμερα [19]



Εικ. 4.3.1.3 : Φίλτρο 16 φασματικών καναλιών [19]

4.3.2 Ψηφιακά αρχεία δεδομένων

Κατά τη διεξαγωγή των πειραμάτων του έργου eVacuate ο υπερφασματικός αισθητήρας κατέγραψε την εξέλιξη πολλών διαφορετικών σεναρίων. Στο πλαίσιο της συγκεκριμένης εργασίας είναι επαρκές μονάχα ένα αρχείο οπτικού βίντεο για ανάλυση. Έτσι ως δεδομένο για ανάλυση λήφθηκε το αρχείο "AIA_HS_04_1.avi", ένα οπτικό βίντεο σε απεικόνιση grayscale που αποτελείται από τα τρία κανάλια του ορατού φάσματος (στα 465 nm 546 nm και 630 nm). Τα χαρακτηριστικά του παρουσιάζονται στον πίνακα που ακολουθεί:

Χαρακτηριστικά οπτικού βίντεο	
Frame width	480 pixels
Frame height	270 pixels
Frame rate	4 fps (frames per second)
Frame captured	232 frames
Video length	58 seconds
Video size	86,0 MB

Εικ. 4.3.2.1 : Χαρακτηριστικά οπτικού βίντεο

Στην εικόνα παρακάτω παρουσιάζεται ένα στιγμιότυπο του οπτικού βίντεο. Αξίζει να σημειωθεί ότι έγινε μια προ - επεξεργασία των αρχικών δεδομένων που λήφθηκαν από την υπερφασματική κάμερα. Στην συγκεκριμένη εργασία όμως σαν πρωτογενή δεδομένα θεωρείται το οπτικό βίντεο που περιγράφηκε πιο πάνω.



Εικ 4.3.2.2 : Στιγμιότυπο οπτικού βίντεο

5

Ανάπτυξη αλγορίθμων για ανίχνευση και παρακολούθηση ανθρώπινης δραστηριότητας.

5.1 Εισαγωγή

Σε αυτό το κεφάλαιο παρουσιάζεται η προγραμματιστική προσέγγιση για λύση του προβλήματος του αυτόματου εντοπισμού και παρακολούθησης της ανθρώπινης δραστηριότητας σε ένα αρχείο βίντεο.

Σε εφαρμογές ανάλυσης εικόνας και συγκεκριμένα για ανίχνευση και παρακολούθηση αντικειμένων, οι συνηθέστερες προσεγγίσεις είναι μέσω διαδικασιών αφαίρεσης παρασκηνίου (background subtraction) ή μέσω οριοθέτησης (thresholding) του σφάλματος μεταξύ μιας εκτίμησης της στατικής εικόνας σε σχέση με τη παρούσα εικόνα. Προϊόν αυτών των

διαδικασιών είναι συνήθως μια δυαδική μάσκα (binary mask) στην οποία παρουσιάζεται με χρώμα μαύρο το φόντο (τιμή = 0) και με άσπρο χρώμα (τιμή = 1) τα αντικείμενα ενδιαφέροντος, στην περίπτωση μας οι άνθρωποι.

Για τους σκοπούς της παρούσας εργασίας, χρησιμοποιήθηκαν αρχικά δύο αλγόριθμοι. Ο πρώτος βασίζεται στην μέθοδο οριοθέτησης (thresholding) Otsu (αναπτύχθηκε σε προηγούμενο κεφάλαιο) για την κατάτμηση της εικόνας και την εξαγωγή της δυαδικής μάσκας για τα αντικείμενα. Τα αποτελέσματα στο στάδιο του εντοπισμού των αντικειμένων δεν ήταν ικανοποιητικά, με αποτέλεσμα να μην αναπτυχθεί αλγόριθμος για την παρακολούθηση βασισμένος σε αυτή την προσέγγιση.

Ο δεύτερος αλγόριθμος αποτελεί μια ολοκληρωμένη προσέγγιση λύσης του προβλήματος, και τα αποτελέσματα του είναι αρκετά ικανοποιητικά.

Στη συνέχεια παρουσιάζονται οι δύο αυτοί αλγόριθμοι.

5.2 Αλγόριθμος 1

5.2.1 Γενικά

Η πρώτη προσέγγιση στο πρόβλημα της κατάτμησης της εικόνας και την ανίχνευση των κινούμενων αντικειμένων (στην περίπτωση μας των ανθρώπων), είναι βασισμένη στη μέθοδο Otsu που αναπτύχθηκε στην ενότητα 3.1. Αποτελεί μια μη επιβλεπόμενη μέθοδο αυτόματης επιλογής κατωφλίου (threshold) για κατάτμηση εικόνας (picture segmentation). Η βασική αρχή της μεθόδου είναι η επιλογή του κατάλληλου ορίου (threshold) γκρι επιπέδου για την εξαγωγή των αντικειμένων από το παρασκήνιο τους. Η επιλογή αυτού του ορίου γίνεται βάση του ιστογράμματος της εικόνας.

Σε ιδανικές περιπτώσεις το ιστόγραμμα της εικόνας παρουσιάζει μια βαθιά και απότομη κοιλάδα ανάμεσα σε δύο κορυφές (οι δυο κορυφές αντιπροσωπεύουν τα αντικείμενα και το παρασκήνιο αντίστοιχα). Ως κατώφλι επιλέγεται μια τιμή κοντά στο κάτω μέρος αυτής της κοιλάδας.

Η διαδικασία της κατάτμισης της εικόνας και εξαγωγής της δυαδικής μάσκας για κάθε frame γίνεται βάση της επιλεγμένης τιμής κατωφλίου. Στις περισσότερες πραγματικές εικόνες είναι πολύ δύσκολη η ανίχνευση της κοιλάδας στο ιστόγραμμα της εικόνα με αποτέλεσμα η επιλογή του κατωφλίου να μὴν δίνει ικανοποιητικά αποτελέσματα.

Στη συνέχεια θα παρουσιαστεί τόσο αναλυτικά όσο και διαγραμματικά ο αλγόριθμος που χρησιμοποιήθηκε καθώς και τα αποτελέσματα του.

5.2.2 Αναλυτική επεξήγηση αλγορίθμου.

Ο αλγόριθμος αποτελεί μια παραλλαγή του αλγορίθμου που ανέπτυξε ο Brett Shoelson για λογαριασμό της επίσημης ιστοσελίδας της Matlab, Mathworks. Στη συνέχεια παρουσιάζονται τα στάδια του αλγορίθμου, συνοδευόμενα από μια λεπτομερή περιγραφή.

Στάδιο 1: Ανάγνωση αρχείου βίντεο.

Υπάρχουν διάφορα formats αρχείων βίντεο. Ο συγκεκριμένος αλγόριθμος καθορίζει πως ο τύπος αρχείων βίντεο που διαβάζει και επεξεργάζεται είναι σε format Audio Video Interleaved (.avi).

Στάδιο 2: Δημιουργία Video Object.

Για την ανάγνωση ενός αρχείου βίντεο υπάρχουν διάφοροι τρόποι. Στον συγκεκριμένο αλγόριθμο επιλέγηκε η μέθοδος ανάγνωσης βίντεο μέσω ενός video object.

Στάδιο 3: Εξαγωγή πληροφορίας από το αρχείο βίντεο.

Μετά την ανάγνωση του αρχείου, καθορίζονται οι τιμές των παραμέτρων number of frames, frame height και frame width. Η πληροφορία αυτή εξάγεται από το video object που δημιουργήθηκε πιο πάνω.

Στάδιο 4: Καθορισμός διαστάσεων πίνακα mean gray levels.

Στον πίνακα αυτό καταγράφεται η μέση τιμή του γκρι της εικόνας. Η καταγραφή γίνεται σε όλα τα frames του βίντεο και έτσι οι διαστάσεις του πίνακα προκύπτουν ίσες με $[1 \times \text{number of frames}]$. Για την αποφυγή διαστασιολογικών ή υπολογιστικών προβλημάτων στη συνέχεια, οι αρχικές τιμές που περιέχονται στον πίνακα είναι μηδενικές (αντί να έχουμε κενό πίνακα, έχουμε μηδενικό πίνακα).

Στάδιο 5: Συνθήκη επανάληψης για επεξεργασία όλων των frames.

Ένα βίντεο αποτελείται από μια ακολουθία εικόνων. Γενικά όταν μιλάμε για επεξεργασία βίντεο, στην ουσία αυτό που γίνεται είναι η επεξεργασία των frames που αποτελούν το βίντεο. Αρχικά Ορίζεται η τιμή της μεταβλητής $\text{frame} = 1$, και η συνθήκη επανάληψης έχει τη μορφή $\text{if frame} < \text{number of frames}$ then. Στην περίπτωση που το αποτέλεσμα είναι TRUE τότε ο αλγόριθμος συνεχίζει στα επόμενα στάδια επεξεργασίας 6-10. Στην περίπτωση που το αποτέλεσμα της συνθήκης είναι FALSE, δηλαδή έχουν επεξεργαστεί όλα τα frames, τότε ο αλγόριθμος τερματίζεται.

Στάδια επεξεργασίας εικόνας σε επανάληψη για κάθε frame :

Στάδιο 6: Ανάγνωση frame (This frame)

Όπως αναφέρθηκε και πιο πάνω, ένα βίντεο αποτελεί μια ακολουθία frames. Έτσι για να επεξεργαστούμε το βίντεο, πρέπει να επεξεργαστούμε το εκάστοτε frame. Το στάδιο αυτό διαβάζει το frame που ορίζεται ως This frame, το οποίο αποτελεί το frame με δείκτη τη μεταβλητή frame. Παράδειγμα αν η μεταβλητή frame έχει τιμή 8, τότε το this frame είναι το όγδοο frame.

Στάδιο 7: Υπολογισμός μέσης τιμή γκρι του εκάστοτε frame

Σε αυτό το στάδιο γίνεται ο υπολογισμός της μέσης τιμής του συγκεκριμένου frame, και συμπληρώνεται ο πίνακας mean gray levels. Η τιμή αυτή προκύπτει από τον υπολογισμό της μέσης τιμής του γκρι από όλα τα pixels του εκάστοτε frame.

Στάδιο 8: Καθορισμός ποσότητας α .

Το α αντιπροσωπεύει μια τιμή η οποία χρησιμοποιείται στη συνέχεια στο στάδιο υπολογισμού του παρασκηνίου του frame. Στο συγκεκριμένο αλγόριθμο η τιμή του α ορίστηκε να είναι ίση με 0,5.

Στάδιο 9: Συνθήκη `if frame = 1` και καθορισμός του παρασκηνίου.

Στην περίπτωση που η πιο πάνω συνθήκη έχει αποτέλεσμα TRUE, δηλαδή πρόκειται για το πρώτο frame, τότε ως παρασκήνιο ορίζεται το πρώτο frame, δηλαδή το πρώτο frame που διαβάστηκε στο στάδιο 6 (this frame).

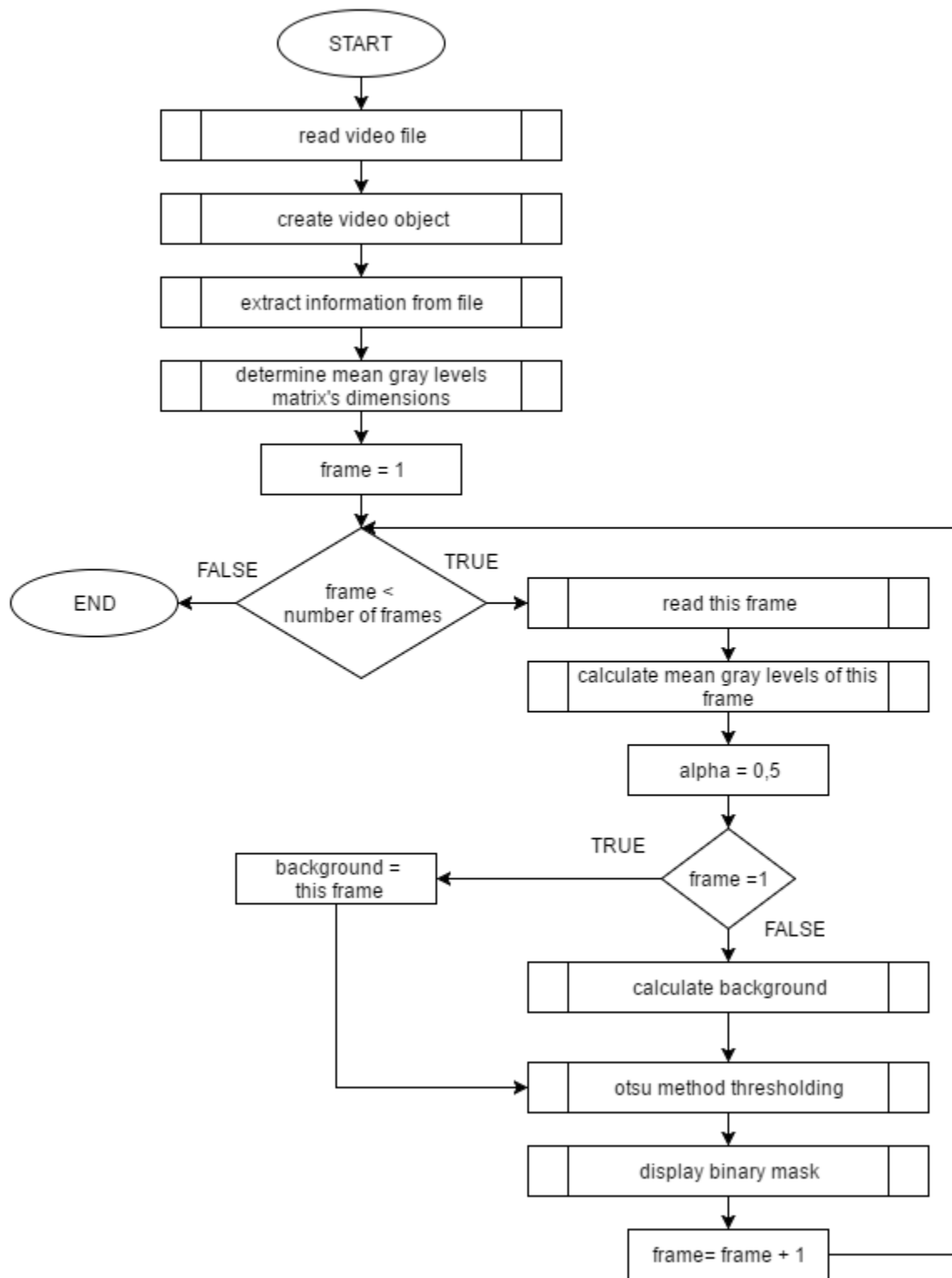
Σε όλες τις άλλες περιπτώσεις, δηλαδή στις περιπτώσεις που το αποτέλεσμα είναι FALSE, το παρασκήνιο προκύπτει από την παρακάτω σχέση.

$$\text{background} = (1 - \alpha) * \text{this frame} + \alpha * \text{background}$$

Στάδιο 10. Εφαρμογή μεθόδου Otsu για την εξαγωγή της δυαδικής μάσκας του frame.

