



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΑΓΡΟΝΟΜΩΝ ΚΑΙ ΤΟΠΟΓΡΑΦΩΝ ΜΗΧΑΝΙΚΩΝ
ΤΟΜΕΑΣ ΤΟΠΟΓΡΑΦΙΑΣ

Διερεύνηση συστήματος αισθητήρων για την τρισδιάστατη καταγραφή κινούμενων αντικειμένων



Διπλωματική Εργασία
Μαγκαφώση Αντωνία

Αθήνα, Μάρτιος 2018



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΑΓΡΟΝΟΜΩΝ ΚΑΙ ΤΟΠΟΓΡΑΦΩΝ ΜΗΧΑΝΙΚΩΝ
ΤΟΜΕΑΣ ΤΟΠΟΓΡΑΦΙΑΣ

Διερεύνηση συστήματος αισθητήρων για την τρισδιάστατη καταγραφή κινούμενων αντικειμένων

Διπλωματική Εργασία
Μαγκαφώση Αντωνία

Επιβλέπων: Ανδρέας Γεωργόπουλος
Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή:

.....
Γεωργόπουλος Ανδρέας

Καθηγητής Ε.Μ.Π.

.....
Δουλάμης Νικόλαος

Επίκουρος Καθηγητής Ε.Μ.Π.

.....
Δουλάμης Αναστάσιος

Επίκουρος Καθηγητής Ε.Μ.Π.

Αθήνα, Μάρτιος 2018

Ευχαριστίες

Θα ήθελα να ευχαριστήσω ξεχωριστά τους ανθρώπους που συνέβαλαν ο καθένας με τον τρόπο του έτσι ώστε να μπορέσω να φέρω εις πέρας αυτή τη διπλωματική εργασία.

Τον Καθηγητή της Σχολής Αγρονόμων και Τοπογράφων Μηχανικών κ. Γεωργόπουλο Ανδρέα για την καθοδήγηση και τη συνεργασία του.

Τους γονείς μου οι οποίοι στάθηκαν δίπλα όλα αυτά τα χρόνια μέχρι την ολοκλήρωση των σπουδών μου.

Τον φίλο μου Daniele για την πολύτιμη βοήθειά του στην πρακτική εφαρμογή της εργασίας.

Όλους τους φίλους μου για την ψυχολογική υποστήριξη καθ' όλη την περίοδο πραγμάτωσης της εργασίας.

Περίληψη

Η παρούσα διπλωματική εργασία εστιάζει στις τεχνολογίες που σχετίζονται με τα συστήματα καταγραφής της ανθρώπινης κίνησης και πιο συγκεκριμένα στο σύστημα Microsoft Kinect και στον αισθητήρα Leap Motion. Γίνεται ανάλυση και σύγκριση των δύο συσκευών που βασίζονται στις ίδιες αρχές λειτουργίας και εξάγονται συμπεράσματα για την καταλληλότητά τους αναλόγως το καταγραφόμενο αντικείμενο. Παρουσιάζεται η μεθοδολογία καταγραφής της κίνησης. Επιπλέον γίνεται η καταγραφή δεδομένων με τον αισθητήρα Leap Motion. Τα δεδομένα αυτά είναι χειρονομίες, κινήσεις των χεριών. Αναλύεται τόσο ο τρόπος καταγραφής τους όσο και τα ίδια τα εξαγόμενα αποτελέσματα. Μετά τη διερεύνηση ενός τέτοιου συστήματος για την καταγραφή κινούμενων αντικειμένων εξάγονται τα συμπεράσματα και γίνονται προτάσεις για βελτίωση της καταγραφής .

Abstract

This diploma thesis focuses on technologies related to motion capture recording systems, more specifically in the Microsoft Kinect and the Leap Motion sensor. The two devices, that are based on the same principles, are analyzed and compared and their suitability according to the recorded object is concluded. The motion capture methodology is presented. In addition, motion capture data are recorded by the Leap Motion sensor. These data are gestures, hand movements. The structure of the data that are recorded and the exported results are analyzed. After analyzing such a system for recording animated objects, conclusions are drawn and suggestions are made to improve the results of the recordings.

Περιεχόμενα εικόνων

Εικόνα 2.1: Αισθητήρας Kinect.....	21
Εικόνα 2.2: Τεχνικά χαρακτηριστικά του αισθητήρα Kinect.....	25
Εικόνα 2.3: Τεχνικά χαρακτηριστικά στο εσωτερικό τμήμα του αισθητήρα Kinect.....	25
Εικόνα 2.4: Τα βήματα της διαδικασίας προσδιορισμού της θέσης του σώματος	26
Εικόνα 2.5: Το σταθερό μοτίβο αναφοράς της δέσμης laser.....	28
Εικόνα 2.6: Το εσωτερικό τμήμα της συσκευής Leap Motion.....	33
Εικόνα 2.7: Το πεδίο αλληλεπίδρασης της συσκευής Leap Motion.....	35
Εικόνα 2.8: Το σύστημα συντεταγμένων της συσκευής Leap Motion.....	36
Εικόνα 2.9: Ιεραρχία των κλάσεων του API της συσκευής.....	39
Εικόνα 3.1: Γωνία που σχηματίζεται μεταξύ των αξόνων του συστήματος συντεταγμένων και του σημείου καταγραφής.....	46
Εικόνα 4.1: Πίνακας ρυθμίσεων της συσκευής.....	55
Εικόνα 4.2: Πίνακας ελέγχου της κατάστασης της συσκευής.....	55
Εικόνα 4.3: Πληροφορίες για την επίσημη ιστοσελίδα και το λογισμικό.....	56
Εικόνα 4.4: Τμήμα εσωτερικού script.....	58
Εικόνα 4.5: Λογισμικό Bracket.....	59
Εικόνα 4.6: Demo της ιστοσελίδας Leap motion (Leap recorder).....	60
Εικόνα 4.7: Ιεραρχική δομή των δεδομένων της κίνησης.....	63
Εικόνα 4.8: Σκελετική δομή ενός παραδείγματος αρχείου BVH (α) αρχική θέση (β) πρώτο καρέ της κίνησης.....	63
Εικόνα 4.9: Καταγραφή δεδομένων.....	68
Εικόνα 4.10: Τμήμα κώδικα.....	69
Εικόνα 4.11: Τμήμα κώδικα.....	70
Εικόνα 4.12: «Τυπωμένα» στοιχεία καταγραφής.....	71
Εικόνα 4.13: Δομή δεδομένων αρχείου json.....	72
Εικόνα 4.14: Interaction Box.....	73

Περιεχόμενα

Ευχαριστίες.....	5
Περίληψη.....	7
Abstract	9
Περιεχόμενα εικόνων.....	11
1. Εισαγωγή	15
1.1 Στόχος.....	15
1.2 Δομή	17
2 Περιγραφή και σύγκριση συσκευών	19
2.2 Ο αισθητήρας Kinect.....	19
2.1.1 Τεχνικά χαρακτηριστικά	22
2.1.2 Αρχές λειτουργίας	26
2.1.3 Software.....	31
2.2 Η συσκευή Leap motion.....	33
2.2.1 Αρχές λειτουργίας	34
2.2.2 Software.....	38
2.3 Σύγκριση Microsoft Kinect - Leap Motion	40
3. Ανασκόπηση Βιβλιογραφίας (State of art).....	42
3.1 Τεχνολογία Leap motion	42
3.2 Τεχνολογία Kinect.....	48
4. Εφαρμογή	54
4.1 Καταγραφή και ανάλυση δεδομένων	68
5. Συμπεράσματα και προτάσεις.....	77
Βιβλιογραφία.....	79

ΚΕΦΑΛΑΙΟ 1

1.Εισαγωγή

1.1 Στόχος

Τα τελευταία χρόνια, οι εξελίξεις στο χώρο της τεχνολογίας είναι ραγδαίες με θετικές επιρροές σε πάρα πολλούς τομείς. Υπάρχει όμως η άποψη ότι η ανάπτυξη αυτή της τεχνολογίας απειλεί την επιθυμητή εξέλιξη και ανάπτυξη του πολιτισμού. Οι έννοιες της τεχνολογίας και του πολιτισμού θεωρούνται από πολλούς αναγωνιστικές, αλλά είναι φανερό ότι η τεχνολογία είναι μέρος του πολιτισμού μιας κοινωνίας και υπό μία έννοια ορίζει και το βαθμό ανάπτυξής της. Η αρχική σκέψη για το θέμα της διπλωματικής αυτής εργασίας ήταν η χρήση της τεχνολογίας για τη διατήρηση της πολιτιστικής κληρονομιάς. Πιο συγκεκριμένα η διαφύλαξη της άυλης πολιτιστικής κληρονομιάς, των προφορικών παραδόσεων, της γλώσσας, των τεχνών (χορός, μουσική), η οποία παρουσιάζει μεγαλύτερες δυσκολίες στην τεκμηρίωσή της. Εστιάζοντας στη μουσική, στο χορό και στη γλώσσα κοινός παρονομαστής όλων αυτών των πεδίων είναι η καταγραφή της κίνησης. Στον τομέα της καταγραφής της ανθρώπινης κίνησης έχει έρθει στην επιφάνεια μια εναλλακτική προσέγγιση. Αντί της απλής καταγραφής βίντεο, οι πραγματικές κινήσεις μπορούν πλέον να καταγραφούν με τη βοήθεια ειδικών αισθητήρων. Τα συστήματα αυτά ταξινομούνται κυρίως σε τρεις κατηγορίες: α) στα οπτικά συστήματα καταγραφής κίνησης, β) στους αδρανειακούς αισθητήρες κίνησης και γ) στα συστήματα καταγραφής κίνησης χωρίς τη χρήση στόχων. Ο τρόπος λειτουργίας των συστημάτων αυτών περιγράφεται αναλυτικότερα παρακάτω.