Στο στάδιο αυτό εφαρμόζεται η μέθοδος Otsu που αναπτύχτηκε στην ενότητα 3.1 που έχει ως προϊόν την εξαγωγή της δυαδικής μάσκας με τα αντικείμενα να απεικονίζονται με λευκό χρώμα ενώ το παρασκήνιο με μαύρο.

Στην επόμενη σελίδα παρουσιάζεται διαγραμματικά ο αλγόριθμος.



Εικ. 5.2.2.1 : Διάγραμμα γενικής περιγραφής του αλγόριθμου 1 (Otsu Thresholding Method)

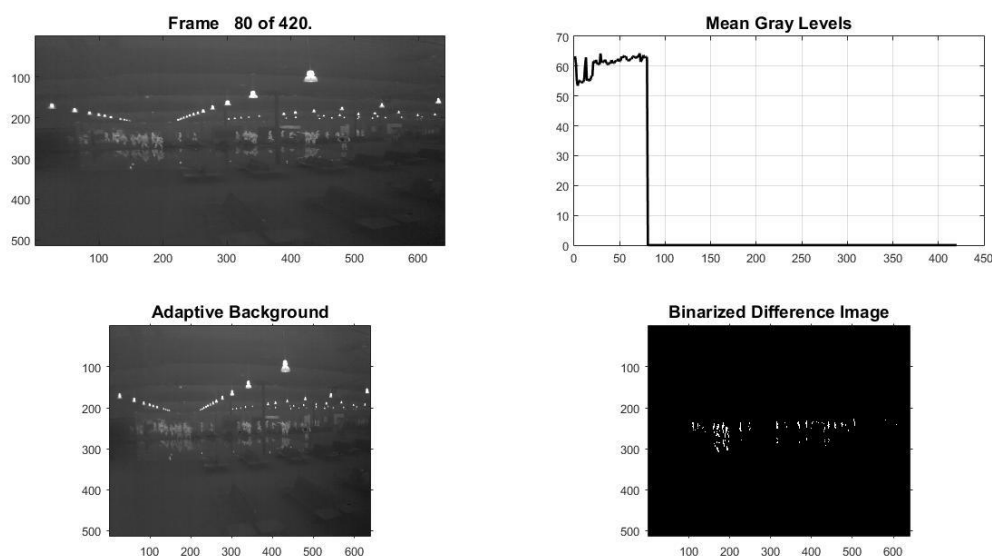
5.2.3 Αποτελέσματα - Σχολιασμός

Κριτήριο αξιολόγησης στο στάδιο της ανίχνευσης της ανθρώπινης δραστηριότητας αποτελεί η αποτελεσματική απεικόνιση των ανθρώπων στη δυαδική μάσκα. Στο παρόν στάδιο αποτέλεσε και το μοναδικό κριτήριο.

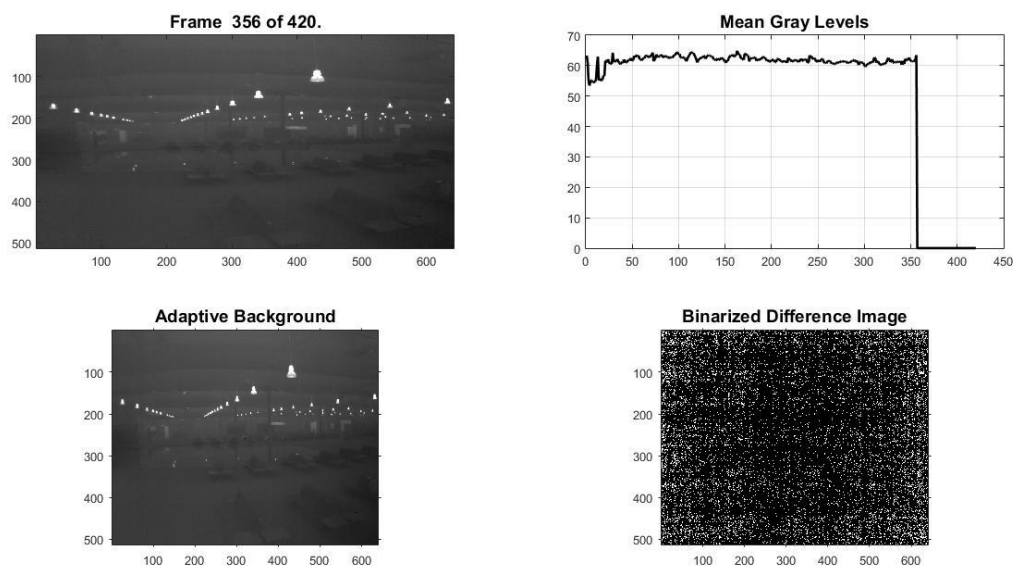
Στην περίπτωση των βίντεο από τη θερμική απεικόνιση, τα αποτελέσματα κατά τη διάρκεια που παρουσιαζόταν ανθρώπινη κίνηση στην εικόνα ήταν πολύ ενθαρρυντικά όπως εύκολα διακρίνουμε στην εικόνα 5.2.3.1. Στις περιπτώσεις όμως που το πεδίο καταγραφής ήταν κενό, ο αλγόριθμος κατέγραφε θόρυβο, γεγονός που θα δημιουργούσε πρόβλημα στο επόμενο στάδιο της παρακολούθησης. Η τελευταία περίπτωση παρουσιάζεται στην εικόνα 5.2.3.2.

Τα αποτελέσματα της εφαρμογής του αλγορίθμου στο βίντεο από την υπερφασματική κάμερα δεν ήταν ικανοποιητικά ούτε στην περίπτωση ύπαρξης ανθρώπινης δραστηριότητας. Παρουσιαζόταν ασυνέχεια στο σχηματισμό των φιγούρων ή ακόμη και καθόλου ανίχνευση των ανθρώπων.

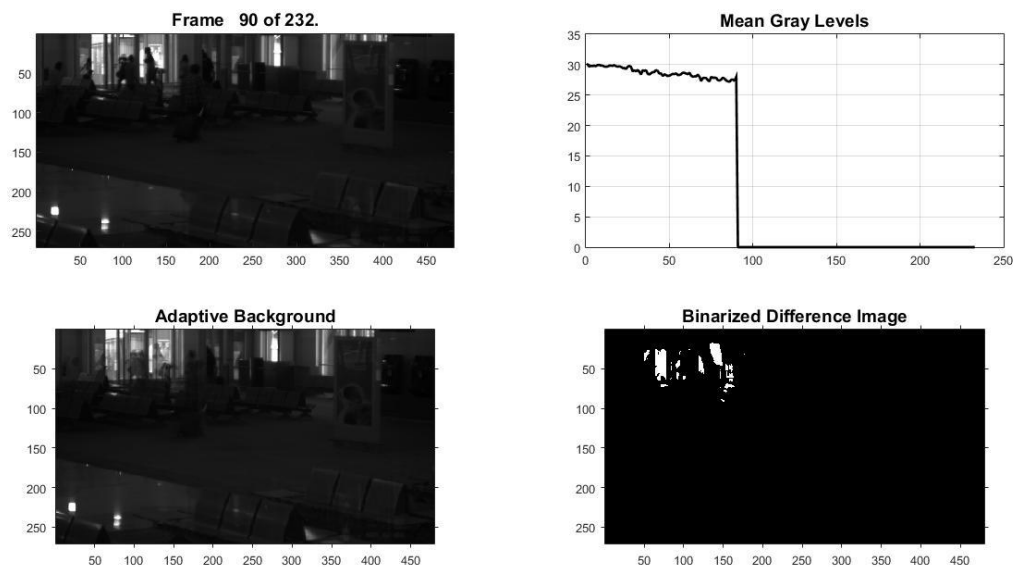
Λαμβάνοντας υπόψη τα παραπάνω αποφασίστηκε πως ο συγκεκριμένος αλγόριθμος δεν θα μπορεί να δώσει ικανοποιητικά αποτελέσματα ούτε στο στάδιο της παρακολούθησης. Έτσι αναπτύχθηκε ένας δεύτερος αλγόριθμος με διαφορετική προσέγγιση στο πρόβλημα.



Εικ.5.2.3.1: Αποτελέσματα αλγορίθμου σε περίπτωση ύπαρξης ανθρώπινης δραστηριότητας (θερμικό βίντεο)



Εικ 5.2.3.2: Αποτελέσματα αλγορίθμου στην περίπτωση μη ύπαρξης ανθρώπινης δραστηριότητας (θερμικό βίντεο)



Εικ 5.2.3.3: Αποτελέσματα αλγορίθμου στο βίντεο που κατέγραψε η υπερφασματική κάμερα

5.3 Αλγόριθμος 2

5.3.1 Γενικά

Ο αλγόριθμος αυτός αποτελεί μια ολοκληρωμένη προσέγγιση λύσης στο πρόβλημα της αυτόματης ανίχνευσης και παρακολούθησης κινούμενων αντικειμένων από ένα βίντεο που καταγράφηκε από ακίνητη κάμερα. Στην περίπτωση μας, ως κινούμενα αντικείμενα θεωρούνται οι άνθρωποι, δηλαδή ο αλγόριθμος ανιχνεύει την ανθρώπινη δραστηριότητα. Ο αλγόριθμος είναι βασισμένος στην κίνηση (motion - based) των αντικειμένων και χωρίζεται σε δύο βασικά μέρη:

- Την ανίχνευση των κινούμενων αντικειμένων σε κάθε frame.
- Την συσχέτιση των ανιχνεύσεων που καταγράφηκαν στα διάφορα frames έτσι ώστε να ανταποκρίνονται στο ίδιο αντικείμενο κατά την πάροδο του χρόνου. (παρακολούθηση)

Για την ανίχνευση των κινούμενων αντικειμένων χρησιμοποιήθηκε ένας αλγόριθμος αφαίρεσης παρασκηνίου (background subtraction algorithm) βασισμένος σε Gaussian mixtures models (ενότητα 2.3). Στη συνέχεια εφαρμόζονται στη μάσκα του προσκηνίου (foreground mask) μορφολογικές διαδικασίες με σκοπό την εξάλειψη του θορύβου. Τέλος μέσω ανάλυσης blob γίνεται η ανίχνευση ομάδων από συνδεδεμένα pixels, τα οποία πιθανόν να ανήκουν σε κινούμενα αντικείμενα.

Για τον αλγόριθμο κάθε αντικείμενο αποτελεί ένα ίχνος (track). Ο συσχετισμός των ανιχνεύσεων για το ίδιο track γίνεται αποκλειστικά βάση της κίνησης των αντικειμένων. Η κίνηση κάθε track εκτιμάται μέσω του φίλτρου Kalman που αναπτύχθηκε στην ενότητα 3.2. Το φίλτρο Kalman χρησιμοποιείται για την εκτίμηση της θέσης που θα έχει το κάθε αντικείμενο σε κάθε frame, και καθορίζει την πιθανότητα κάθε ανίχνευσης να αποτελεί την επόμενη θέση των προϋπαρχόντων tracks.

Η διατήρηση της συνοχής των tracks κατά τη διάρκεια εξέλιξης του βίντεο αποτελεί σημαντικό κομμάτι του αλγορίθμου. Σε κάθε δοσμένο frame, μέρος του συνόλου των ανιχνεύσεων, ανατίθενται σε tracks (assigned tracks), ενώ κάποιες ανιχνεύσεις και tracks παραμένουν απροσδιόριστα (unassigned).

Στην περίπτωση των προσδιορισμένων tracks γίνεται ο εκσυγχρονισμός τους (update) βάση της αντίστοιχης ανίχνευσης. Τα μη προσδιορισμένα tracks θεωρούνται από τον αλγόριθμο αόρατα. Μια μη προσδιορισμένη (unassigned) ανίχνευση συνεπάγεται τη δημιουργία καινούριου track.

Για κάθε track καταγράφεται ο αριθμός των frames που παραμένει συνεχόμενα αόρατο. Αν ο αριθμός αυτός ξεπερνά ένα συγκεκριμένο όριο, τότε το track αυτό θεωρείται ότι έχει φύγει από το οπτικό πεδίο της κάμερας, με αποτέλεσμα το συγκεκριμένο track να διαγράφεται.

Καθ' όλη τη διάρκεια που ο αλγόριθμος επεξεργάζεται τα frames του βίντεο, ταυτόχρονα με το frame που τυγχάνει επεξεργασίας παρουσιάζεται και η δυαδική μάσκα που παρουσιάζει τις ανιχνεύσεις με τον αύξον αριθμό των tracks.

Στη συνέχεια ακολουθεί η αναλυτική και διαγραμματική επεξήγηση του αλγορίθμου, καθώς και τα αποτελέσματα του.

5.3.2 Αναλυτική επεξήγηση αλγορίθμου.

Ο αλγόριθμος που αναπτύχθηκε αποτελεί μια παραλλαγή του αλγορίθμου "Motion-based Multiple Object Tracking" της Mathworks. Λόγο της έκτασης του αλγορίθμου και του γεγονότος ότι αποτελείται από πολλές συναρτήσεις, θεωρήθηκε πως μια γραμμική λεπτομερής επεξήγηση του αλγορίθμου στο σύνολο του, θα ήταν δυσνόητη. Έτσι επιλέγηκε η παρουσίαση μιας γενικής περιγραφής - επεξήγησης του αλγορίθμου, με ονομαστικές αναφορές στις συναρτήσεις, που θα συνοδεύεται στη συνέχεια με λεπτομερή περιγραφή των συναρτήσεων που χρίζουν περαιτέρω επεξήγησης.

Στο σημείο αυτό αξίζει να σημειωθεί ότι ο αλγόριθμος που αναπτύχθηκε, εκτός από την ανίχνευση και παρακολούθηση της ανθρώπινης δραστηριότητας, είχε επίσης ως σκοπό την εξαγωγή ενός dataset που θα χρησιμοποιηθεί στη συνέχεια για περαιτέρω επεξεργασία. Το dataset αυτό θα αποτελεί στην ουσία μια χρονοσειρά των centroids των tracks καθ' όλη τη διάρκεια του βίντεο.

Οι διαδικασίες και τα βήματα του αλγορίθμου και των συναρτήσεων παρουσιάζονται και σε διαγραμματική μορφή. Στο τέλος της ενότητας θα παρουσιαστούν τα αποτελέσματα και ο σχολιασμός τους.

❖ **Γενική περιγραφή αλγορίθμου σε στάδια**

Στάδιο 1: Δημιουργία System Objects.

Στο πρώτο στάδιο γίνεται η προετοιμασία για την επεξεργασία (καθορίζονται input / output) του αρχείου βίντεο. Η συνάρτηση που αποτελεί το πρώτο αυτό στάδιο, δημιουργεί objects, δηλαδή λογικές δομές (logical structures), από την ανάγνωση του αρχείου. Στη συνέχεια εξαγονται τα objects που ανιχνεύονται σε κάθε frame.

Στάδιο 2: Προετοιμασία των tracks.

Κάθε ξεχωριστό αντικείμενο που εντοπίζει ο αλγόριθμος το θεωρεί ως ένα track. Κάθε track αποτελεί ένα διάνυσμα από πληροφορίες για το αντικείμενο. Στο στάδιο αυτό καθορίζονται ποιές είναι αυτές οι πληροφορίες που περιέχει το κάθε track.

Στάδιο 3: Καθορισμός δεικτών.

Καθορίζονται δυο δείκτες να είναι ίσοι με το 1. Ο πρώτος ονομάζεται nextId και αναφέρεται στην ταυτότητα του επόμενου track. Ο δεύτερος με ονομασία counter, είναι ο δείκτης που θα χρησιμοποιηθεί στο στάδιο καταγραφής του dataset.

Στάδιο 4: Διαστασιολόγηση πινάκων που αφορούν τη δημιουργία του dataset.

Ο πίνακας x περιέχει την συντεταγμένη x του centroid κάθε track.

Ο πίνακας y περιέχει την συντεταγμένη y του centroid κάθε track.

Ο πίνακας id περιέχει την ταυτότητα κάθε track.

Οι τρεις αυτοί πίνακες έχουν διαστάσεις 1 x 100. Το 100 επιλέγηκε ως ο μέγιστος αριθμός tracks που μπορούν να εμφανιστούν σε ένα βίντεο.

Ο πίνακας `data of frames` κατά τη διάρκεια εκτέλεσης του αλγορίθμου θα περιέχει στην πρώτη γραμμή του τον πίνακα `x`, στη δεύτερη τον πίνακα `y` και στη τρίτη τον πίνακα `id`. Έτσι οι διαστάσεις του προκύπτουν 3×100 .

Ο πίνακας `new data of frames` ορίζεται αρχικά ως ένας μηδενικός πίνακας ίδιων διαστάσεων με τον `data of frame`.

Ο πίνακας `dataset` είναι ένας πίνακας τριών διαστάσεων ο οποίος στην πράξη θα καταγράφει τον πίνακα `new data of frames` σε κάθε `frame` του βίντεο. Έτσι οι διαστάσεις του πίνακα είναι $3 \times 100 \times \text{number of frames}$. Αρχικά ορίζουμε αυτόν τον πίνακα ως μηδενικό πίνακα.

Στάδιο 5: Συνθήκη επανάληψης.

Τα επόμενα στάδια του αλγορίθμου αποτελούν μια επαναληπτική διαδικασία για όλα τα `frames` του βίντεο. Έτσι το στάδιο αυτό αποτελεί μια συνθήκη της μορφής `while object reader is not done then`:

- Στην περίπτωση που το αποτέλεσμα είναι `TRUE` τότε ο αλγόριθμος συνεχίζει στα επόμενα στάδια (στάδια 6-13)
- Στην περίπτωση που το αποτέλεσμα είναι `FALSE` σημαίνει ότι έχουν επεξεργαστεί όλα τα `frame` και το επόμενο βήμα του αλγορίθμου είναι το στάδιο 14.

Στάδιο 6: Ανάγνωση `frame`.

Στο στάδιο αυτό γίνεται η ανάγνωση του εκάστοτε `frame` του βίντεο. Η διαδικασία αυτή γίνεται μέσω του `object reader`, `step by step`.

Στάδιο 7: Ανίχνευση αντικειμένων και εξαγωγή πληροφοριών.

Μέσω της συνάρτησης `detectObject` γίνεται η ανίχνευση των `tracks` και η καταγραφή των πληροφοριών : `centroids`, `bboxes`, `mask`.

Στάδιο 8: Εκτίμηση καινούριας τοποθεσίας των tracks.

Εδώ γίνεται η εκτίμηση της τοποθεσίας που θα έχουν τα tracks, μέσω του Kalman Filter

Στάδιο 9: Ανάθεση ανίχνευσης στα tracks.

Η ανάθεση των αντικειμένων που ανιχνεύτηκαν στο παρόν frame στα ήδη υπάρχοντα tracks γίνεται μέσω της ελαχιστοποίησης του κόστους. Το κόστος ορίζεται ως ο αρνητικός λογάριθμος των ανιχνεύσεων που αντιστοιχούν σε κάθε track.

Στάδιο 10: Εκσυγχρονισμός προσδιορισμένων tracks.

Στην περίπτωση που καταγεγραμμένα tracks προσδιορίζονται στο frame (δηλαδή υπάρχει αντιστοίχιση ανίχνευσης σε προϋπάρχον track) τότε γίνεται ο εκσυγχρονισμός τους σύμφωνα με την τοποθεσία της ανίχνευσης.

Στάδιο 11: Εκσυγχρονισμός μη προσδιορισμένων tracks.

Στην περίπτωση που σε κάποια tracks δεν βρεθεί αντιστοιχία με κάποια ανίχνευση, αυτά τα tracks μαρκάρονται ως αόρατα. Κάθε φορά που δεν υπάρχει ανίχνευση που να αντιστοιχεί σε αυτά τα tracks, αυξάνεται η "ηλικία" τους κατά 1.

Στάδιο 12: Διαγραφή χαμένων tracks.

Η συνάρτηση που αποτελεί αυτό το στάδιο, διαγράφει τα tracks που είναι αόρατα για πολλά συνεχόμενα frames. Διαγράφει επίσης και νεοεισαχθείσα tracks τα οποία ήταν αόρατα για πάρα πολλά frames στο σύνολο.

Στάδιο 13: Δημιουργία καινούριων tracks.

Οι ανιχνεύσεις που γίνονται και δεν έχουν ανατεθεί σε κάποιο προϋπάρχον track, δημιουργούν καινούρια tracks.

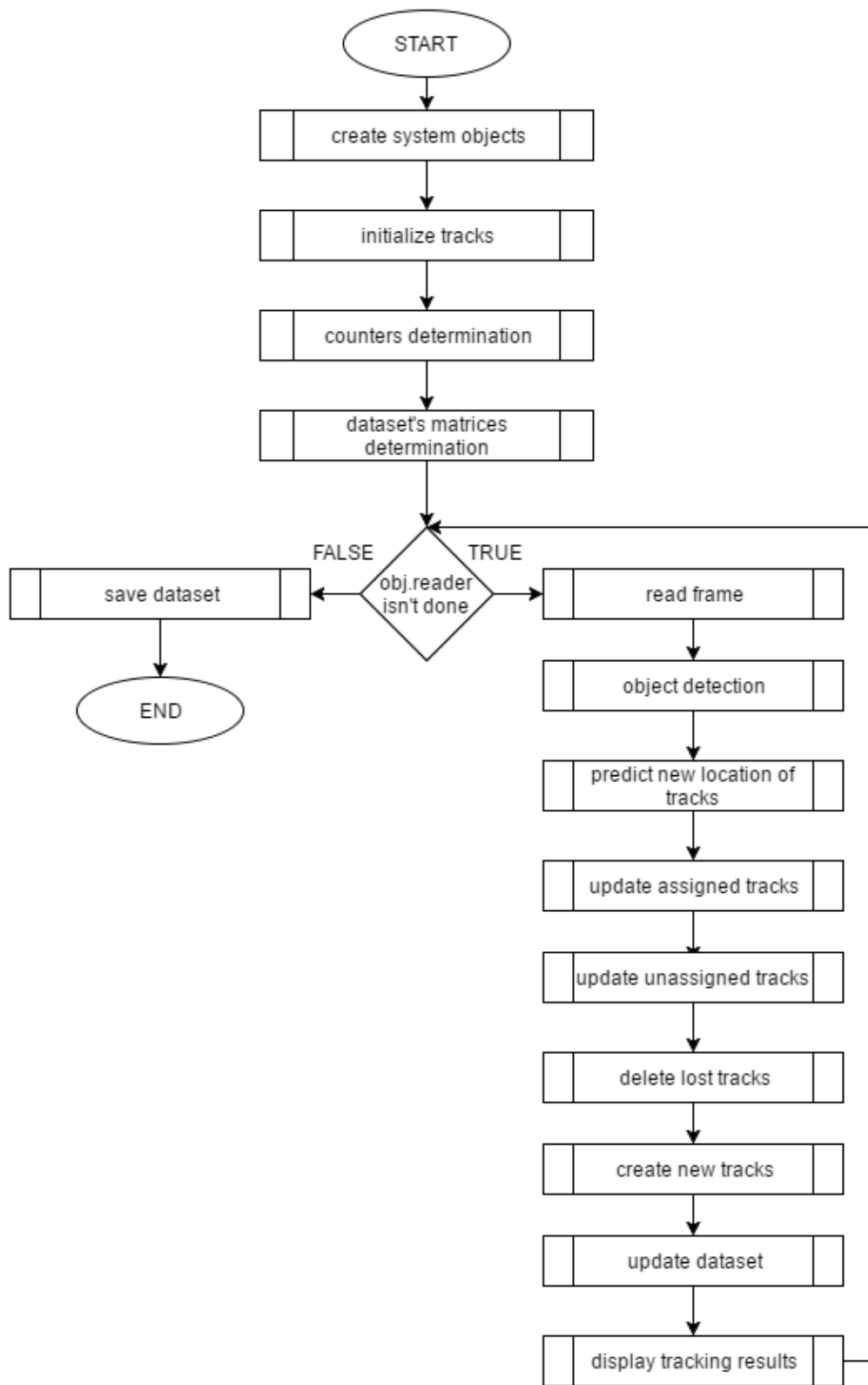
Στάδιο 14: Υπολογισμών πινάκων που αποτελούν το dataset και παρουσίαση αποτελεσμάτων.

Στο στάδιο αυτό καθορίζονται ποιά tracks είναι αξιόπιστα (Reliable tracks). Για να είναι αξιόπιστο ένα track πρέπει να έχει εμφανιστεί σε τουλάχιστον οκτώ συνεχόμενα frames. Στη συνέχεια, εφόσον το track είναι αξιόπιστο, καταγράφονται οι πληροφορίες του στους πίνακες id,x,y. Τέλος γίνεται η προβολή του αρχικού εκάστοτε frame καθώς και της δυαδικής μάσκας, με ζωγραφισμένα πάνω τους τα bbox και το trackId των ανιχνεύσεων. Για καλύτερη κατανόηση του αλγορίθμου, στις περιπτώσεις που το track δεν κινείται προς την τοποθεσία που εκτιμήθηκε πως θα έχει στο επόμενο frame σύμφωνα με το Kalman filter, ζωγραφίζεται στη εικόνα αυτή η εκτιμώμενη περιοχή με ετικέτα "predicted".

Στάδιο 15: Καταγραφή dataset.

Στο στάδιο αυτό καταγράφεται το dataset που θα χρησιμοποιηθεί σε επόμενη ενότητα. Αρχικά κατασκευάζεται ο 2D πίνακας data of frames ο οποίος περιέχει τους πίνακες x,y,id με τα αντίστοιχα στοιχεία των tracks του frame. Για την αποφυγή διαστασιολογικών και υπολογιστικών προβλημάτων καταχωρούμε τις εγγραφές του πίνακα data of frame στον πίνακα new data of frame. Τέλος ανάλογα με το δείκτη count γίνεται η καταγραφή του πίνακα new data of frame στη θέση του δείκτη count στον 3D πίνακα dataset. Τέλος αυξάνεται ο δείκτης count κατά 1.

Μετά την αποθήκευση του dataset, τερματίζεται ο αλγόριθμος. Ακολουθεί το διάγραμμα 5.3.2.1 με την γενική περιγραφή του αλγορίθμου.



Εικ 5.3.2.1 : Διάγραμμα γενικής περιγραφής του αλγόριθμου 2.

- ❖ Λεπτομερής περιγραφή συναρτήσεων αλγορίθμου.

Στο σημείο αυτό γίνεται η λεπτομερής περιγραφή των συναρτήσεων που αναφέρθηκαν πιο πάνω ως στάδια. Δηλαδή θα αναλυθούν εις βάθος τα στάδια που χρίζουν περαιτέρω επεξήγησης.

Στάδιο 1: Δημιουργία System Objects.

Στο στάδιο αυτό γίνεται η προετοιμασία για την επεξεργασία του αρχείου βίντεο.

Βήμα 1: Δημιουργία object για ανάγνωση του αρχείου.

```
obj.reader = vision.VideoFileReader('file_name.avi')
```

Βήμα 2: Δημιουργία δυο Objects ως Video players.

Το πρώτο video player παρουσιάζει το αρχικό βίντεο ενώ το δεύτερο τη δυαδική μάσκα του αρχικού βίντεο. Για τα δυο βίντεο player καθορίζονται και οι συντεταγμένες του τετράπλευρου που θα καταλαμβάνουν στην οθόνη.

```
obj.videoPlayer = vision.VideoPlayer('Position', [20, 400, 700, 400]);  
obj.maskPlayer = vision.VideoPlayer('Position', [740, 400, 700, 400]);
```

Βήμα 3: Δημιουργία object για ανίχνευση προσκηνίου (foreground detection).

Η ανίχνευση του προσκηνίου χρησιμοποιείται για τον διαχωρισμό των κινούμενων αντικειμένων από το παρασκήνιο. Προϊόν της συνάρτησης είναι η εξαγωγή της δυαδικής μάσκας (binary mask) του frame, στην οποία τα pixels, με τιμή 1 αντιπροσωπεύουν το προσκήνιο και αυτά με τιμή 0 αντιπροσωπεύουν το παρασκήνιο.

Καθορίζονται οι παράμετροι της συνάρτησης `ForegroundDetector`

```
obj.detector = vision.ForegroundDetector('NumGaussians', 3,  
... 'NumTrainingFrames', 10, 'MinimumBackgroundRatio', 0.7);
```

Η παράμετρος `'NumGaussians'` αναφέρεται στον αριθμό των Gaussian που θα χρησιμοποιηθούν στο μοντέλο μίξης (Gaussian mixture model). Στη συγκεκριμένη εργασία επιλέγηκε ο αριθμός αυτός να είναι ίσος με 3. Η τιμή αυτή συνήθως κυμαίνεται μεταξύ 3 - 5.