Οπτικά Συστήματα Καταγραφής Κίνησης: Τα οπτικά συστήματα καταγραφής κίνησης (optical motion capture systems) είναι μία από τις πιο ευρέως χρησιμοποιούμενες τεχνολογίες στην ανάλυση της κίνησης, κυρίως λόγω της υψηλής ακρίβειας που προσφέρουν. Τα συστήματα αυτά διακρίνονται σε δύο βασικές κατηγορίες: i) στα παθητικά και ii) στα ενεργητικά συστήματα. Και στις δύο περιπτώσεις ο χώρος καταγραφής περιστοιχίζεται από έναν αριθμό καμερών, ο οποίος για πλήρη καταγραφή σώματος κυμαίνεται από 8 έως 16. Στην περίπτωση των παθητικών συστημάτων, προσαρμόζονται στην κάμερα φωτιόδοδοι (LEDs) υπέρυθρης ακτινοβολίας, ενώ κατάλληλα υπέρυθρα (IR) φίλτρα που επιτρέπουν τη διέλευση της υπέρυθρης ακτινοβολίας τοποθετούνται στον φακό της κάμερας. Ταυτόχρονα, αντανακλαστικοί στόχοι (markers) που βρίσκονται σε συγκεκριμένα σημεία του σώματος του χορευτή, χρησιμοποιούνται για τη συνεχή παρακολούθηση της κίνησής του σε πραγματικό χρόνο. Αντίθετα, τα οπτικά συστήματα καταγραφής, που βασίζονται σε ενεργητικούς markers τύπου LED, μετρούν την υπέρυθρη ακτινοβολία που εκπέμπεται από τα LED αυτά. Η θέση τους στον τρισδιάστατο χώρο υπολογίζεται στη συνέχεια με τη βοήθεια τεχνικών τριγωνισμού. Η τεχνολογία ενεργών LED βοηθά στην επίλυση σημαντικών προβλημάτων που συναντώνται στα οπτικά συστήματα καταγραφής, όπως είναι ο θόρυβος στα

δεδομένα, η πιθανή απώλεια της πληροφορίας και οι εσφαλμένες αντανakλάσεις. Το μεγάλο μειονέκτημα, ωστόσο, των συστημάτων αυτών παραμένει το υψηλό κόστος.

Αδρανειακοί Αισθητήρες Κίνησης: Οι αδρανειακοί αισθητήρες κίνησης προσαρμόζονται στο σώμα του χορευτή με τη βοήθεια κατάλληλων κουστουμιών. Η τεχνολογία αυτή βασίζεται στη χρήση επιταχυνσιόμετρων και γυροσκοπίων, ενώ η μετάδοση της πληροφορίας στον υπολογιστή πραγματοποιείται ασύρματα. Η τεχνολογία αυτή είναι λιγότερο ακριβής από αυτήν της προηγούμενης κατηγορίας, παρ' όλα αυτά είναι αρκετά σταθερή, ενώ ταυτόχρονα δεν απαιτείται η χρήση καμερών ή κατάλληλων συνθηκών φωτισμού στη σκηνή.

Συστήματα Καταγραφής Κίνησης Χωρίς τη Χρήση Στόχων: Ιδανικά, η καταγραφή της κίνησης μπορεί να γίνει χωρίς την τοποθέτηση στόχων στο σώμα του χορευτή ή τη χρήση αδρανειακών αισθητήρων. Παρά το γεγονός ότι η τεχνολογία αυτή υστερεί σε ακρίβεια σε σχέση με τις τεχνολογίες των προηγούμενων κατηγοριών συστημάτων, αποτελεί σίγουρα το μέλλον στο χώρο της καταγραφής της ανθρώπινης κίνησης. Για αυτόν το λόγο, τα τελευταία χρόνια υπάρχει σημαντικό ερευνητικό ενδιαφέρον προς την κατεύθυνση αυτή. Ιδιαίτερα οι πρόσφατες εξελίξεις στις τεχνολογίες καμερών βάθους χαμηλού κόστους, όπως για παράδειγμα είναι ο αισθητήρας Microsoft Kinect, επιτρέπουν τη δημιουργία τρισδιάστατων χαρτών βάθους, αλλά και πληροφορίας σχετικά με τη θέση των αρθρώσεων του ανθρώπινου σώματος (skeletal data) σε πραγματικό χρόνο. Ο συνδυασμός τέτοιου είδους πληροφορίας από πολλαπλούς αισθητήρες βάθους επιτρέπει την αποδοτικότερη μοντελοποίηση του ανθρώπινου σκελετού στις τρεις διαστάσεις, την καλύτερη κάλυψη της σκηνής και την αντιμετώπιση προβλημάτων απόκρυψης που οφείλονται είτε στη στάση του σώματος του ίδιου του χορευτή είτε στην ύπαρξη άλλων χορευτών στη σκηνή (Kitsikidis, 2014). Λόγω των πλεονεκτημάτων αυτών, η συγκεκριμένη τεχνολογία έχει χρησιμοποιηθεί, μεταξύ άλλων, και στα πλαίσια του ερευνητικού έργου FP7 iTreasures (<http://www.i-treasures.eu/>) που έχει ως στόχο την ανάλυση, μοντελοποίηση, μετάδοση και ανάδειξη της άυλης πολιτιστικής κληρονομιάς. (Dimitropoulos, 2014)

Λαμβάνοντας υπόψη τους παραπάνω τύπους καταγραφής κίνησης έγινε αρχικά η επιλογή της συσκευής Kinect λόγω της ευρείας χρήσης της και της οικονομικότητάς της. Αρχικό θέμα της έρευνας αυτής ήταν η χρήση πολλαπλών αισθητήρων για την καταγραφή της κίνησης, η ανάλυση και η αξιολόγηση των δεδομένων. Δυσκολίες στην απόκτηση των συσκευών αλλά και τα περιορισμένα χρονικά περιθώρια οδήγησαν στην επιλογή μιας άλλης εξίσου πρωτοπόρας και αρκετά πιο οικονομικής συσκευής, του αισθητήρα Leap Motion. Καταλήγοντας, ο κύριος στόχος της έρευνας είναι η διερεύνηση του σχετικά νέου συστήματος καταγραφής κίνησης Leap Motion, η ανάλυση του τρόπου λειτουργίας του, η ανάλυση της ακρίβειας και της αξιοπιστίας τόσο ενός αλλά και πολλαπλών αισθητήρων.

1.2 Δομή

Η διπλωματική αυτή αποτελείται από 6 κεφάλαια και είναι οργανωμένη ως εξής:

- Στο 1ο κεφάλαιο, στην εισαγωγή παρουσιάζεται ο στόχος και προσδιορίζεται το αντικείμενο της εργασίας.
- Στο 2ο κεφάλαιο γίνεται η περιγραφή των συσκευών Kinect και Leap Motion. Αναλύονται οι βασικές αρχές λειτουργίας των συσκευών, τα τεχνικά τους χαρακτηριστικά αλλά και τα διαθέσιμα λογισμικά που μπορούν να χρησιμοποιηθούν. Επίσης γίνεται και η σύγκριση των δύο αισθητήρων.
- Στο 3ο κεφάλαιο γίνεται μια αναφορά στις εφαρμογές που έχουν ήδη χρησιμοποιήσει τις συσκευές Kinect και Leap Motion.
- Στο 4ο κεφάλαιο περιγράφεται η πρακτική εφαρμογή και τα αποτελέσματά της.
- Στο 5ο κεφάλαιο περιλαμβάνονται τα συμπεράσματα, τα σχόλια και οι προτάσεις που απορρέουν από την ανάλυση και την εφαρμογή.

Τέλος παρατίθεται η βιβλιογραφία.

ΚΕΦΑΛΑΙΟ 2

2 Περιγραφή και σύγκριση συσκευών

2.20 αισθητήρας Kinect

Η συσκευή ή το σύστημα Kinect ανακοινώθηκε για πρώτη φορά τον Ιούνιο του 2009 στο Electronic Entertainment Expo με το κωδικό όνομα "Project Natal". Ακολουθώντας την παράδοσή της να δίνει ονόματα πόλεων η Microsoft έδωσε αυτόν τον τίτλο ως φόρο τιμής προς τον γεννημένο στην Βραζιλία διευθυντή της Alex Kipman ο οποίος ξεκίνησε αυτό το project. Το όνομα Natal επιλέχτηκε επίσης επειδή σημαίνει «γενέθλιος», αντικατοπτρίζοντας έτσι την άποψη της Microsoft ότι το project σηματοδοτεί «την έναρξη νέας γενιάς συσκευών αναψυχής για το σπίτι». Παρουσιάστηκαν τρία demos για το Kinect όταν ανακοινώθηκε: Ricochet, Paint Party και Milo & Kate. Η τεχνολογία ανίχνευσης και καταγραφής σκελετού που παρουσιάστηκε έδινε την δυνατότητα ταυτόχρονης καταγραφής τεσσάρων ατόμων, με το χαρακτηριστικό των 48 σκελετικών σημείων για ένα ανθρώπινο σώμα στα 30Hz (ή 30 frames/sec).