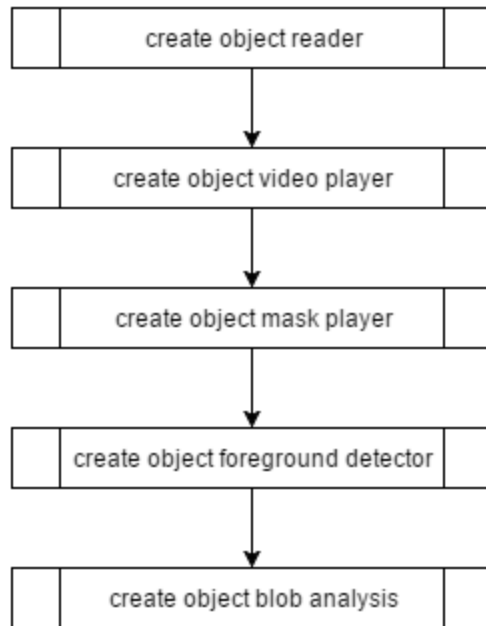
Η παράμετρος `'NumTrainingFrames'` αφορά τον αριθμό των frames που θα θεωρηθούν από τον αλγόριθμο ως frames εκπαίδευσης (training frames). Μετά από δοκιμές η τιμή αυτή ορίστηκε να είναι ίση με 10.

Τέλος η παράμετρος `'MinimumBackgroundRatio'` είναι η ελάχιστη αναλογία παρασκηνίου. Στην ουσία αποτελεί ένα όριο για τον καθορισμό του background model. Η τιμή του ορίου αυτού επιλέγηκε να είναι ίση με 0,7 στα θερμικά δεδομένα και 0,4 στα οπτικά.

Βήμα 4: Δημιουργία object για blob ανάλυση.

Συνδεδεμένες ομάδες από pixels προσκηνίου είναι πιθανό να αντιστοιχούν σε κινούμενα αντικείμενα. Για τον εντοπισμό τέτοιων ομάδων γίνεται η χρήση της συνάρτησης `blobAnalysis`, η οποία υπολογίζει τα χαρακτηριστικά αυτών των ομάδων (pixels). Τα χαρακτηριστικά είναι : το μέγεθος (area), το centroid και το bounding box.

```
obj.blobAnalyser =  
vision.BlobAnalysis('BoundingBoxOutputPort', ...  
true, 'AreaOutputPort', true, 'CentroidOutputPort', true,  
... 'MinimumBlobArea', 0);
```



Εικ 5.3.2.2 : Διάγραμμα σταδίου δημιουργίας System objects

Στάδιο 2: Προετοιμασία των tracks.

Η συνάρτηση αυτού του σταδίου δημιουργεί μια παράταξη (array) από tracks, όπου κάθε track είναι μια δομή (structure) που αντιπροσωπεύει ένα κινούμενο αντικείμενο στο βίντεο. Ο σκοπός αυτής της δομής είναι η διατήρηση της κατάστασης (maintain the state) των αντικειμένων που ανιχνεύτηκαν.

Η κατάσταση περιέχει πληροφορίες που χρησιμοποιούνται για την ανάθεση ανιχνεύσεων σε tracks, για τον τερματισμό των tracks και για την παρουσίαση τους.

Η δομή (structure) των tracks περιέχει τα ακόλουθα πεδία:

- id : είναι η ταυτότητα του track (ακέραιος αριθμός).
- bbox: αποτελεί το παρόν πλαίσιο οριοθέτησης (bounding box) του αντικειμένου. Χρησιμοποιείται στο κομμάτι της παρουσίασης.

- Kalman filter: Αναφέρεται στο object Kalman filter που χρησιμοποιείται για την παρακολούθηση των αντικειμένων σε motion based εφαρμογές.
- age : Αντιπροσωπεύει τον αριθμό των frames που έχουν περάσει από τότε που εντοπίστηκε για πρώτη φορά το track.
- total visible count : Είναι το πεδίο που αποθηκεύει τον αριθμό του συνόλου των frames που έχει εντοπιστεί το track.
- consecutive invisible count : Στο πεδίο αυτό αποθηκεύεται ο αριθμός των συνεχόμενων frames για τα οποία το track δεν έχει ανιχνευτεί (αόρατο).

```
function tracks = initializeTracks()

    tracks = struct(...
        'id', {}, ...
        'bbox', {}, ...
        'kalmanFilter', {}, ...
        'age', {}, ...
        'totalVisibleCount', {}, ...
        'consecutiveInvisibleCount', {});

end
```

Ανιχνεύσεις που αποτελούν θόρυβο, τείνουν να παρουσιάζουν μικρή "διάρκεια ζωής". Δηλαδή εμφανίζονται σε πολύ λίγα frames.

Στάδιο 7: Ανίχνευση αντικειμένων και εξαγωγή πληροφοριών (σε κάθε frame).

Η συνάρτηση detect objects επιστρέφει τα centroids και τα bounding boxes των αντικειμένων που ανιχνεύτηκαν. Επιστρέφει επίσης τη δυαδική μάσκα η οποία έχει το ίδιο μέγεθος με το frame εισόδου.

Η συνάρτηση εκτελεί κατάτμηση κίνησης (motion segmentation) χρησιμοποιώντας ανιχνευτή προσκηνίου. Τέλος εκτελούνται μορφολογικές

διαδικασίες στη δυαδική μάσκα για την αφαίρεση του θορύβου (noisy pixels) και για το "γέμισμα" των καινών στα εναπομείναντα blobs.

Βήμα 1: Ανίχνευση προσκηνίου.

```
mask = obj.detector.step(frame);
```

Βήμα 2: Εφαρμογή μορφολογικών διαδικασιών για αφαίρεση θορύβου και γέμισμα καινών.

```
mask = imopen(mask, strel('rectangle', [3,3]));  
mask = imclose(mask, strel('rectangle', [10,10]));  
mask = imfill(mask, 'holes');
```

Βήμα 3: Εκτέλεση blob analysis για την εύρεση συνδεδεμένων pixels.

```
[~, centroids, bboxes] = obj.blobAnalyser.step(mask);
```

Στάδιο 9: Ανάθεση ανίχνευσης στα tracks.

Η ανάθεση των αντικειμένων που ανιχνεύτηκαν στο παρόν frame, στα ήδη υπάρχοντα tracks, γίνεται μέσω της ελαχιστοποίησης του κόστους. Ως κόστος ορίζεται ο αρνητικός λογάριθμος των ανιχνεύσεων που αντιστοιχούν σε κάθε track.

Ο αλγόριθμος της παρούσας συνάρτησης εκτελείται σε δυο στάδια.

Στάδιο 1:

Ο υπολογισμός του κόστους ανάθεσης κάθε ανίχνευσης σε κάθε track γίνεται με τη χρήση της μεθόδου <distance> (απόστασης) από τη συνάρτηση <vision.KalmanFilter>. Το κόστος λαμβάνει υπόψη την Ευκλείδεια απόσταση μεταξύ του centroid της εκτίμησης της θέσης του track (που έγινε μέσω του Kalman filter) και της θέσης του centroid της ανίχνευσης του track. Μέσα στη διαδικασία υπολογισμού του κόστους περιέχεται και ο παράγοντας αξιοπιστία της πρόβλεψης, που παρέχεται από το Kalman filter. Τα αποτελέσματα αυτής της διαδικασίας

αποθηκεύονται σε ένα πίνακα διαστάσεων $M \times N$, όπου M είναι ο αριθμός των tracks και N ο αριθμός των ανιχνεύσεων.

Στάδιο 2:

Στο στάδιο αυτό γίνεται η λύση του προβλήματος ανάθεσης του πίνακα κόστους που δημιουργήθηκε πριν, με τη χρήση της συνάρτησης `<assignedDetectionToTracks>`. Η συνάρτηση παίρνει ως δεδομένα τον πίνακα κόστους και τη τιμή του κόστους που προκύπτει από τη μη ανάθεση καμίας ανίχνευσης στο track.

Η τιμή του κόστους της μη ανάθεσης καμίας ανίχνευσης στο track (`costOfNonAssignment`) εξαρτάται από το φάσμα των τιμών που επιστρέφει η μέθοδος `<distance>` της συνάρτησης `<vision.KalmanFilter>`. Αυτή η τιμή πρέπει να καθοριστεί πειραματικά. Αν τεθεί πολύ χαμηλά, αυξάνεται η πιθανότητα της δημιουργίας καινούριου track, ή και την διάσπαση των tracks. Αν τεθεί πολύ υψηλή μπορεί ο αλγόριθμος να εντοπίζει μόνο ένα track που να περιέχει όλα τα κινούμενα αντικείμενα.

Μετά τη διεξαγωγή αρκετών επαναλήψεων στον αλγόριθμο, τα καλύτερα αποτελέσματα προέκυψαν θέτοντας `costOfNonAssignment = 20`.

Αξίζει να σημειωθεί ότι η συνάρτηση `AssignDetectiontoTracks` χρησιμοποιεί την έκδοση του Hungarian algorithm που πρότεινε ο Munkres [20]. Ο Hungarian algorithm υπολογίζει και αντιστοιχεί τις αναθέσεις των ανιχνεύσεων στα tracks, με τρόπο που να ελαχιστοποιείται το συνολικό κόστος.

Η συνάρτηση επιστρέφει ένα πίνακα $M \times 2$ που περιέχει τους δείκτες των προσδιορισμένων ανιχνεύσεων και tracks στις δυο του στήλες. Επιστρέφει επίσης και τους δείκτες των ανιχνεύσεων και tracks για τα οποία προέκυψε κάποια αντιστοίχιση.

```

function[assignments,unassignedTracks,unassignedDetections]
    = ...detectionToTrackAssignment()

nTracks = length(tracks);
nDetections = size(centroids, 1);

% Compute the cost of assigning each detection to each
track.
cost = zeros(nTracks, nDetections);
for i = 1:nTracks
    cost(i, :) = distance(tracks(i).kalmanFilter,
        centroids);
end

% Solve the assignment problem.
costOfNonAssignment = 20;
[assignments, unassignedTracks, unassignedDetections] =
    assignDetectionsToTracks(cost, costOfNonAssignment);
end

```

5.3.3 Αποτελέσματα - Σχολιασμός.

Τα αποτελέσματα της λειτουργίας του αλγορίθμου είναι οπτικοποιημένα (σε μορφή εικόνων) για να μπορεί να γίνει η σωστή αξιολόγησή τους. Σαν αποτελέσματα προς αξιολόγηση θα είναι η ικανοποιητική απεικόνιση των αντικειμένων στην δυαδική μάσκα, η σωστή ανίχνευση των αντικειμένων και ο διαχωρισμός τους καθώς και η συνέπεια του αλγορίθμου στο κομμάτι της παρακολούθησης της πορείας κάθε αντικειμένου στο χώρο και στον χρόνο (tracking).

Παρακάτω ακολουθεί ο σχολιασμός των αποτελεσμάτων. Λόγω του ότι τα αποτελέσματα που προέκυψαν από τα δύο θερμικά βίντεο είναι πανομοιότυπα, για να αποφευχθεί η επανάληψη, ο σχολιασμός των αποτελεσμάτων στην περίπτωση των θερμικών δεδομένων θα γίνει μόνο για την περίπτωση του αρχείου "AIA_TH_04_1.avi".

5.3.3.1 Σχολιασμός αποτελεσμάτων εφαρμογής του αλγορίθμου στα θερμικά δεδομένα.

Καταρχάς πρέπει να διαχωριστούν τα frames του βίντεο προς επεξεργασία σε 3 κατηγορίες ανάλογα με το μέγεθος του πλήθους που απεικονίζεται στο εκάστοτε frame. Οι κατηγορίες αυτές είναι:

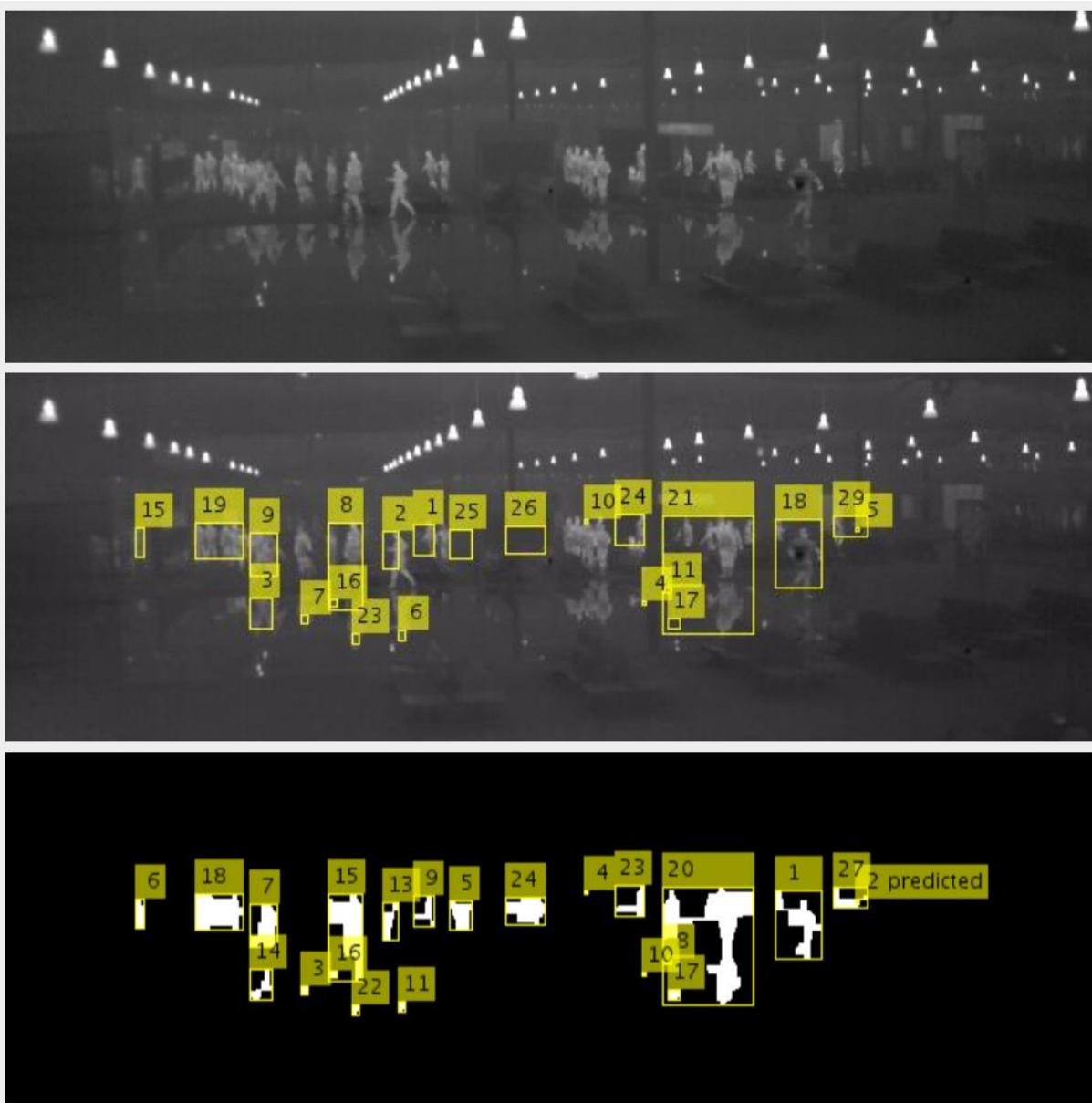
1. frames στα οποία εμφανίζεται μεγάλο πλήθος (υπάρχουν πολλοί άνθρωποι).
2. frames στα οποία φεύγουν οι περισσότεροι άνθρωποι από το πεδίο καταγραφής της κάμερας.
3. frames στα οποία απεικονίζεται μικρός αριθμός ατόμων, χωρίς να υπάρχει αλληλοεπικάλυψη μεταξύ τους.

Η σειρά των παραπάνω κατηγοριών επιλέχτηκε με τρόπο που να συμβαδίζει χρονικά με τα γεγονότα που εκτυλίσσονται στο βίντεο. Το πρώτο κομμάτι του βίντεο αποτελείται από frames στα οποία παρατηρείται υψηλή συγκέντρωση ατόμων. Στη συνέχεια ένα μικρό μέρος του βίντεο αποτελείται από frames που παρουσιάζουν το πλήθος να συνωστιάζεται αρχικά στις δύο εξόδους και στη συνέχεια να εξαφανίζονται όλοι οι άνθρωποι από το πλάνο. Ακολουθεί ένα μέρος που δεν υπάρχει καθόλου ανθρώπινη δραστηριότητα προς ανίχνευση και τέλος, στα τελευταία frames κάνουν την εμφάνιση του δύο άτομα στο πλάνο (ένα άτομο στην περίπτωση του αρχείου AIA_TH_04_1.avi).

❖ Σχολιασμός πρώτης κατηγορίας frames.

Στην κατηγορία αυτή όπως προαναφέρθηκε ανήκουν τα frames (εικόνα 5.3.3.1.1) που παρουσιάζουν υψηλή συγκέντρωση ατόμων. Ο αλγόριθμος που χρησιμοποιήθηκε είναι βασισμένος κυρίως στην κίνηση (motion-based) των αντικειμένων και γι' αυτό το λόγο τα αποτελέσματα στο κομμάτι της εξαγωγής της δυαδικής μάσκας, που στην οποία με άσπρο χρώμα εμφανίζονται τα κινούμενα αντικείμενα (στην περίπτωση μας οι άνθρωποι), ήταν αρκετά ικανοποιητικά. Το πρόβλημα που αντιμετώπισε ο αλγόριθμος ήταν στο κομμάτι του διαχωρισμού των tracks. Αυτό οφείλεται σε διάφορους παράγοντες. Ο κυριότερος εξ αυτών είναι το γεγονός της αλληλοεπικάλυψης των ανθρώπων με αποτέλεσμα ο αλγόριθμος να θεωρεί σαν ένα track την ανίχνευση μιας ομάδας ανθρώπων. Ένα άλλο πρόβλημα αποτελεί το δάπεδο του χώρου, στο οποίο παρουσιάζεται η ανάκλαση των φιγούρων των ατόμων.

Ο συγκεκριμένος αλγόριθμος αποθηκεύει και αυτές τις σκιές σαν tracks, γεγονός που δημιουργεί πρόβλημα στην περαιτέρω επεξεργασία των δεδομένων (βλέπε το κάτω μέρος των frames). Τέλος το μικρό μέγεθος των φιγούρων σε συνδυασμό με την αδυναμία ανίχνευσης αλλαγής θέσης (κίνησης) των ατόμων στην περίπτωση που τα άτομα παραμένουν ακίνητα ή κινούνται στην κατεύθυνση του βάθους έχουν σαν αποτέλεσμα να μην καταγράφεται όλη η ανθρώπινη δραστηριότητα που απεικονίζεται στο εκάστοτε frame.



Εικ 5.3.3.1.1 : Αποτελέσματα object detection στο frame #70.

❖ Σχολιασμός δεύτερης κατηγορίας frames.

Στην κατηγορία αυτή έχουμε τα frames στα οποία απεικονίζονται τα τελευταία στάδια της εκκένωσης του κτιρίου (εικόνα 5.3.3.1.2). Πολλοί άνθρωποι δηλαδή πολλά tracks χάνονται απο το πεδίο καταγραφής της κάμερας. Ταυτόχρονα παρουσιάζεται μεγάλος συνωστισμός στις εξόδους με αποτέλεσμα να γίνεται η καταγραφή πολλών ατόμων σαν ένα track λόγω αλληλοεπικάλυψης. Οι άνθρωποι που βρίσκονταν ανάμεσα στις 2 εξόδους βλέπουμε να μην εντοπίζονται ή να μην κατηγοριοποιούνται σαν reliable tracks. Αυτό οφείλεται στο γεγονός ότι μπορεί να ήταν ακίνητοι ή στο ότι η κίνηση τους γινόταν κυρίως στον άξονα του βάθους. Αξίζει να σημειωθεί ότι σ' αυτά τα frames παρατηρείται και η αδυναμία του Kalman Filter να προβλέψει σωστά την κίνηση των tracks, με αποτέλεσμα να βλέπουμε αυτές τις περιοχές με ετικέτα "predicted" που αναφέρονται στην εκτίμηση της τοποθεσίας των tracks μέσω του Kalman Filter. Αυτό δημιουργεί προβλήματα στο κομμάτι του tracking καθώς δεν αποθηκεύεται σωστά η καινούρια τοποθεσία κάθε track.

❖ Σχολιασμός τρίτης κατηγορίας frames.

Εδώ ανήκουν τα frames στα οποία παρουσιάζεται μικρή ανθρώπινη δραστηριότητα, δηλαδή απεικονίζουν λίγους ανθρώπους. Στην περίπτωση των θερμικών βίντεο, τέτοια είναι τα τελευταία frames των βίντεο όπου απεικονίζονται ένας ή δύο άνθρωποι στο σκηνικό. Σε τέτοιες ιδανικές περιπτώσεις ο αλγόριθμος δουλεύει άριστα και μπορεί κάποιος να το διαπιστώσει πολύ εύκολα κοιτάζοντας την εικόνα 5.3.3.1.3.



Εικ 5.3.3.1.2 : Αποτελέσματα object detection στο frame #166.



Εικ 5.3.3.1.3 : Αποτελέσματα object detection στο frame #393.

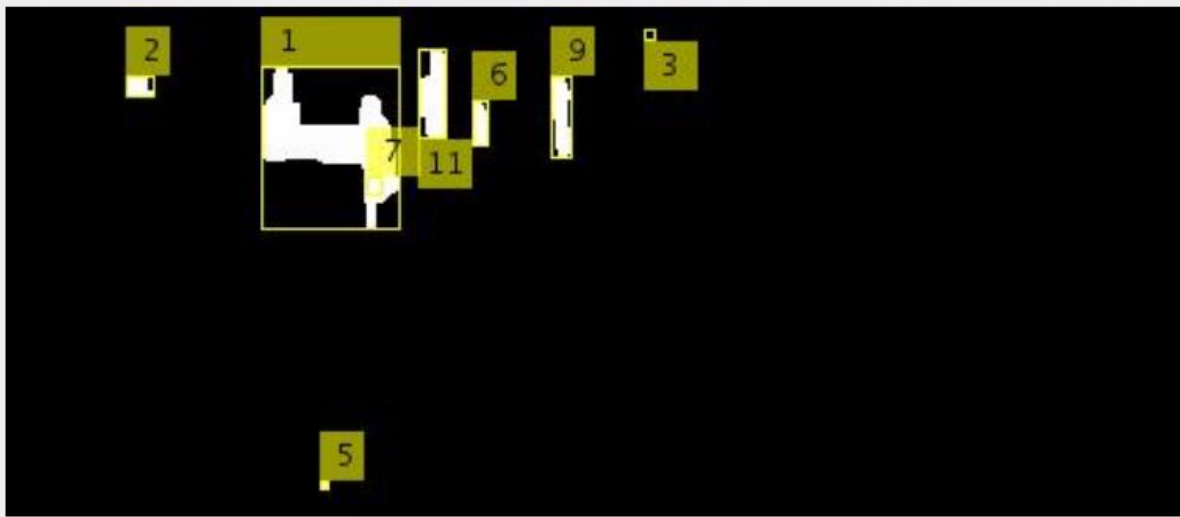
5.3.3.2 Σχολιασμός αποτελεσμάτων εφαρμογής του αλγορίθμου στα οπτικά δεδομένα.

Η εφαρμογή του αλγορίθμου στα οπτικά δεδομένα, δηλαδή στο βίντεο "ΑΙΑ_HS_04_1.avi" που καταγράφηκε από την υπερφασματική κάμερα (στα κανάλια R,G,B), δεν έδωσε ικανοποιητικά αποτελέσματα. Τα περισσότερα προβλήματα σχετίζονται με τη φύση των δεδομένων. Όπως μπορεί να διαπιστωθεί εύκολα από μια απλή παρατήρηση των ανεπεξέργαστων frames, υπάρχει πολύ μικρή διαφοροποίηση στους τόνους του γκρι στα pixels της εικόνας. Αυτό το πρόβλημα είναι εντονότερο στην περιοχή μπροστά από τα καθίσματα (το κάτω μέρος των frames), η οποία είναι πολύ σκοτεινή με αποτέλεσμα ο αλγόριθμος να μην εντοπίζει εύκολα τις μεταβολές στις τιμές των pixels αφού σχεδόν όλη η περιοχή, με ή χωρίς ανθρώπινη παρουσία, είναι πολύ σκούρα. Το αντίθετο συμβαίνει στην περιοχή στο βάθος της απεικόνισης όπου υπάρχουν παράθυρα. Εκεί είναι πολύ έντονες οι διαφοροποιήσεις στον τόνο του γκρι που ο αλγόριθμος ανιχνεύει και αποθηκεύει ακόμα και τις ανακλάσεις στο γυαλί (φαίνεται εντονότερα στην εικόνα 5.3.3.2.3). Τέλος υπάρχει και το πρόβλημα της αλληλοεπικάλυψης των φιγούρων όπου ο αλγόριθμος καταγράφει ως μια ανίχνευση - ένα track μία ομάδα ανθρώπων.

Αν και η δυαδική μάσκα εξάχθηκε με ικανοποιητική ακρίβεια, δεδομένων των συνθηκών, εντούτοις ο διαχωρισμός της σε tracks δεν ήταν ο αναμενόμενος. Όσο μεγάλωνε ο αριθμός των ατόμων που εμφανίζονταν στο frame τόσο μειωνόταν η αποδοτικότητα του αλγορίθμου.



Εικ 5.3.3.2.1 : Αποτελέσματα object detection στο frame #36.



Εικ 5.3.3.2.2 : Αποτελέσματα object detection στο frame #54.



Εικ 5.3.3.2.3 : Αποτελέσματα object detection στο frame #140.

5.3.3.1 Σύγκριση αποτελεσμάτων μεταξύ θερμικών και οπτικών δεδομένων

Με την εφαρμογή του ίδιου αλγόριθμου και στα δύο είδη δεδομένων τα αποτελέσματα προέκυψαν όπως περιγράφηκε πιο πάνω. Το "δυνατό" κομμάτι του αλγορίθμου είναι ο εντοπισμός της ανθρώπινης δραστηριότητας σαν σύνολο, γεγονός που το επιβεβαιώνουν οι δυαδικές μάσκες τόσο των θερμικών όσο και των οπτικών δεδομένων.

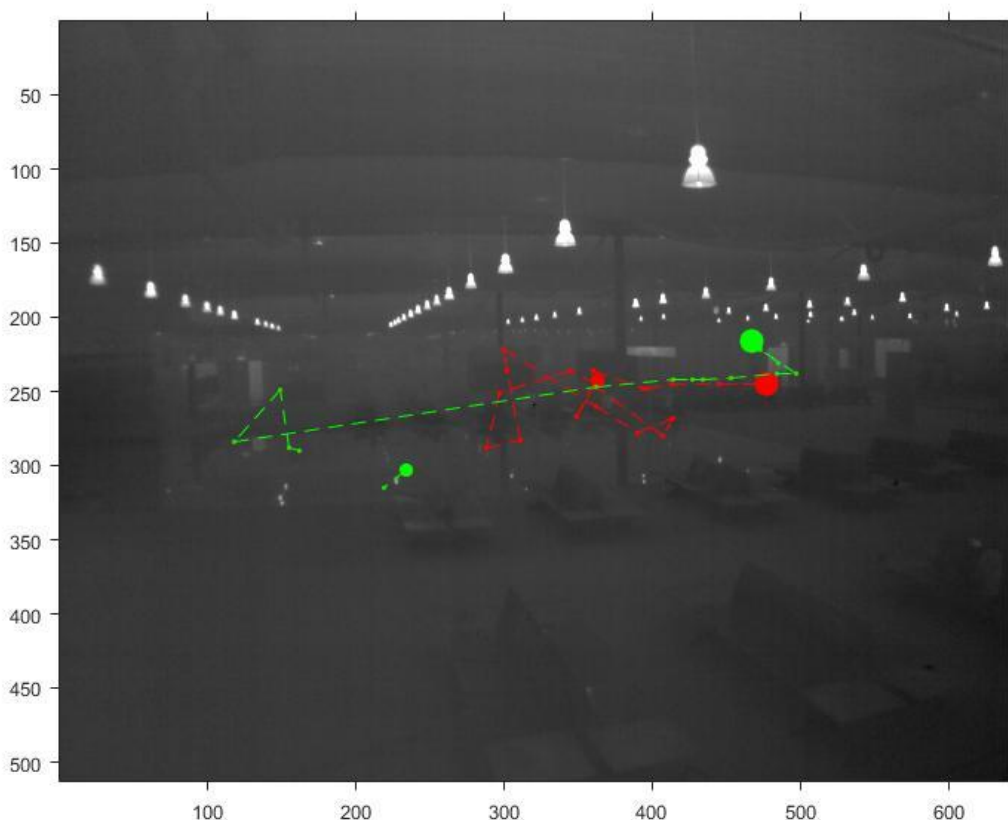
Στην περίπτωση των θερμικών δεδομένων, παρατηρείται μια υπεροχή στην ποιότητα των αποτελεσμάτων, η οποία οφείλεται κυρίως στην ιδιότητα της θερμικής απεικόνισης να παρουσιάζει το ανθρώπινο σώμα σε πιο ανοικτό τόνο (λόγω της θερμοκρασίας του) σε σύγκριση με το παρασκήνιο που το περιβάλλει. Δηλαδή στην περίπτωση των θερμικών δεδομένων ήταν πολύ πιο εύκολη και αποδοτική η διαδικασία της αφαίρεσης παρασκηνίου (background segmentation).