Παρόλο που το Kinect στα αρχικά πλάνα είχε σχεδιαστεί να περιέχει ένα μικροεπεξεργαστή που θα του επέτρεπε να εκτελεί διεργασίες όπως την ανίχνευση-καταγραφή σκελετού, τον Ιανουάριο του 2010 αποκαλύφθηκε ότι η συσκευή τελικώς δεν θα περιέχει δικό της επεξεργαστή. Αντιθέτως η επεξεργασία θα καλυπτόταν από ένα πυρήνα του επεξεργαστή Xenon του Xbox 360. Στις 13 Ιουνίου, η Microsoft ανακοίνωσε ότι το σύστημα θα ονομαζόταν επίσημα Kinect από τις λέξεις kinetic (κινητικός) και connect (συνδέω), χαρακτηριστικές της φύσεως του project. Επίσης ανακοινώθηκε ότι στην Βόρειο Αμερική θα έβγαινε στην κυκλοφορία τον Νοέμβριο του 2010 όπως και ότι θα κυκλοφορούσε νέα έκδοση Xbox 360 με σχεδιασμένη για Kinect θύρα σύνδεσης. Τον Ιούνιο του 2011, η Microsoft δημοσίευσε το πακέτο ανάπτυξης λογισμικού SDK για μη εμπορική χρήση. Ενώ τον Ιούλιο κυκλοφόρησε ένα άσπρο Kinect μαζί με το «Xbox 360 Limited Edition Kinect Star Wars Bundle». Μέχρι τότε όλα τα μοντέλα Kinect είχαν κυκλοφορήσει σε μαύρο χρώμα. Στα τέλη του Οκτωβρίου του 2011, η Microsoft ανακοίνωσε την έναρξη της εμπορικής έκδοσης του προγράμματος Kinect for Windows και την προώθηση του πακέτου SDK σε εταιρείες. Τον Φεβρουάριο του 2012, η Microsoft ανακοίνωσε την κυκλοφορία της εμπορικής έκδοσης του πακέτου Kinect for Windows SDK και δήλωσε ότι πάνω από 300 εταιρείες σε 25 χώρες εργάζονται πάνω σε Kinect εφαρμογές.

Στις 21 Φεβρουαρίου το 2011, η Microsoft ανακοίνωσε ότι θα κυκλοφορούσε μια μη-εμπορική έκδοση του πακέτου λογισμικού για Windows την ερχόμενη άνοιξη, η οποία προοριζόταν για 12 χώρες. Το πακέτο περιλάμβανε οδηγούς προγραμμάτων (drivers) Windows 7 για την συσκευή Kinect. Έδωσε στους προγραμματιστές την δυνατότητα να δημιουργήσουν και να αναπτύξουν εφαρμογές χρησιμοποιώντας C++, C# ή Visual Basic με το Microsoft Visual Studio 2010, οι οποίες έχουν τα εξής χαρακτηριστικά:

- Πρόσβαση και επεξεργασία δεδομένων χαμηλού επιπέδου από τους αισθητήρες βάθους, αισθητήρες κάμερας χρώματος και τεσσάρων μικροφωνικών στοιχείων.
- Δυνατότητα εντοπισμού της εικόνας του σκελετού ενός ή δύο ανθρώπων που κινούνται μπροστά από τον πεδίο όρασης του Kinect χωρίς να χρειάζεται αρχική στάση βαθμονόμησης. Ανίχνευση περισσότερων αρθρώσεων ανά χρήστη (αστραγάλους και καρπούς), σε σχέση με άλλα SDKs.
- Προηγμένες δυνατότητες ηχητικής επεξεργασίας όπως περιορισμός θορύβου, ακύρωση ανάκλασης ήχου, προσδιορισμός ηχητικής πηγής μέσω της κατάλληλης διάταξης των μικροφώνων και ολοκλήρωση με την εφαρμογή Windows speech recognition API για φωνητική αναγνώριση.
- Παροχή βοηθητικού κώδικα και τεκμηριωμένων πληροφοριών μέσω εγγράφων (Documentation). (Κόντος, 2014)

Ο αισθητήρας Kinect βασίζεται σε τεχνολογία λογισμικού η οποία έχει αναπτυχθεί από την Rare, θυγατρική της Microsoft. Η ενσωματωμένη κάμερα χρησιμοποιεί τεχνολογία η οποία αναπτύχθηκε από την ισραηλινή εταιρία PrimeSense. Η τελευταία ανέπτυξε ένα σύστημα που μπορεί να ερμηνεύσει συγκεκριμένες χειρονομίες, καθιστώντας εφικτό τον έλεγχο ηλεκτρονικών συσκευών χωρίς κάποια άλλη συσκευή εισόδου, αλλά μόνο μέσω χειρονομιών του σώματος. Το σύστημα αυτό λειτουργεί χρησιμοποιώντας μια υπέρυθρη κάμερα, έναν προβολέα υπέρυθρων και ένα ειδικό μικροσίπ για να παρακολουθεί την κίνηση αντικειμένων και ατόμων και στις τρεις διαστάσεις. Αυτό το σύστημα τρισδιάστατου σαρωτή (3D scanner) που ονομάζεται Light Coding χρησιμοποιεί μια παραλλαγή εικόνας (Χάρτης Βάθους) η οποία μπορεί να αναπαρασταθεί σε τρισδιάστατο χώρο. Ο αισθητήρας Kinect έχει σχεδιαστεί για να τοποθετείται κατά μήκος πάνω ή κάτω από την οθόνη. Η συσκευή διαθέτει μία οπτική (RGB) κάμερα, αισθητήρα βάθους και multi-array μικρόφωνο και είναι ικανή να παρέχει πλήρη σωματική απεικόνιση τρισδιάστατη καταγραφή κίνησης, αναγνώριση προσώπου και δυνατότητες αναγνώρισης φωνής. Ο αισθητήρας βάθους αποτελείται από ένα υπέρυθρο προβολέα λέιζερ σε συνδυασμό με ένα μονόχρωματικό αισθητήρα CMOS, ο οποίος καταγράφει δεδομένα βίντεο σε 3D κάτω από οποιοδήποτε συνθήκες φωτισμού. Η τεχνολογία του λογισμικού επιτρέπει την προηγμένη αναγνώριση χειρονομιών, αναγνώριση προσώπου και αναγνώριση φωνής. Σύμφωνα με τις πληροφορίες που παρέχονται, το Kinect είναι σε θέση να εντοπίζει ταυτόχρονα έως έξι άτομα. Λειτουργεί ιδανικά σε ένα εύρος απόστασης

μεταξύ 1.2m και 3.5m, αν και μπορεί να διατηρήσει τον εντοπισμό του χρήστη σε εύρος 0.7m μέχρι 7m. Ο χώρος κίνησης που απαιτείται από το χρήστη είναι περίπου έξι τετραγωνικά μέτρα. Ο χάρτης βάθους που δημιουργείται από τον αισθητήρα είναι μία εικόνα η οποία αναλόγως με την απόσταση των αντικειμένων χρησιμοποιεί διαφορετικές αποχρώσεις για να τα αποτυπώσει και να τα απεικονίσει.

Οι πρόσφατες τεχνολογικές εξελίξεις που αφορούν στις κάμερες βάθους, όπως ο αισθητήρας Kinect της Microsoft, δημιούργησαν νέες δυνατότητες στην τρισδιάστατη καταγραφή ανθρώπων και αντικειμένων. Η συσκευή Kinect χρησιμοποιεί ως βασικό στοιχείο την κατανόηση της γλώσσας του σώματος και προσπαθεί κατά κάποιο τρόπο να κατανοήσει τις κινήσεις του χρήστη ή του αντικειμένου πριν τις αποδώσει.. Με άλλα λόγια το Kinect είναι μια νέα προσέγγιση της εικονικής πραγματικότητας. Είναι ένα από τα νεότερα και οικονομικότερα χειριστήρια για εικονική πραγματικότητα και επιτρέπει στην κονσόλα ή στον υπολογιστή να παρακολουθεί τις κινήσεις σε πραγματικό χρόνο. Η τεχνολογία αυτή είχε ξεκινήσει πριν από 15 χρόνια όμως μέχρι πριν από λίγο καιρό ήταν απαγορευτική για να την αποκτήσει κάποιος. (Βωβός,2014)



Εικόνα 2.1: Αισθητήρας Kinect

2.1.1 Τεχνικά χαρακτηριστικά

Το Kinect αρχικά παρουσιάσθηκε ως εξάρτημα κονσόλας παιχνιδιών αλλά στη συνέχεια, λόγω της κάμερας βάθους αρκετά υψηλής ακρίβειας, άρχισε να χρησιμοποιείται σε καινοτόμες εφαρμογές επεξεργασίας εικόνας ή να αντικαθιστά παλιές. Ο αισθητήρας αποτελείται από δύο τμήματα. Αυτά είναι το μηχανικό μέρος που επιτρέπει την κίνηση του οργάνου στον οριζόντιο άξονα για προσαρμογή γωνίας λήψης και το κύριο μέρος που φέρει τους αισθητήρες και τις κάμερες (κάμερα RGB, πομπός υπέρυθρων, υπέρυθρος αισθητήρας βάθους, μικρόφωνα, LED). Πιο αναλυτικά τα μέρη του αισθητήρα Kinect περιγράφονται παρακάτω.