Το γενικό πρόβλημα του αλγορίθμου ήταν η αδυναμία σωστού διαχωρισμού των ανιχνεύσεων σε ανεξάρτητα tracks. Το πρόβλημα αυτό οφείλεται σε ένα τεράστιο βαθμό στο γεγονός της αλληλοεπικάλυψης των ατόμων, στο μικρό μέγεθος των φιγούρων και στην κίνηση τους στον άξονα του βάθους της εικόνας. Ο λανθασμένος διαχωρισμός των ανιχνεύσεων σε ανεξάρτητα tracks οδήγησε σε μη ικανοποιητικά αποτελέσματα στο κομμάτι της παρακολούθησης. Κατά την εξέλιξη των frames, πολλές ανιχνεύσεις αποθηκεύονταν κάτω από διαφορετικά tracks, αρκετές ανιχνεύσεις ομαδοποιούνταν και καταγράφονταν κάτω από ένα track και μερικές ανιχνεύσεις δεν καταγράφονταν καν σαν tracks καθώς δεν ικανοποιούσαν το κριτήριο των οκτώ συνεχόμενων εμφανίσεων σε frames.

Αποτελέσματα της διαδικασίας παρακολούθησης φαίνονται στις εικόνες παρακάτω εικόνες.

Στην εικόνα 5.3.3.1.1 βλέπουμε την πορεία που είχαν δύο tracks κατά τη διάρκεια καταγραφής του θερμικού βίντεο. Με πράσινο χρώμα απεικονίζεται το track #4 και με κόκκινο το track #7. Ο μικρός κύκλος σηματοδοτεί την τοποθεσία της πρώτης καταγραφής του track ενώ ο μεγάλος την τοποθεσία εξόδου, δηλαδή την τοποθεσία της τελευταίας καταγραφής του track. Η

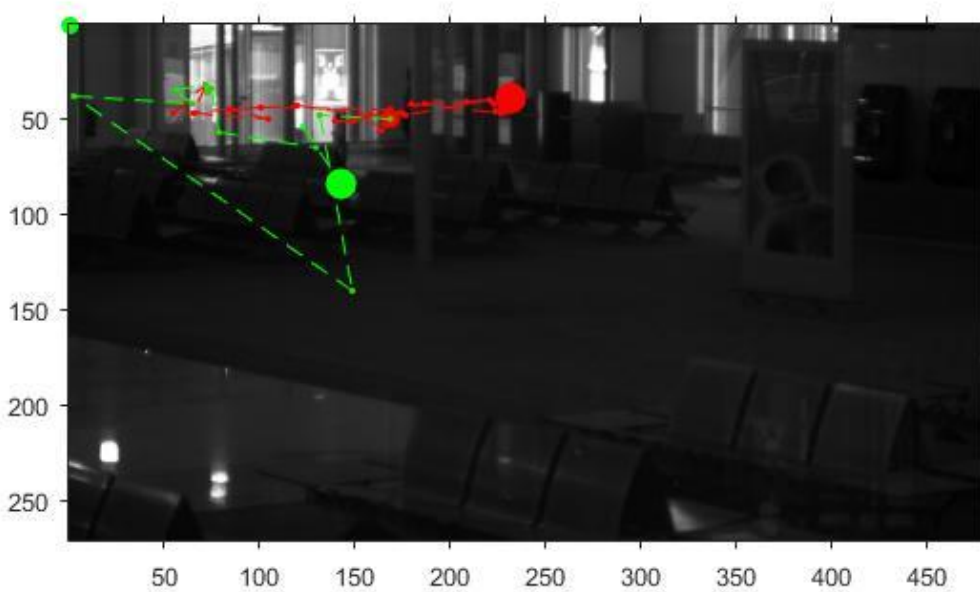
πορεία που ακολουθεί το track #4 (πράσινο χρώμα) έχει καταγραφεί με ικανοποιητική ακρίβεια, αν εξαιρεθούν μερικά frames λίγο μετά την πρώτη εμφάνιση του. Το track #7 είναι χαρακτηριστικό παράδειγμα της καταγραφής ανίχνευσης που δεν ανήκει στο ίδιο track. Η περιοχή στο κέντρο της εικόνας ήταν η πιο πολυσύχναστη, και αποτελεί την πιο προβληματική περιοχή επεξεργασίας για τον αλγόριθμο. Το θετικό στοιχείο στην περίπτωση των θερμικών δεδομένων είναι το γεγονός ότι ο αλγόριθμος κατέγραφε σχεδόν σε όλες τις περιπτώσεις ως σημείο εξόδου των tracks, σημείο κοντά στις εξόδους.



Εικ 5.3.3.1.1 : Παρακολούθηση (tracking) στο βίντεο θερμικών δεδομένων

Στην περίπτωση των οπτικών δεδομένων, για λόγους που αναπτύχτηκαν πιο πάνω, ο διαχωρισμός των ανιχνεύσεων και η καταγραφή τους ως ανεξάρτητα tracks δεν ήταν ο αναμενόμενος. Έτσι το αποτέλεσμα της διαδικασίας παρακολούθησης των αντικειμένων, σε επίπεδο παρακολούθησης ανεξάρτητων tracks (tracking individual tracks) ήταν λανθασμένο και μη αντιπροσωπευτικό. Στην εικόνα 5.3.3.1.2 βλέπουμε την πορεία που κατέγραψε ο αλγόριθμος για τα track #1 (κόκκινο χρώμα) και track #9

(πράσινο χρώμα). Επειδή όλη η "δράση" λάμβανε χώρα στο πάνω αριστερά κομμάτι της εικόνα, βλέπουμε ότι εκεί υπάρχει σωρεία ανιχνεύσεων, αλλά η πορεία τους είναι μη φυσιολογική. Ο πιθανότερος λόγος είναι ότι κάτω απο ίδιο track καταγράφονταν ανιχνεύσεις διαφορετικών ατόμων. Σε περιοχές με παρουσία πολλών αλληλεπικαλυπτόμενων κινούμενων αντικειμένων η χρήση του Kalman Filter, που είναι το κύριο εργαλείο στο κομμάτι της παρακολούθησης, δεν δίνει ικανοποιητικά αποτελέσματα.



Εικ 5.3.3.1.2 : Παρακολούθηση (tracking) στο βίντεο οπτικών δεδομένων

Σαν τελικό συμπέρασμα, τα αποτελέσματα σε επίπεδο ατομικής ανίχνευσης αντικειμένων ήταν ικανοποιητικά και μπορούν να χρησιμοποιηθούν για περαιτέρω ανάλυση και επεξεργασία. Στην επόμενη ενότητα γίνεται η επεξεργασία των ανιχνεύσεων, η κατηγοριοποίηση τους σε ομάδες και η μελέτη αυτών το ομάδων αναφορικά με την κίνηση τους στον χώρο και το χρόνο.

6

Συσταδοποίηση ανιχνεύσεων και παρακολούθηση των συστάδων που δημιουργούνται

Στα συστήματα δυναμικής εκκένωσης κτιρίων το ενδιαφέρον επικεντρώνεται στη μελέτη της συμπεριφοράς και της κίνησης των ανθρώπων σε ομάδες, και όχι σαν ξεχωριστές οντότητες. Τα αποτελέσματα στο κομμάτι της ανίχνευσης αντικειμένων (ανθρώπων) που προέκυψαν από την εφαρμογή του αλγορίθμου (στη προηγούμενη ενότητα), δύναται να χρησιμοποιηθούν σαν δεδομένα για περαιτέρω επεξεργασία. Το κεφάλαιο αυτό ασχολείται με την ταξινόμηση των ανιχνεύσεων κάθε frame σε συστάδες με σκοπό τη μελέτη της συμπεριφοράς των ομάδων αυτών.

Η κύρια ιδέα του αλγορίθμου k -μέσων είναι η ταξινόμηση των παρατηρήσεων σε k συστάδες με κριτήριο την κοντινότερη απόσταση μεταξύ κέντρου συστάδας και παρατήρησης. Πρόκειται για μια επαναληπτική διαδικασία 2

βημάτων, ανάθεσης και εκσυγχρονισμού, η οποία τερματίζεται όταν υπάρξει σύγκλιση στα αποτελέσματα της διαδικασίας.

Στο πλαίσιο της εργασίας εφαρμόστηκε ο αλγόριθμος k-μέσων στα αρχεία των ανιχνεύσεων που προέκυψαν από την εφαρμογή του αλγορίθμου 2 (Multi Object Tracking) της ενότητας 5.3 στα θερμικά αρχεία δεδομένων (βίντεο). Ο λόγος που δεν παρουσιάζονται τα αποτελέσματα της εφαρμογής του αλγορίθμου στα οπτικά δεδομένα είναι ότι τα αποτελέσματα δεν ήταν αξιόπιστα και δεν παρουσιάζουν κάποια χρησιμότητα. Αυτό οφειλόταν σε λόγους που δεν αφορούσαν την σωστή λειτουργία των αλγορίθμων που χρησιμοποιήθηκαν, αλλά στην ίδια φύση των δεδομένων. Ο αριθμός των ανιχνεύσεων σε κάθε frame είναι πολύ μικρός και δεν δικαιολογεί την χρήση αλγορίθμου ταξινόμησης. Επίσης, όλη η ανθρώπινη δραστηριότητα στο οπτικό βίντεο λαμβάνει χώρα σε ένα πολύ μικρό κομμάτι του πεδίου καταγραφής (πάνω αριστερά) με αποτέλεσμα όλες οι ανιχνεύσεις να συνωστιζονται σε αυτό το χώρο. Η εφαρμογή ενός αλγορίθμου συσταδοποίησης σε τέτοιας φύσεως ανιχνεύσεις δεν έχει νόημα καθώς το αποτέλεσμα δεν θα οδηγήσει σε χρήσιμα συμπεράσματα.

Σχετικά με την επεξεργασία των θερμικών δεδομένων, σκοπός ήταν η εξαγωγή εικόνων που να παρουσιάζουν την κατηγοριοποίηση των ανιχνεύσεων κάθε frame σε συστάδες, και ακολούθως η παρακολούθηση αυτών των συστάδων στην εξέλιξη του χρόνου.

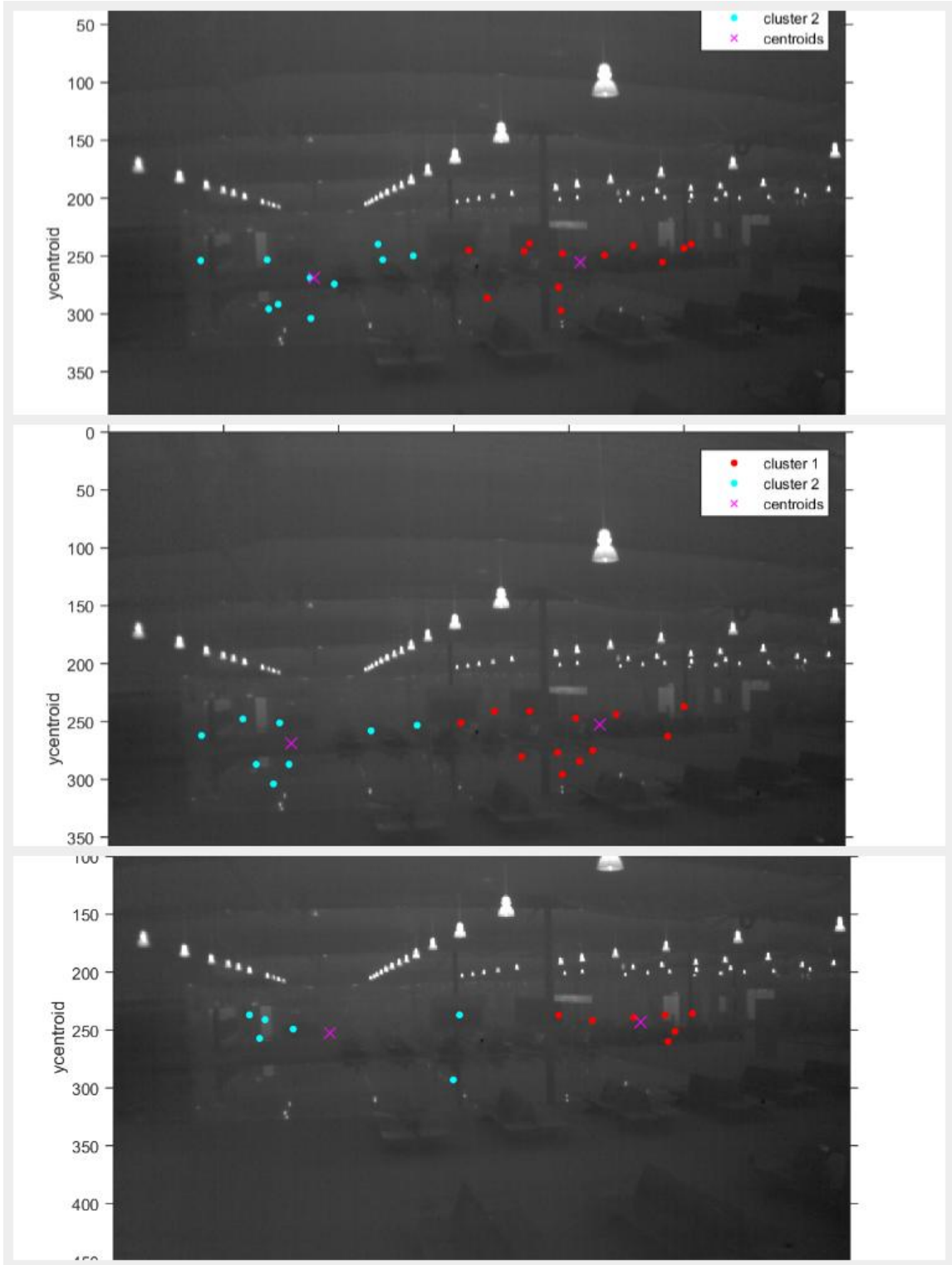
6.1 Εφαρμογή αλγορίθμου στα θερμικά δεδομένα

Τα θερμικά δεδομένα εισόδου στον αλγόριθμο k-μέσων αποτελούν τα centroids των ανιχνεύσεων κάθε frame του θερμικού αρχείου βίντεο. Είναι δηλαδή το αποτέλεσμα της διαδικασίας ανίχνευσης της προηγούμενης ενότητας το οποίο ο αλγόριθμος αποθήκευε στον πίνακα - αρχείο dataset.

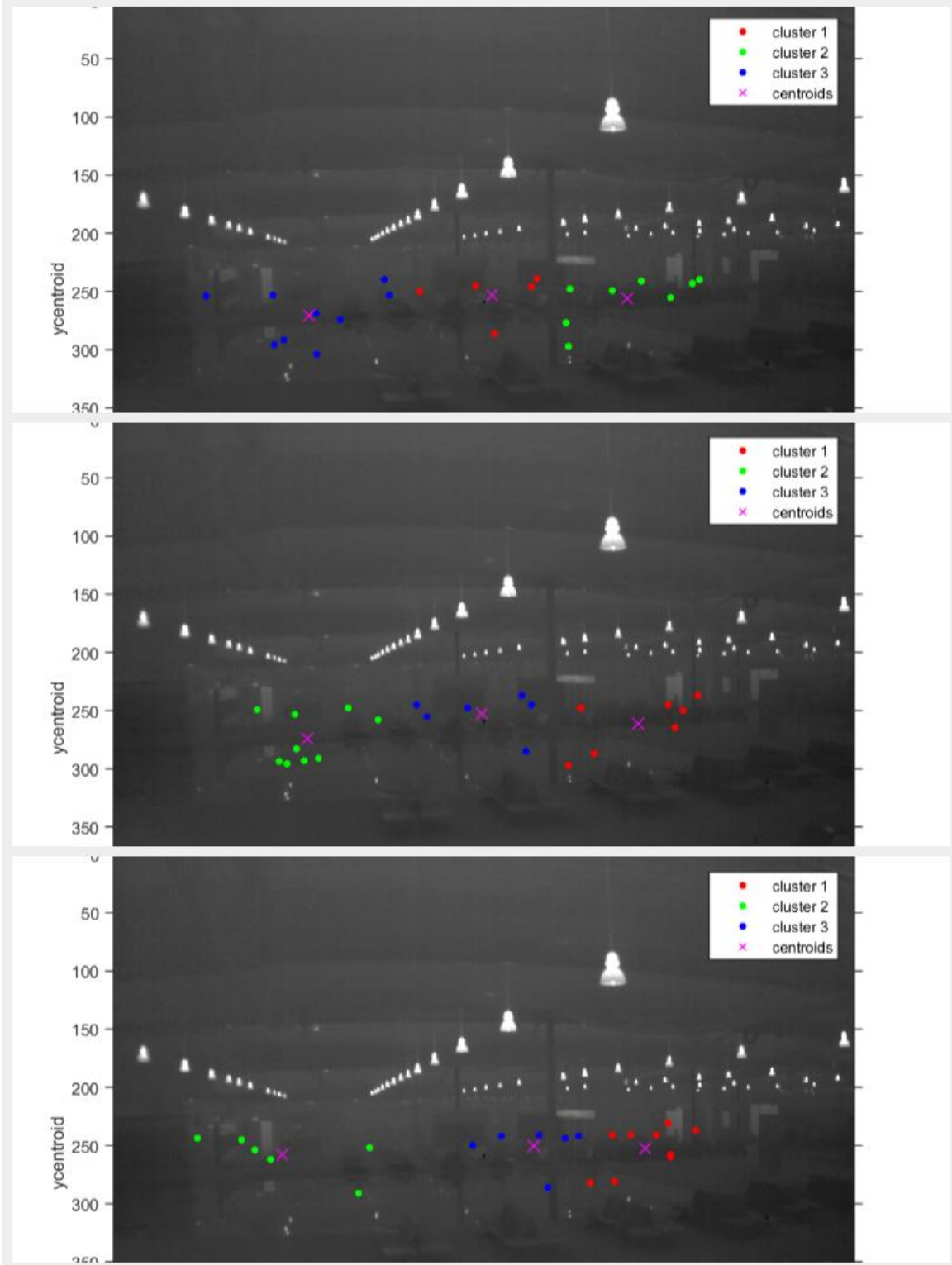
Στην περίπτωση των θερμικών δεδομένων, εφαρμόστηκε ο αλγόριθμος για δύο διαφορετικές τιμές k (αριθμός συστάδων). Ο αριθμός των ανιχνεύσεων καθώς και η κατανομή τους στο χώρο, δικαιολογούσε τον διαχωρισμό τους σε δύο ή τρεις ομάδες. Το αποτέλεσμα του αλγορίθμου για k =2 διαφαίνεται στην Εικ. 6.1.1 (για τα frames 80,110,150) και για k=3 στην Εικ 6.1.2 (για τα frames 80, 100, 120).

Για $k=2$ δημιουργούνται δυο συστάδες που τα κέντρα τους, κατά την εξέλιξη των frames μετατοπίζονται όλο και πιο κοντά στις δύο εξόδους. Με άλλα λόγια το αποτέλεσμα της παρακολούθησης (tracking) των ανιχνεύσεων σε επίπεδο δυο ομάδων παρουσιάζει μεγάλη ακρίβεια και είναι σύμφωνο με την πραγματική εξέλιξη της εκκένωσης στο αρχικό βίντεο. Στην εικόνα 6.1.3 παρουσιάζονται τα κέντρα των συστάδων ανά 5 frames για όλη την διάρκεια του βίντεο. Εύκολα μπορεί να παρατηρήσει κανείς ότι τα κέντρα αυτά ομαδοποιούνται σε 2 περιοχές, περιοχές που είναι κοντά στις δυο εξόδους. Ερμηνεύοντας την εικόνα αντιλαμβάνεται κανείς ότι κοντά στην δεξιά έξοδο που υπάρχουν πιο πολλά κέντρα συστάδων, είτε έχει καταγραφεί περισσότερη ανθρώπινη δραστηριότητα ή ότι σε αυτή την περιοχή καταγράφηκε ανθρώπινη δραστηριότητα σε περισσότερα frames σε σύγκριση με την περιοχή κοντά στην αριστερή έξοδο.

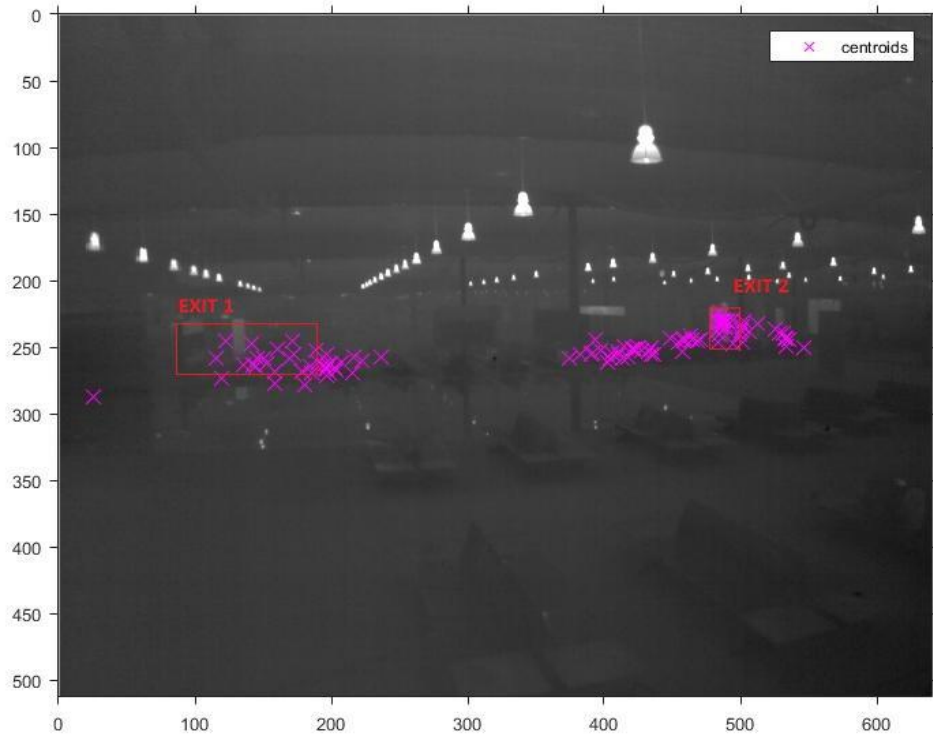
Για $k=3$ δημιουργούνται τρεις συστάδες όπως φαίνονται στην εικόνα 6.1.4 που παρουσιάζει τα κέντρα των συστάδων ανά 5 frames για όλη τη διάρκεια του βίντεο. Τα κέντρα δυο εξ αυτών όπως ήταν αναμενόμενο ομαδοποιούνται κοντά στις δύο εξόδους. Η τρίτη συστάδα αποτελείται από τα κέντρα των ανιχνεύσεων που βρίσκονταν στο κέντρο του πεδίου. Η ανάλυση της τρίτης συστάδας επιδεικνύει ιδιαίτερο ενδιαφέρον καθώς αποτελείται από τις ανιχνεύσεις των αναποφάσιστων, ως προς το πια έξοδο θα χρησιμοποιήσουν, ατόμων. Η μελέτη αυτής της ομάδας και η πρόβλεψη της κίνησης της είναι ένα από τα βασικά ζητούμενα των δυναμικών συστημάτων εκκένωσης κτιρίων.



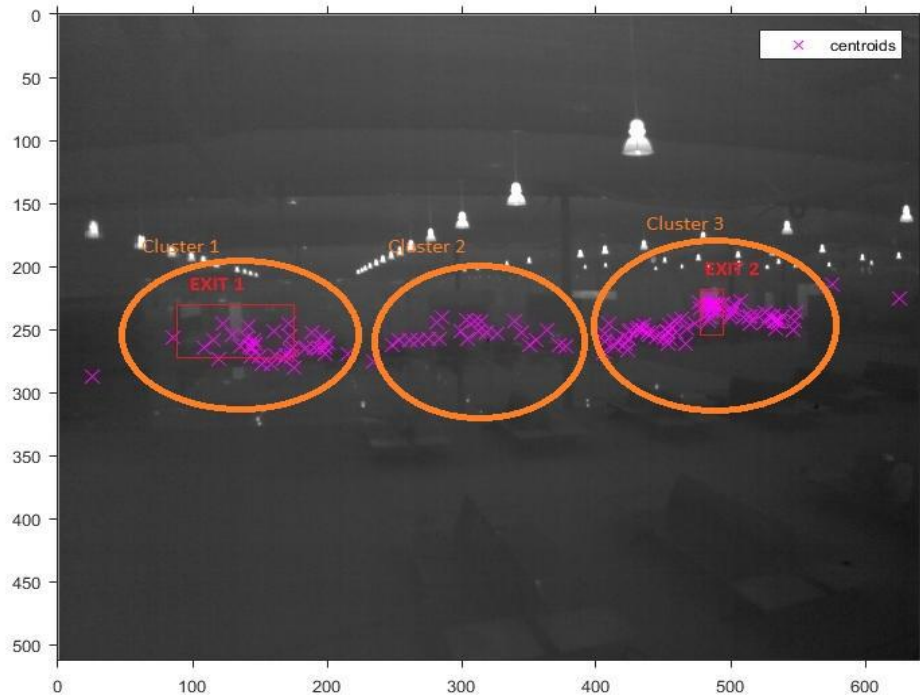
Εικ 6.1.1 : Εφαρμογή αλγορίθμου k-μέσων στα θερμικά δεδομένα για k=2



Εικ 6.1.2 : Εφαρμογή αλγορίθμου k-μέσων στα θερμικά δεδομένα για k=3



Εικ 6.1.3 : Παρουσίαση κέντρων συστάδων για $k=2$



Εικ 6.1.3 : Παρουσίαση κέντρων συστάδων για $k=3$

7

Συμπεράσματα και μελλοντικές κατευθύνσεις

Στο κεφάλαιο αυτό παρουσιάζονται τα συμπεράσματα, για τα επιμέρους τμήματα, αλλά και για το σύνολο (της προσέγγισης), που προέκυψαν από την εφαρμογή της παρούσας προσέγγισης ως λύση στο πρόβλημα της ανίχνευσης και παρακολούθησης ανθρώπινης δραστηριότητας σε οπτικά και θερμικά αρχεία βίντεο, μέσω τεχνικών Όρασης Υπολογιστών. Στη συνέχεια γίνεται μια αναφορά στις κατευθύνσεις στις οποίες μπορεί να κινηθεί η παρούσα εργασία για να βελτιστοποιηθεί.

7.1 Συμπεράσματα

Τα τελευταία χρόνια γίνεται προσπάθεια ανάπτυξης δυναμικών συστημάτων τα οποία θα διαχειρίζονται καταστάσεις έκτακτης αναγκαστικής εκκένωσης κτηρίων. Για την ανάπτυξη τέτοιων συστημάτων απαιτείται η καταγραφή και η κατανόηση της παρούσας κατάστασης από το πρόγραμμα.

Η παρούσα διπλωματική ασχολήθηκε με τη λύση του προβλήματος της ανίχνευσης και παρακολούθησης της ανθρώπινης δραστηριότητας τόσο σε επίπεδο ατόμου, όσο και σε επίπεδο ομάδων, μέσα από την ανάλυση οπτικών και θερμικών βίντεο. Για την επίλυση του προβλήματος αναλύθηκαν και χρησιμοποιήθηκαν τεχνικές και προσεγγίσεις που ανήκουν στον θεματικό τομέα της Όρασης Υπολογιστών.

Για την αντιμετώπιση του πρώτου σταδίου, της ανίχνευσης της ανθρώπινης δραστηριότητας (human detection) αναπτύχθηκε ένας αλγόριθμος βασισμένος στη μέθοδο καθορισμού κατωφλίου Otsu (thresholding method). Τα αποτελέσματα της εφαρμογής του αλγορίθμου, τόσο στα οπτικά όσο και στα θερμικά δεδομένα, δεν ήταν ικανοποιητικά. Έτσι δημιουργήθηκε η ανάγκη αναζήτησης καινούριας προσέγγισης για τη λύση του προβλήματος.