Έγχρωμη κάμερα

Ο αισθητήρας Kinect έχει μια ενσωματωμένη κάμερα χρώματος με υποστηριζόμενες αναλύσεις 640x480 pixels στα 30 καρέ ανά δευτερόλεπτο (FPS) και 1280x960 pixels στα 12 καρέ ανά δευτερόλεπτο (FPS). Η ορατή περιοχή της κάμερας είναι 43° κατακόρυφα με 57° οριζόντια. Λειτουργεί ως πηγή video είτε για απλή καταγραφή ή για video chat μεταξύ χρηστών, ενώ μέσω αυτής γίνεται η αναγνώριση προσώπου του χρήστη. Γενικά, το frame rate ανάλογα με την ανάλυση κυμαίνεται από 9Hz μέχρι 30 Hz. Πέρα από την προφανή λειτουργία του ως παροχή video για τις δυνατότητες video chat μεταξύ χρηστών, η RGB κάμερα προσφέρει τα απαραίτητα στοιχεία για να είναι εφικτή η αναγνώριση προσώπου.

Ο ρόλος της RGB κάμερας είναι πολύ μικρότερος σε σύγκριση με τις υπόλοιπες κάμερες που είναι ενσωματωμένες στο Kinect. Η έγχρωμη κάμερα είναι υπεύθυνη για την λήψη και μετάδοση των έγχρωμων δεδομένων. Οι τύποι των δεδομένων χρώματος που υποστηρίζει ο αισθητήρας Kinect είναι οι RGB, YUV, BAYER.

Στην RGB (Red - Green - Blue) μορφή χρώματος κάθε pixel έχει μέγεθος 4 Byte όπου τα 3 πρώτα byte χρησιμοποιούνται για το Κόκκινο χρώμα (1 Byte), το Πράσινο χρώμα (1 Byte) και το Μπλε χρώμα (1 Byte), καθώς και ένα τελευταίο Byte που χρησιμοποιείται για την διαφάνεια (Transparency).

Στην YUV μορφή χρώματος το Y αντιπροσωπεύει το κανάλι της φωτεινότητας, το U αντιπροσωπεύει το κανάλι του Μπλε χρώματος και το V αντιπροσωπεύει το κανάλι του Κόκκινου χρώματος. Επειδή η YUV μορφή χρώματος χρησιμοποιεί 2 Bytes (16bits) ανά pixel, δεσμεύει λιγότερη μνήμη για την αποθήκευση των δεδομένων. Οι αναλύσεις που υποστηρίζει η κάμερα χρώματος του αισθητήρα Kinect για την YUV μορφή είναι 640x480 pixels (15 frames per second).

Τόσο η YUV όσο και η RGB μορφή χρώματος αντιπροσωπεύουν την ίδια εικόνα μιας και οι δυο μορφές υπολογίζονται από τα ίδια δεδομένα της κάμερας. Η κάμερα χρώματος του Kinect μπορεί επίσης να επιστρέψει ένα frame εικόνας, χρησιμοποιώντας ένα συνδυασμό των χρωμάτων κόκκινου, πράσινου και μπλε (50% πράσινο, 25% κόκκινο και 25% μπλε). Το φίλτρο που δημιουργεί την εικόνα λέγεται φίλτρο Bayer, και χρησιμοποιείται με αυτόν τον τρόπο επειδή το ανθρώπινο μάτι είναι πιο ευαίσθητο στο πράσινο χρώμα και μπορεί να εντοπίσει περισσότερες αλλαγές στο πράσινο παρά στα άλλα χρώματα. Οι αναλύσεις που υποστηρίζει η κάμερα χρώματος για την Bayer μορφή είναι 640x480 pixels (30 frames per second) και 1280x960 pixels (12 frames per second).

Πομπός και δέκτης υπέρυθρων

Για τον υπολογισμό του βάθους των σημείων ενός αντικειμένου, ο αισθητήρας KINECT χρησιμοποιεί τις δύο αυτές συσκευές υπέρυθρων. Συνοπτικά ο πομπός υπέρυθρων εκπέμπει ένα πλαίσιο από υπέρυθρες κουκκίδες φως σε ένα "ψευδο-τυχαίο" μοτίβο στο οπτικό πεδίο του αισθητήρα των οποίων την αντανάκλαση ανιχνεύει ο υπέρυθρος αισθητήρας βάθους. Καθώς οι υπέρυθρες ακτίνες ανακλώνται από τα διάφορα αντικείμενα ο αισθητήρας βάθους διαβάζει και μετατρέπει σε πληροφορίες βάθους, μετρώντας την απόσταση μεταξύ του αισθητήρα και του αντικειμένου από όπου η κάθε κουκίδα διαβάστηκε. Πιο συγκεκριμένα κάθε τιμή βάθους είναι ανάλογη του χρόνου που θα κάνει μια ανακλώμενη ακτίνα να επιστρέψει πίσω στο δέκτη υπέρυθρων. Όσο μεγαλύτερη η απόσταση ενός σημείου από τον δέκτη τόσο μεγαλύτερη η τιμή βάθους για το σημείο αυτό. Παρέχεται επίσης μονόχρωμο βίντεο βάθους σε VGA ανάλυση (320x240) με 16 bit βάθος και 2.048 επίπεδα ευαισθησίας. Το Kinect μπορεί επίσης να προβάλει εικόνα απευθείας από την IR κάμερα ακόμα και σε ανάλυση 1280x1024 σε χαμηλότερο όμως frame rate. Το γωνιακό πεδίο όρασης της κάμερας βάθους είναι 57° οριζόντια και 43° κατακόρυφα και η απόσταση βέλτιστης λειτουργίας είναι μεταξύ 0.8m και 3.5m.

Σειρά μικροφώνων

Η σειρά μικροφώνων αποτελείται από 4 μικρόφωνα, ποιότητας ήχου 16 bit με συχνότητα δειγματοληψίας 16kHz. Το κύριο πλεονέκτημα της χρησιμοποίησης 4 μικροφώνων αντί του ενός είναι ότι μπορεί να γίνει πιο αποτελεσματική αναγνώριση της φωνής, συγκρίνοντας την χρονική στιγμή που το κάθε μικρόφωνο λαμβάνει το ίδιο ηχητικό σήμα ο αισθητήρας μπορεί να εξακριβώσει την γωνία της προέλευσης του σήματος. Τα μικρόφωνα αυτά έχουν δυνατότητα απομόνωσης θορύβου και ενίσχυσης φωνής καθώς και αναγνώρισης φωνητικών εντολών σε πολλές γλώσσες. Στην αρχή, η φωνητική

αναγνώριση ήταν διαθέσιμη μόνο στην Ιαπωνία, τις ΗΠΑ, τον Καναδά και το Ηνωμένο Βασίλειο. Πλέον, αυτή η δυνατότητα υποστηρίζεται σε 11 χώρες σε όλο τον κόσμο.

Μηχανοκίνητη Βάση (Motorized Tilt)

Η συσκευή διαθέτει επίσης μια μηχανοκίνητη βάση η οποία ρυθμίζεται δυναμικά μέσω κώδικα και ο σκοπός της είναι η ρύθμιση της κλίσης του αισθητήρα. Η κίνηση της βάσης προέρχεται από ένα μικρό μοτέρ, έναν κινητήρα στο μέγεθος νομίσματος και τρία εύθραυστα πλαστικά γρανάζια, τα οποία είναι ευαίσθητα στη θερμότητα και ίσως αποτελούν το πιο αδύναμο σημείο της συσκευής. Ο κινητήρας κλίσης στη βάση του Kinect μπορεί να στρέφει τη συσκευή κατά 27° προς τα πάνω ή προς τα κάτω, αλλάζοντας το πεδίο όρασης των αισθητήρων αντίστοιχα. Έτσι δίνεται η δυνατότητα προσαρμογής του οπτικού πεδίου του αισθητήρα για βέλτιστη λήψη εικόνων, 27 μοίρες προς τα πάνω και 27 μοίρες προς τα κάτω όταν είναι σταθερά τοποθετημένος (σύνολο 54° στην κάθετη διεύθυνση). Επειδή, το μοτέρ κλίσης χρειάζεται περισσότερη ενέργεια από όση μπορούν να δώσουν οι USB θύρες του Xbox 360, η συσκευή χρησιμοποιεί έναν επιπρόσθετο προσαρμογέα που συνδυάζει την επικοινωνία μέσω USB και την παροχή επιπρόσθετης ενέργειας. Τα επόμενα μοντέλα Xbox 360 περιλαμβάνουν μία ειδική AUX θύρα που καλύπτει αυτή την ανάγκη.

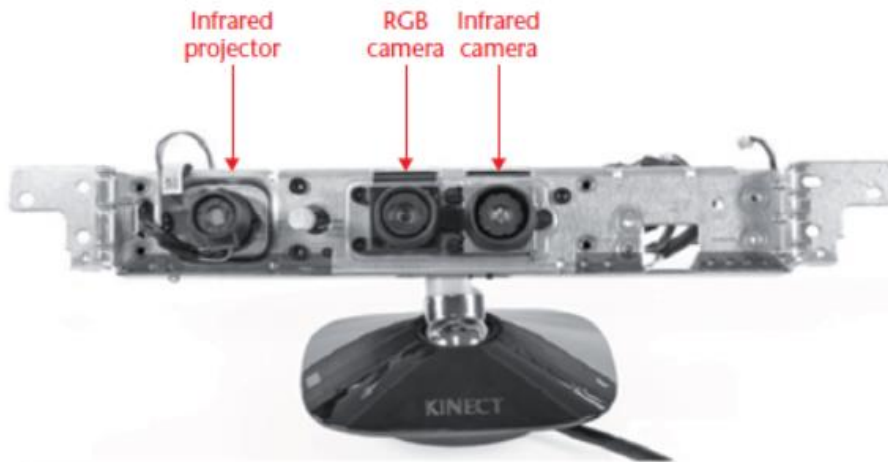
Επιταχυνσιόμετρο (Accelerometer)

Το Kinect παρέχει ένα επιταχυνσιόμετρο τριών αξόνων (KXSD9-1026) το οποίο παρέχει την πληροφορία της θέσης της συσκευής επιστρέφοντας ένα τρισδιάστατο διάνυσμα με την θετική κατεύθυνση του άξονα Z να είναι προς την κατεύθυνση του αισθητήρα. Οι τιμές του X και του Y καθορίζουν την κύλιση και την κλίση, ενώ το Z καθορίζει εάν το Kinect είναι ανάποδα ή όχι. Το επιταχυνσιόμετρο μετρά μονάχα την κλίση, κι όχι τον προσανατολισμό της συσκευής. Το χαρακτηριστικό αυτό του Kinect έχει χρησιμοποιηθεί ευρέως στον τομέα της ρομποτικής. (Συντυχάκης, 2012).



Figure copied from Kinect for Windows SDK Quickstarts

Εικόνα 2.2: Τεχνικά χαρακτηριστικά του αισθητήρα Kinect

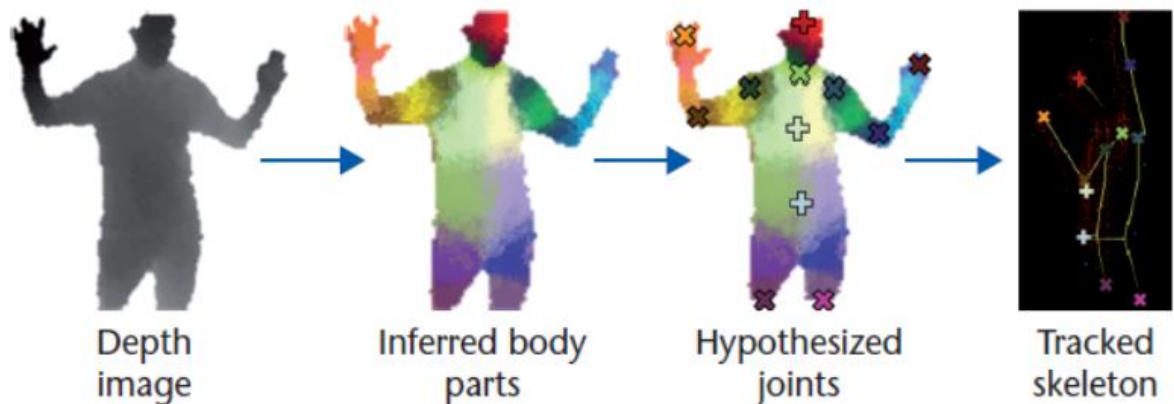


Εικόνα 2.3: Τεχνικά χαρακτηριστικά στο εσωτερικό τμήμα του αισθητήρα Kinect

2.1.2 Αρχές λειτουργίας

Αντίθετα με το προφανές συμπέρασμα ότι το Kinect μέσω του εξοπλισμού και μιας σειράς αλγορίθμων υπολογίζει τις κινήσεις του σώματος του χρήστη, η αλήθεια είναι ελαφρώς διαφορετική. Πρακτικά, το Kinect «συμπεραίνει» τις κινήσεις που εκτελεί ο χρήστης.

Η διαδικασία του προσδιορισμού της θέσης του σώματος αποτελείται από 3 στάδια. Πρώτα, κατασκευάζεται ένας χάρτης βάθους (depth map) χρησιμοποιώντας την τεχνική δομημένου φωτός (structured light), στην συνέχεια προσδιορίζονται τα μέλη σώματος μέσω εκμάθησης μηχανής (machine learning) που έχει προηγηθεί και, τέλος, υπολογίζονται οι θέσεις των αρθρώσεων ενός 3D σκελετού που αναπαριστά τη θέση και την στάση του χρήστη.



Εικόνα 2.4: Τα βήματα της διαδικασίας προσδιορισμού της θέσης του σώματος

Κατασκευή πίνακα βάθους

Το Kinect χρησιμοποιεί έναν προβολέα και έναν αισθητήρα υπέρυθρων και όχι την RGB κάμερα για τον υπολογισμό του βάθους. Η RGB κάμερα, ωστόσο, χρησιμοποιείται σε μεγάλο βαθμό κατά την βελτιστοποίηση των δεδομένων. Η τεχνική της ανάλυσης ενός ήδη γνωστού μοτίβου ονομάζεται τεχνική δομημένου φωτός και ορίζει ότι προβάλλοντας ένα γνωστό μοτίβο πάνω στη σκηνή, μπορείς να συμπεράνεις το βάθος από την παραμόρφωση του μοτίβου σε κάθε σημείο.

Η απεικόνιση βάθους είναι η διαδικασία σχηματισμού μιας εικόνας στην οποία το χαρακτηριστικό που ενδιαφέρει τον χρήστη είναι η απόσταση των σημείων από την κάμερα. Πρέπει να αναφερθεί ότι στην απεικόνιση βάθους, η τιμή κάθε εικονοστοιχείου

είναι η απόσταση του σημείου του αντικειμένου που αντιπροσωπεύεται από το συγκεκριμένο εικονοστοιχείο σε σχέση με το σύστημα λήψης της απεικόνισης.

Ο αισθητήρας KINECT περιέχει ένα υπέρυθρο πομπό ακτίνων laser, μια υπέρυθρη κάμερα και μία έγχρωμη κάμερα. Ο πομπός laser εκπέμπει μία δέσμη laser η οποία διαχωρίζεται σε πολλές δέσμες μετά την διάθλασή της για την δημιουργία ενός σταθερού μοτίβου από κουκκίδες οι οποίες προβάλλονται στην σκηνή. Αυτό το μοτίβο απεικονίζεται από την υπέρυθρη κάμερα και συσχετίζεται με ένα μοτίβο αναφοράς. Το μοτίβο αναφοράς αποκτάται από την απεικόνιση του χώρου σε μία γνωστή απόσταση από τον αισθητήρα και αποθηκεύεται στην μνήμη του αισθητήρα. Όταν μία κουκκίδα προβληθεί στο αντικείμενο προς απεικόνιση τότε αν η απόστασή της είναι μικρότερη ή μεγαλύτερη από την απόσταση όπου αποκτήθηκε το μοτίβο αναφοράς, η θέση της κουκκίδας στην υπέρυθρη εικόνα θα μετακινηθεί στην διεύθυνση προς την νοητή γραμμή που ενώνει την κάμερα με την πηγή υπέρυθρων. Έτσι η απόσταση για κάθε εικονοστοιχείο μπορεί να ανακτηθεί από την αντίστοιχη διαφορά.

Το Kinect συνδυάζει το δομημένο φως, εκπέμποντας μέσω του προβολέα ένα γνωστό μοτίβο στιγμάτων, με δύο τεχνικές της υπολογιστικής όρασης: τον υπολογισμό βάθους μέσω εστίασης και τον στερεοσκοπικό υπολογισμό βάθους. Ο υπολογισμός βάθους μέσω εστίασης βασίζεται στο γεγονός ότι τα αντικείμενα φαίνονται πιο θολά, όσο πιο μακριά βρίσκονται. Το Kinect βελτιώνει δραματικά την ακρίβεια αυτής τεχνικής, καθώς χρησιμοποιεί έναν αστιγματικό φακό με διαφορετικά μήκη εστίασης για την κατεύθυνση στον x και στον y άξονα. Οπότε, όταν ο προβολέας δείχνει έναν κύκλο, πρακτικά γίνεται μία έλλειψη της οποίας ο προσανατολισμός μεταβάλλεται ανάλογα με το βάθος.