Τη δεύτερη προσέγγιση αποτέλεσε ο αλγόριθμος 2 (Multi Object Tracking) της ενότητας 5.3. Ο αλγόριθμος είναι βασισμένος στην ανίχνευση των κινούμενων αντικειμένων, δηλαδή είναι ένας αλγόριθμος motion-based. Για την εξαγωγή των κινούμενων αντικειμένων (προσκήνιο) από το ακίνητο παρασκήνιο, δηλαδή για την διαδικασία της αφαίρεσης παρασκηνίου, χρησιμοποιεί μίξη Γκαουσιανών μοντέλων, και εφαρμόζει στο κάθε frame ανάλυση blob για τον εντοπισμό ομάδων pixels που πιθανόν να αποτελούν αντικείμενα. Κάθε αντικείμενο που ανιχνεύεται, καταγράφεται στον αλγόριθμο σαν track. Στη συνέχεια με τη χρήση του φίλτρου Kalman δημιουργείται μια χρονοσειρά με τη θέση που είχε το κάθε track σε κάθε frame του βίντεο.

Τα αποτελέσματα της εφαρμογής του αλγορίθμου Multi Object Tracking στο κομμάτι της ανίχνευσης ήταν ικανοποιητικά σε μεγάλο βαθμό και στους δύο τύπους δεδομένων (οπτικά και θερμικά). Τα προβλήματα που παρουσίασε το κομμάτι της ανίχνευσης οφείλονται κυρίως στη φύση των δεδομένων. Τα κύρια προβλήματα ήταν :

- Οι σκιές και οι ανακλάσεις των ανθρώπινων φιγούρων σε πάτωμα και τζάμια καταγράφονταν από τον αλγόριθμο ως tracks.
- Στην περίπτωση των οπτικών δεδομένων στο κάτω μισό τμήμα της απεικόνισης δεν υπήρχαν διαφοροποιήσεις στους τόνους μεταξύ ανθρώπου και παρασκηνίου. Πρόβλημα που προκύπτει λόγω ανεπαρκούς φωτισμού στην περιοχή αυτή.

- Το συχνό φαινόμενο της αλληλοεπικάλυψης των ανθρώπων είχε ως αποτέλεσμα την καταγραφή μιας ομάδας ανθρώπων κάτω από ένα μόνο track.
- Το μικρό μέγεθος απεικόνισης των φιγούρων σε συνδυασμό με τυχόν κίνηση του ανθρώπου προς τον άξονα του βάθους, είχε ως αποτέλεσμα τη μη αναγνώριση του από τον αλγόριθμο.

Τα παραπάνω προβλήματα στο κομμάτι της ανίχνευσης είχαν αλυσιδωτές συνέπειες στο κομμάτι της παρακολούθησης. Επιπρόσθετο εμπόδιο για τη σωστή παρακολούθηση των ανιχνεύσεων στον χρόνο, αποτέλεσε το γεγονός ότι στις απεικονίσεις υπήρχε μεγάλος αριθμός ατόμων (άρα μεγάλος αριθμός ανιχνεύσεων) συγκεντρωμένος σε μικρό χώρο. Όταν ο αριθμός των ανιχνεύσεων σε μια ακολουθία εικόνων είναι μεγάλος, τότε η χρήση του φίλτρου Kalman που είναι το κύριο "εργαλείο" χρήσης στο κομμάτι της παρακολούθησης, αδυνατεί να αντιμετωπίσει με αποτελεσματικότητα το πρόβλημα της σωστής ανάθεσης ανίχνευσης σε track.

Τα αποτελέσματα της διαδικασίας παρακολούθησης της κίνησης των ανεξάρτητων tracks κρίθηκαν αναξιόπιστα και για τους δύο τύπους δεδομένων (οπτικών και θερμικών). Γι' αυτό το λόγο αποφασίστηκε όπως μελετηθεί το κομμάτι της παρακολούθησης σε επίπεδο ομάδων στις οποίες χωρίζεται το πλήθος κατά τη διάρκεια εκκένωσης του κτιρίου.

Για τον διαχωρισμό του πλήθους σε ομάδες, επιστρατεύτηκε ο αλγόριθμος συσταδοποίησης k-μέσων. Αρχικά έγινε η εφαρμογή του αλγορίθμου σε όλες τις ανιχνεύσεις όλων των frames του βίντεο και αποθηκεύτηκαν τα κέντρα (centroids) των συστάδων, που προέκυψαν σε κάθε frame, σε ένα πίνακα. Στη συνέχεια με τη χρήση αυτού του πίνακα προέκυψε η λύση του προβλήματος της παρακολούθησης της ανθρώπινης δραστηριότητας σε επίπεδο ομάδων. Τα αποτελέσματα ήταν αρκετά ικανοποιητικά καθώς η ερμηνεία που προκύπτει από την ανάγνωση τους βρίσκεται πολύ κοντά με την πραγματική εξέλιξη των γεγονότων στο βίντεο.

Αξιίζει να σημειωθεί ότι τα αποτελέσματα εφαρμογής του αλγορίθμου k-μέσων, αν και εφαρμόστηκε και στους δύο τύπους δεδομένων, επιλέχτηκε να παρουσιαστούν στην παρούσα εργασία μόνο για την περίπτωση των θερμικών δεδομένων. Γενικά ο αλγόριθμος k-μέσων για να έχει κάποια χρήσιμη ερμηνευτική σημασία πρέπει οι ανιχνεύσεις να είναι κατανομημένες διάσπαρτα

στο χώρο και να παρουσιάζουν κάποια συγκέντρωση κοντά σε περισσότερα από δύο σημεία. Στην περίπτωση των θερμικών δεδομένων, τέτοια σημεία ήταν οι δύο έξοδοι.

Αν γίνει σύγκριση της αποδοτικότητας της προσέγγισης μεταξύ των οπτικών και των θερμικών δεδομένων, εύκολα διαπιστώνεται ότι τα καλύτερα αποτελέσματα παρουσίασε η εφαρμογή στα θερμικά δεδομένα. Ο κύριος λόγος είναι η ιδιότητα της θερμικής απεικόνισης να παρουσιάζει το ανθρώπινο σώμα με πιο ανοικτό τόνο από το παρασκήνιο του επειδή έχει υψηλότερη θερμοκρασία. Έτσι γίνεται πιο εύκολο ο διαχωρισμός των ανθρώπων από το παρασκήνιο, που αλυσιδωτά οδηγεί σε καλύτερα αποτελέσματα σε όλα τα στάδια της λύσης.

7.2 Μελλοντικές κατευθύνσεις

Το αντικείμενο της παρούσας εργασίας αποτελεί μέρος ενός σύγχρονου πεδίου έρευνας στον Τομέα της Όρασης Υπολογιστών. Προσεγγίζει το ενδιαφέρον μεγάλης μερίδας επιστημόνων με αποτέλεσμα να υπάρχει τεράστια εξέλιξη και πρόοδος ως προς την βελτιστοποίηση των μεθόδων ανίχνευσης και παρακολούθησης της ανθρώπινης δραστηριότητας μέσα από την ανάλυση εικόνων.

Το ποιο υποσχόμενο κομμάτι της παρούσας εργασίας είναι η εφαρμογή της προτεινόμενης προσέγγισης σε θερμικά δεδομένα. Εντούτοις υπάρχει χώρος για σημαντικές βελτιώσεις για να αποτελέσει αξιόπιστη και ικανοποιητική μέθοδο λύσης στο πρόβλημα. Προς την κατεύθυνση αυτή προτείνονται παρακάτω δύο σημεία που θα βελτιώσουν την αποδοτικότητα της παρούσας προσέγγισης και θα της δώσουν καινούρια χρησιμότητα.

- Αρχικά προτείνεται η δημιουργία ενός 3D μοντέλου απεικόνισης της υπάρχουσας κατάστασης το οποίο θα προκύπτει από την λειτουργία δύο θερμικών καμερών. Η μια θα είναι τοποθετημένη στο έδαφος (με τον ίδιο τρόπο που ήταν τοποθετημένη και η θερμική κάμερα στην εργασία μας), ενώ η δεύτερη θα είναι τοποθετημένη στην οροφή. Η πρώτη θα απεικονίζει την πρόσοψη των ανθρώπων και η δεύτερη την κάτοψη. Μέσα από συναρτήσεις αντιστοιχίας πρόσοψης - κάτοψης στο

3D μοντέλο μπορεί να επιτευχθεί ο διαχωρισμός και η καταγραφή των ανιχνεύσεων με περισσότερη ακρίβεια.

- Σαν προέκταση του αλγορίθμου προτείνεται η κατηγοριοποίηση και χαρακτηρισμός της εκάστοτε ανθρώπινης δραστηριότητας μέσω νευρωτικών δικτύων.

Βιβλιογραφία

- [1] : D. M. Gavrila, "The Visual Analysis of Human Movement : A Survey", in *Computer Vision and Image Understanding*, January 1999, Vol. 73, No. 1, pp. 82-98.
- [2] : W.H. Liang Wang, and Tieniu Tan, "Recent developments in human motion analysis", *Pattern Recognition*, 2003, Vol 36, pp. 585-601.
- [3] : J. K. Aggarwal, Q. Cai, W. Liao, and B. Sabata, "Articulated and elastic non-rigid motion: a review", *Proceedings of the 1994 IEEE Workshop on Motion of Non- Rigid and Articulated Objects*, 1994, pp. 2-14.
- [4] : J. K. Aggarwal, Q. Cai, "Human Motion Analysis : A Review", *Computer Vision and Image Understanding*, 1999, vol. 73, pp. 428-440.
- [5] : Thomas B. Moeslund, and Erik Granum, "A Survey of Computer Vision-Based Human Motion Capture", *Computer Vision and Vision and Image Understanding*, 2001, vol. 81, pp. 231-268.
- [6] : Jessica Jun, Lin Wang, and Sameer Singh, "Video analysis of human dynamics - a survey", *Real-Time Imaging*, vol. 9,2003, pp. pp. 321 -346
- [7] : Kandylakis Z, Karantzalos K, Doulamis A, Doulamis N., : "Multiple Object Tracking with Background Estimation in Hyperspectral Video Sequences" 2015, *IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*.
- [8] : Maria Vakalopoulou and Konstantinos Karantzalos, "Automatic Descriptor-Based Co-Registration of Frame Hyperspectral Data," *Remote Sensing*, vol. 6, no. 4, pp. 3409–3426, 2014.
- [9] : R. Bodor, B. Jackson, N. Papanikolopoulos, "Vision-Based Human Tracking and Activity Recognition".
- [10] : C. Stauffer, W.E.L Grimson, "Adaptive background mixture models for real-time tracking", *The Artificial Intelligence Laboratory Massachusetts Institute of Technology Cambridge*.
- [11] : J. M. S. Prewitt and M. L. Mendelsohn, "The analysis of cell images," *Ann. N.Y Acad. Sci.*, vol. 128, pp. 1035-1053, 1966
- [12] : K. Fukunage, *Introduction to Statistical Pattern Recognition*. New York: Academic, 1972, pp. 260-267.

- [13] : Otsu, N., "A Threshold Selection Method from Gray-Level Histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 9, No. 1, 1979, pp. 62-66.
- [14] : K. P. Murphy, *Dynamic Bayesian Networks: Representation, Inference and Learning*. PhD thesis, University of California, Berkeley, 2002.
- [15] : M. S. Arulampalam, S. Maskell, and N. Gordon, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, pp. 174–188, 2002.
- [16] : R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *Transactions of the ASME–Journal of Basic Engineering*, vol. 82, no. Series D, pp. 35–45, 1960
- [17] : Negenborn, R. (2003), 'Robot Localization and Kalman Filters On finding your position in a noisy world' Diploma Thesis, Utrecht University
- [18] : TAU 2 Uncooled Cores (<http://www.flir.eu>)
- [19] : HYPERSPECTRAL MOSAIC SNAPSHOT IMAGER (<http://www2.imec.be>)
- [20] : J. Munkres, "Algorithms for the Assignment and Transportation Problems", *Journal of the Society for Industrial and Applied Mathematics*, Vol.5, No.1, pp32-38, 1957.
- [21] : J. B. MacQueen (1967): "Some Methods for classification and Analysis of Multivariate Observations, *Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability*", Berkeley, University of California Press, 1:281-297
- [22] : Q Lloyd, S. P. (1957). "Least square quantization in PCM". *Bell Telephone Laboratories Paper*. Published in journal much later: Lloyd., S. P. (1982). "Least squares quantization in PCM" *IEEE Transactions on Information Theory* pp 129–137.
- [23] : Hamerly, G.; Elkan, C. (2002). "Alternatives to the k-means algorithm that find better clusterings". *Proceedings of the eleventh international conference on Information and knowledge management (CIKM)*.
- [24] : Arthur, David, and Sergi Vassilvitskii. "K-means++: The Advantages of Careful Seeding." *SODA '07: Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*. 2007, pp. 1027–1035.
- [25] : Seber, G. A. F. *Multivariate Observations*. Hoboken, NJ: John Wiley & Sons, Inc., 1984.
- [26] : Spath, H. *Cluster Dissection and Analysis: Theory, FORTRAN Programs, Examples*. Translated by J. Goldschmidt. New York: Halsted Press, 1985.

- [27] : MacKay, David (2003). "Chapter 20. An Example Inference Task: Clustering". *Information Theory, Inference and Learning Algorithms*. Cambridge University Press. pp. 284–292. ISBN 0-521-64298-1. MR 2012999.
- [28] : Voulodimos, A.S., Doulamis, N.D., Kosmopoulos, D.I., Varvarigou, T.A.: *Improving multicamera activity recognition by employing neural network based readjustment*. *Applied Artificial Intelligence* 26 (2012) 97–118
- [29] : Brutzer, S., Hoferlin, B., Heidemann, G.: *Evaluation of background subtraction techniques for video surveillance*. In: *2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. (2011) 1937–1944
- [30] : El Baf, F., Bouwmans, T., Vachon, B.: *Fuzzy statistical modeling of dynamic backgrounds for moving object detection in infrared videos*. In: *IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2009. CVPR Workshops 2009*. (2009) 60–65
- [31] : Messelodi, S., Modena, C.M., Segata, N., Zanin, M.: *A kalman filter based background updating algorithm robust to sharp illumination changes*. In Roli, F., Vitulano, S., eds.: *Image Analysis and Processing ICIAP 2005*. Number 3617 in *Lecture Notes in Computer Science*. Springer Berlin Heidelberg (2005) 163–170
- [32] : Stauffer, C., Grimson, W.: *Adaptive background mixture models for real-time tracking*. In: *Computer Vision and Pattern Recognition. IEEE Conf. on. Volume 2*. (1999) –252 Vol. 2
- [33] : Zivkovic, Z.: *Improved adaptive gaussian mixture model for background subtraction*. In: *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004. Volume 2*. (2004) 28–31 Vol.2
- [34] : Nathan Hagen and Michael W. Kudenov, "Review of snapshot spectral imaging technologies," *Optical Engineering*, vol. 52, no. 9, pp. 090901–090901, 2013.