Αντίστοιχα, ο στερεοσκοπικός υπολογισμός βάθους προκύπτει από την έννοια της παράλλαξης, η οποία δηλώνει ότι η αλλαγή της θέσης του παρατηρητή οδηγεί σε αλλαγή της γωνίας του προς παρατήρηση αντικειμένου σε σχέση με αυτόν. Οπότε, όταν παρατηρείται η σκηνή από διαφορετική γωνία, τα αντικείμενα τα οποία βρίσκονται κοντά μετακινούνται περισσότερο σε σχέση με τα αντικείμενα που βρίσκονται μακριά. Το Kinect αναλύει τη μεταβολή της γωνίας ενός μοτίβου σημείων προβάλλοντάς το από μία θέση και παρατηρώντας το από μια άλλη. Αυτός είναι και ο ρόλος του προβολέα και του αισθητήρα υπέρυθρων.



Shpunt et al, PrimeSense patent application
US 2008/0106746

Εικόνα 2.5: Το σταθερό μοτίβο αναφοράς της δέσμης laser

Δημιουργία βάσης δεδομένων

Γενικά, το ανθρώπινο σώμα είναι ικανό να εκτελέσει ένα τεράστιο εύρος στάσεων - κινήσεων (poses), οι οποίες είναι δύσκολο να προσομοιωθούν. Για αυτό το λόγο, η ομάδα έρευνας της Microsoft δημιούργησε μία μεγάλη βάση δεδομένων ανθρώπινων κινήσεων μέσω motion capture (mocap). Ο στόχος ήταν να κρατήσουν την ευρεία ποικιλία στάσεων που παίρνουν οι άνθρωποι σε μία υποθετική δραστηριότητα. Η βάση δεδομένων αποτελείται από περίπου 500.000 frames μερικών εκατοντάδων ενεργειών όπως οδήγηση, χορός, τρέξιμο. Στην περίπτωση μη ορατών στάσεων, το σύστημα προβαίνει σε γενίκευση των δεδομένων. Πιο συγκεκριμένα, δεν χρειάζεται να καταγράφονται όλοι οι πιθανοί συνδυασμοί των διαφόρων άκρων, καθώς είναι επαρκές ένα μεγάλο εύρος στάσεων. Επίσης, δεν είναι αναγκαία η καταγραφή σε mocap των περιστροφών, των συντεταγμένων του χώρου, της συμμετρίας στον κατακόρυφο άξονα, του σχήματος και του μεγέθους του σώματος ή της θέσης της κάμερας, καθώς όλα αυτά προκύπτουν ημιαυτόματα.

Καθώς το σύστημα δεν χρησιμοποιεί συνεχή δεδομένα, αυτό που πρακτικά ενδιαφέρει είναι μόνο στατικές πόζες και όχι κίνηση. Συχνά, αλλαγές στην στάση από ένα mocap frame σε ένα άλλο είναι τόσο μικρές που είναι αμελητέες. Οπότε, αποβάλλονται πολλές παρόμοιες, περιττές στάσεις σε σχέση με το αρχικό mocap δεδομένο, χρησιμοποιώντας τον αλγόριθμο ομαδοποίησης «πιο μακρινού γείτονα» (furthest neighbor clustering), όπου η απόσταση ανάμεσα σε δύο στάσεις ορίζεται ως η μέγιστη Ευκλείδεια απόσταση

των μελών του σώματος. Το τελικό αποτέλεσμα είναι ένα υποσύνολο 100.000 στάσεων, τέτοιες ώστε καμία στάση να μην έχει απόσταση μικρότερη από 5 εκατοστά (cm) σε σχέση με μία άλλη. Για την βελτιστοποίηση της ποιότητας της βάσης δεδομένων mocap με περιοχές από τμήματα στάσεων που δεν είχαν συμπεριληφθεί, απαραίτητη ήταν η διαρκής επανάληψη της διαδικασίας της δειγματοληψίας από το μοντέλο, της εκπαίδευσης του συστήματος και του ελέγχου της ακρίβειας πρόβλεψης των αρθρώσεων του σκελετού. Τα πρώτα πειράματα έγιναν με τη βάση δεδομένων mocap του Carnegie Mellon University (CMU) που είχαν αποδεκτά αποτελέσματα, αν και καλύφθηκε πολύ μικρότερο εύρος περιοχών από στάσεις.

Οι ερευνητές δημιούργησαν μια τυχαία δρομολόγηση εκτέλεσης διεργασιών μέσω της οποίας μπορούσαν να πειραματιστούν με δοκιμαστικές εικόνες με ήδη ονομασμένα τα μέλη του σώματος. Οι στόχοι ήταν δύο, ρεαλισμός και ποικιλία. Για να δουλέψει σωστά το υπό εκπαίδευση μοντέλο, έπρεπε τα δείγματα να μοιάζουν πολύ με κανονικές εικόνες και να περιλαμβάνουν μεγάλο εύρος των πιθανών διαφοροποιήσεων στην εμφάνιση που ήλπιζαν να αναγνωριστούν κατά τη διάρκεια των δοκιμών. Τυχόν διαφοροποιήσεις στο βάθος/κλίμακα της εικόνας και στη μετατόπιση του σώματος του χρήστη μπορούν να υπολογιστούν από επεξεργασία των δεδομένων, ενώ άλλες αλλαγές όπως η θέση της κάμερας, η θέση, το μέγεθος και το σχήμα του σώματος δε γίνεται να υπολογιστούν. Οπότε προκύπτουν κατευθείαν ως δεδομένα που εξάγονται από τις εικόνες. Η διαδικασία δειγματοληψίας ξεκινά παίρνοντας ως δείγματα από συνθετικά μοντέλα ένα τυχαίο σύνολο παραμέτρων και στη συνέχεια, χρησιμοποιώντας κλασικές τεχνικές γραφικών, υπολογίζεται το βάθος και σχεδιάζονται εικόνες τμημάτων του σώματος έχοντας ως βάση 3D μοντέλα. Στη συνέχεια, γίνεται αντιστοίχιση με κάποιο από τα μοντέλα αναφοράς, ανάλογα με το μέγεθος και το σχήμα του σώματος.

Προσδιορισμός θέσης σώματος

Η αρχική και πλέον κλασική προσέγγιση που έχει χρησιμοποιηθεί στους αισθητήρες κίνησης είναι η εξής: ο χρήστης έπρεπε αρχικά να μείνει για μερικά δευτερόλεπτα σε μία στάση προσαρμογής (PSI Pose), έτσι ώστε ο αλγόριθμος να εντοπίσει τα μέλη του σώματος κάνοντας απλή αντιστοίχιση. Στη συνέχεια αρκούσε ο αλγόριθμος εντοπισμού να ακολουθεί την κίνηση του σώματος. Η βασική ιδέα πίσω από αυτό είναι ότι έχοντας μία περιοχή αναγνωρισμένη ως χέρι στο πρώτο frame, στο επόμενο frame το χέρι δεν γίνεται να έχει μετατοπιστεί πολύ μακριά, οπότε απλώς προσπαθεί να κάνει αντιστοίχιση σε περιοχές κοντινές στην προηγούμενη θέση. Ένα ιδιαίτερα σημαντικό κομμάτι είναι η ονοματοδοσία των ενδιάμεσων τμημάτων του σώματος. Ορίζονται πολλές ετικέτες για τα μέλη του σώματος, χρησιμοποιώντας χρωματικό κώδικα, οι οποίες καταλήγουν να καλύπτουν πλήρως το σώμα. Κάποια από αυτά τα μέλη προσδιορίζονται για να προκύπτουν άμεσα από αυτά οι αρθρώσεις του σκελετού που αναπαριστά τις κινήσεις

του χρήστη, ενώ άλλα απλώς καλύπτουν τα ενδιάμεσα κενά ή είναι απαραίτητα για την πρόβλεψη της θέσης άλλων αρθρώσεων. Η ακρίβεια αυτού του τρόπου προσδιορισμού μπορεί να αλλάξει ανάλογα με την αντίστοιχη εφαρμογή. Για παράδειγμα μία εφαρμογή που χρειάζεται μόνο το πάνω μέρος του σώματος, μπορεί να δεχθεί το κάτω μέρος ως ένα ενιαίο κομμάτι, για εξοικονόμηση χώρου και πόρων. Τα μέλη πρέπει να διαχωρίζονται με τέτοιο τρόπο ώστε να είναι μεν επαρκώς μικρά για να γίνεται σωστά η αντιστοίχιση, αλλά όχι τόσο πολυάριθμα ώστε να προκύπτει σπατάλη χώρου στο σύστημα.

Θεωρητικά οι αλγόριθμοι που χρησιμοποιούνται είναι αρκετά αποτελεσματικοί, ωστόσο πρακτικά αποτυγχάνουν όταν η θέση του σώματος χαθεί για κάποιο λόγο και είναι ιδιαίτερα κακοί στο να αντιμετωπίζουν την ύπαρξη άλλων αντικειμένων, τα οποία εμποδίζουν τον εντοπισμό του χρήστη, ακόμα και για πολύ μικρό χρονικό διάστημα. Επιπρόσθετα, ο εντοπισμός παραπάνω του ενός χρηστών ταυτόχρονα είναι δύσκολος, ενώ μόλις χαθεί η θέση του σώματος, χρειάζεται χρόνος για να επανεντοπιστεί, αν και εφόσον είναι εφικτό.

Η αντιμετώπιση της Microsoft σε αυτό το ζήτημα ήταν να σχεδιαστεί ένα σύστημα εντοπισμού της θέσης του σώματος, το οποίο δεν θα βασιζόταν στον εντοπισμό κάθε μέλους του σώματος, αλλά στην τοπική ανάλυση κάθε pixel. (Δουλφής, 2016)

2.1.3 Software

Με τον όρο λογισμικό υπολογιστών ή λογισμικό (software) ορίζεται η συλλογή από προγράμματα υπολογιστών, διαδικασίες και οδηγίες χρήσης που εκτελούν ορισμένες εργασίες σε ένα υπολογιστικό σύστημα.

Ο όρος περιλαμβάνει:

- το λογισμικό εφαρμογών, όπως οι επεξεργαστές κειμένου, που εκτελούν παραγωγικές εργασίες για τους χρήστες,
- το λογισμικό συστήματος, όπως τα λειτουργικά συστήματα, που παρέχει τις αναγκαίες υπηρεσίες του υλικού στο λογισμικό εφαρμογών,
- το ενδιάμεσο λογισμικό (middleware), που ελέγχει και συντονίζει τα κατανεμημένα συστήματα, και
- το υλικολογισμικό που προγραμματίζει σε χαμηλό επίπεδο το υλικό ενός υπολογιστή ή τα περιφερειακά του.

Το λογισμικό περιλαμβάνει τα προγράμματα, τους ιστότοπους, τα βιντεοπαιχνίδια και άλλα προγράμματα που έχουν αναπτυχθεί σε μια γλώσσα προγραμματισμού όπως για παράδειγμα η C, η C++ ή η Java. Το λογισμικό είναι κωδικοποιημένο με συγκεκριμένο τρόπο, με τη βοήθεια ενός δυαδικού ψηφιακού συστήματος, ώστε να είναι «κατανοητό» από το υλικό.

Παρακάτω θα γίνει ειδικότερη αναφορά στο λογισμικό που χρησιμοποιείται κατά κύριο λόγο για τον αισθητήρα Kinect.

Μόλις κυκλοφόρησε το Kinect, πέρα από την αποδεδειγμένα εντυπωσιακή αποδοχή που είχε από τους χρήστες του Xbox 360, υπήρξε άμεσα τεράστιο ενδιαφέρον για δημιουργία νέων εφαρμογών. Ωστόσο, δεν συνοδεύονταν από κάποιο Πακέτο Ανάπτυξης Λογισμικού (SDK).

Ένα Πακέτο Ανάπτυξης Λογισμικού (Software Development Kit - SDK ή devkit) είναι ένα σύνολο εργαλείων ανάπτυξης που επιτρέπουν σε έναν προγραμματιστή να δημιουργήσει λογισμικό εφαρμογών για ένα συγκεκριμένο πακέτο λογισμικού, πλατφόρμα, παιχνιδομηχανή, λειτουργικά συστήματα κτλ.

Η αρχιτεκτονική του SDK περιλαμβάνει

- KinectHardware: Τον αισθητήρα Kinect καθώς και το USB μέσω του οποίου είναι συνδεδεμένος ο αισθητήρας με τον υπολογιστή.
- KinectDrivers : Τους οδηγούς των Windows για το Kinect οι οποίοι εγκαθίστανται ως μέρος της διαδικασίας εγκατάστασης του SDK μέσω των οποίων η συστοιχία μικροφώνων του Kinect μπορεί να χρησιμοποιηθεί και ως συσκευή ήχου ώστε να υπάρχει πρόσβαση και μέσα από το API των Windows, γίνεται ο έλεγχος των ρευμάτων ήχου και βίντεο (για το χρώμα, το βάθος και το σκελετό) και η εκάστοτε εφαρμογή μπορεί να χρησιμοποιεί περισσότερους από έναν αισθητήρες Kinect.

- NUIAPI : Φυσική Διεπαφή Χρήστη παρακολουθώντας τα ρεύματα εξόδου για το σκελετό, το χρώμα και το βάθος απεικόνισης.
- DirectX Media Object (DMO) για beamforming και τις συστοιχίες μικροφώνων.
- WindowAPIs για τον ήχο και την ομιλία

Το API είναι μια συντομογραφία του όρου application-programming interface (Διεπαφή προγραμματισμού εφαρμογών), το οποίο είναι ένα σύνολο από εργαλεία, πρωτόκολλα και μεθόδους για την επικοινωνία με το λειτουργικό σύστημα ή πρόγραμμα ελέγχου ή συσκευή. Το API είναι διεπαφή ανάμεσα σε εφαρμογές, κι όχι διεπαφή χρήστη. Με τα APIs οι εφαρμογές επικοινωνούν χωρίς να γίνεται αντιληπτό από το χρήστη. Η χρήση ενός API σημαίνει ότι ένα πρόγραμμα οδήγησης ή άλλη εφαρμογή, είναι διαθέσιμο στον υπολογιστή για να κάνει κάποιες διαδικασίες.

Μέχρι στιγμής υπάρχουν πέντε αρκετά δημοφιλή APIs για την αποκωδικοποίηση των δεδομένων που παράγει το Kinect, ανάμεσα σε αυτά και το official SDK της Microsoft, τα οποία είναι ελεύθερα προς χρήση: Libfreenect, OpenNI και NiTE, CLNUI, Microsoft Kinect SDK, Evoluce SDK (το οποίο είναι βασισμένο στο OpenNI).

OpenNI και NiTE

Δημιουργήθηκε από έναν αφιλοκερδή οργανισμό ο οποίος απαρτίζεται από διάφορες εταιρίες, συμπεριλαμβανομένης και της PrimeSense Ltd. οι οποίες θέλησαν να θέσουν ένα βιομηχανικό πρότυπο λειτουργικότητας για τις συσκευές φυσικής διεπαφής χρήστη-υπολογιστή (Natural User Interaction Devices) και κυκλοφόρησε το Δεκέμβριο του 2010. Το OpenNI αναπτύχθηκε σε C/C++, έτσι ώστε να μπορεί να χρησιμοποιηθεί από διάφορα λειτουργικά συστήματα, όπως Mac OSX, Ubuntu, Windows. Είναι το επίσημο λογισμικό των Χτίον συσκευών της Asus, αλλά μπορεί να λειτουργήσει και με το Kinect. Παρέχει επικοινωνία με τον αισθητήρα βάθους, την κάμερα χρώματος, τα μικρόφωνα και τη μηχανοκίνητη βάση. Το OpenNI, όμως, συνοδεύεται και από μια ενδιάμεση βιβλιοθήκη η οποία λέγεται NiTE και είναι εξοπλισμένο με τεχνολογίες αναγνώρισης φωνής, αναγνώρισης χειρονομιών χεριών, και ανίχνευσης σκελετού. Η ανίχνευση του σκελετού του χρήστη απαιτεί αρχική στάση βαθμονόμησης αλλά στις πιο πρόσφατες εκδόσεις του, το OpenNI/NiTE, δεν χρειάζεται αρχική στάση βαθμονόμησης (χαρακτηριστικό που είχε μόνο το MS SDK). (Αρβανίτης, 2015).

2.2 Η συσκευή Leap motion

Η Leap Motion είναι μια αμερικάνικη εταιρεία που κατασκευάζει τη συσκευή Leap Motion controller. Το Leap Motion controller είναι μια συσκευή αισθητήρα φυσικής αλληλεπίδρασης χρήστη, ανάλογη με το υπολογιστικό ποντίκι, που υποστηρίζει την είσοδο δεδομένων στον υπολογιστή χωρίς άγγιγμα ή επαφή, δηλαδή απελευθερώνει τα χέρια του χρήστη. Σκοπός του αισθητήρα ήταν να αντικαταστήσει το ποντίκι ή το πληκτρολόγιο κάτι το οποίο εγκαταλείφθηκε στην πορεία. Η τεχνολογία του όμως μπορεί πέρα από την βιομηχανία του gaming να εφαρμοστεί και σε άλλους τομείς, όπως στην χειρουργική, την αρχιτεκτονική, την μηχανική και τον σχεδιασμό.

Η τεχνολογία Leap Motion αναπτύχθηκε για πρώτη φορά το 2008, από τον David Holz, ο οποίος ίδρυσε την εταιρεία σε συνεργασία με τον Micheal Buckward. Το 2010 για πρώτη φορά η εταιρεία έβγαλε ανακοίνωση για το πρώτο της προϊόν που αρχικά ονομαζόταν «The Leap». Στο πλαίσιο της ανάπτυξης εφαρμογών για το προϊόν, κυκλοφόρησε περίπου 12.000 συσκευές για τους προγραμματιστές που ενδιαφέρονταν να δημιουργήσουν εφαρμογές για το προϊόν. Σύντομα οι πωλήσεις εκτοξεύτηκαν και η εταιρεία κυκλοφόρησε την δεύτερη έκδοση λογισμικού για προγραμματιστές το 2014. Η εταιρεία συνεργάζεται με την ASUS και την HP για να ενσωματώσει την τεχνολογία της σε μια νέα γενιά υπολογιστών, ακόμα και σε πληκτρολόγια.

Το Leap Motion controller είναι μια περιφερειακή συσκευή usb, μικρή σε μέγεθος, που έχει σχεδιαστεί για να τοποθετείται σε μια φυσική επιφάνεια εργασίας στραμμένη προς τα πάνω. Έχει μήκος 80 χιλιοστά, πλάτος 12,7 χιλιοστά και βάρος 45 γραμμάρια. Παρακολουθεί όλες τις κινήσεις και στα δέκα δάχτυλα του χεριού μέχρι και 1/100 του χιλιοστού. Ένα χέρι έχει 29 οστά, 29 αρθρώσεις, 123 συνδέσμους, 48 νεύρα και 30 αρτηρίες. Αποτελεί την πιο ευαίσθητη από την υπάρχουσα τεχνολογία ελέγχου κίνησης και με αυτόν τον τρόπο μας δίνει την δυνατότητα προγραμματισμού και σχεδίασης πολύ απαιτητικών εφαρμογών σε σχέση με την ακρίβεια της κίνησης. (Κωτσιδου, 2015)



Εικόνα 2.6: Το εσωτερικό τμήμα της συσκευής Leap Motion

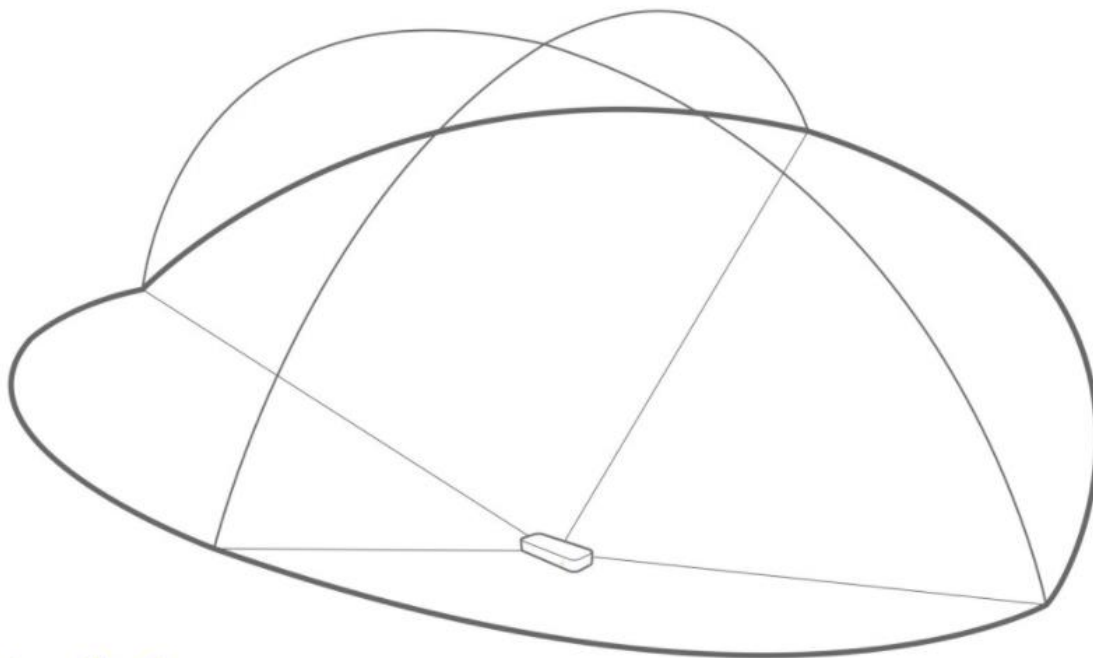
2.2.1 Αρχές λειτουργίας

Η συσκευή Leap Motion χρησιμοποιεί υπέρυθρη απεικόνιση (IR) για να προσδιορίσει τη θέση προκαθορισμένων αντικειμένων σε περιορισμένο χώρο και σε πραγματικό χρόνο. Η περιοχή που καλύπτει η ανίχνευση της συσκευής είναι ημισφαιρική. Τα LED δημιουργούν ένα 3D μοτίβο από κουκκίδες υπέρυθρου φωτός και οι κάμερες παράγουν σχεδόν 300 καρέ ανά δευτερόλεπτο, τα οποία αποστέλλονται μέσω ενός καλωδίου USB στον υπολογιστή που είναι συνδεδεμένη η συσκευή. Εκεί αναλύονται από το λογισμικό του αισθητήρα Leap Motion με χρήση πολύπλοκων μαθηματικών μεθόδων, οι οποίες δεν έχουν αποκαλυφθεί από την εταιρεία. Από τεχνική άποψη, πολύ λίγες λεπτομέρειες είναι γνωστές σχετικά με την ακριβή φύση των αλγορίθμων που χρησιμοποιούνται λόγω περιορισμών στον τομέα της ευρεσιτεχνίας και λόγω εμπορικότητας. Ωστόσο, από την επιθεώρηση της συσκευής, είναι σαφές ότι χρησιμοποιούνται τρεις ξεχωριστοί υπέρυθροι πομποί (IR LED) σε συνδυασμό με δύο υπέρυθρες (IR) κάμερες. Επομένως, μπορεί να κατηγοριοποιηθεί ως οπτικό σύστημα παρακολούθησης βάσει της αρχής της στερεοσκοπικής όρασης.

Σύμφωνα με τις επίσημες δημοσιεύσεις, το λογισμικό Leap αναλύει τα αντικείμενα που παρατηρούνται στο οπτικό πεδίο της συσκευής. Η εικόνα παρέχει ρυθμίσεις που περιέχουν τις τιμές φωτεινότητας του αισθητήρα και ρυθμίσεις που περιέχουν τα στοιχεία βαθμονόμησης της κάμερας, η οποία μπορεί να χρησιμοποιηθεί για τη διόρθωση, την παραμόρφωση του φακού και άλλες οπτικές ατέλειες στα δεδομένα της εικόνας. Τα δεδομένα εικόνας παρέχονται ως μια σειρά τιμών *pixel*. Η υπέρυθρη μορφή χρησιμοποιεί ένα *byte* ανά *pixel*, που καθορίζει τη φωτεινότητα που μετρήθηκε για την εν λόγω θέση του αισθητήρα. Τα δεδομένα εμφανίζονται σε υπέρυθρη μορφή ως κλίμακα του γκρι της εικόνας. Όταν μια ακτίνα φωτός εισέρχεται σε μία από τις κάμερες Leap Motion, ο φακός κάμπτεται την ακτίνα φωτός, έτσι ώστε να φτάσει στον αισθητήρα, ο οποίος καταγράφει σε κλίμακα του γκρι μια τιμή φωτεινότητας σε μια συγκεκριμένη θέση *pixel*. Φυσικά, υπάρχουν αποκλίσεις από την πραγματικότητα. Με τη παροχή των στοιχείων της βαθμονόμησης της κάμερας γίνεται εφικτή η διόρθωση της απόκλισης αυτής, η οποία επιτρέπει τον υπολογισμό της πραγματικής γωνίας του αρχικού φωτός.

Ο αισθητήρας αναγνωρίζει τα χέρια, τα δάχτυλα, αναφέρει διακριτές θέσεις, χειρονομίες και κίνηση. Το οπτικό πεδίο του Leap Motion είναι μια ανεστραμμένη πυραμίδα που ξεκινά από το κέντρο της συσκευής. Η αποτελεσματική εμβέλειά του εκτείνεται από περίπου 25 έως 600 χιλιοστά πάνω από τη συσκευή (1 ίντσα έως 2 πόδια) και έχει ρυθμό ανίχνευσης κίνησης στα 200 frames/sec. Οι αισθητήρες κατευθύνονται κατά μήκος του άξονα *y* προς τα πάνω όταν η συσκευή βρίσκεται σε κανονική θέση λειτουργίας και έχουν ένα οπτικό πεδίο περίπου 150 μοιρών. Ο ίδιος ο αισθητήρας αποκτά πρόσβαση και προγραμματίζεται μέσω APIs, με την υποστήριξη μιας ποικιλίας γλωσσών προγραμματισμού, όπως η C++ και η Python. Οι θέσεις των αναγνωρισμένων

αντικειμένων αποκτώνται μέσω αυτών των API. Τα καρτεσιανά και σφαιρικά συστήματα συντεταγμένων που χρησιμοποιούνται για την περιγραφή των θέσεων στο αισθητήριο χώρο περιγράφονται παρακάτω. Ωστόσο, πρέπει να σημειωθεί ότι η συχνότητα δειγματοληψίας δεν είναι σταθερή, δεν μπορεί να ρυθμιστεί και ποικίλλει σημαντικά.



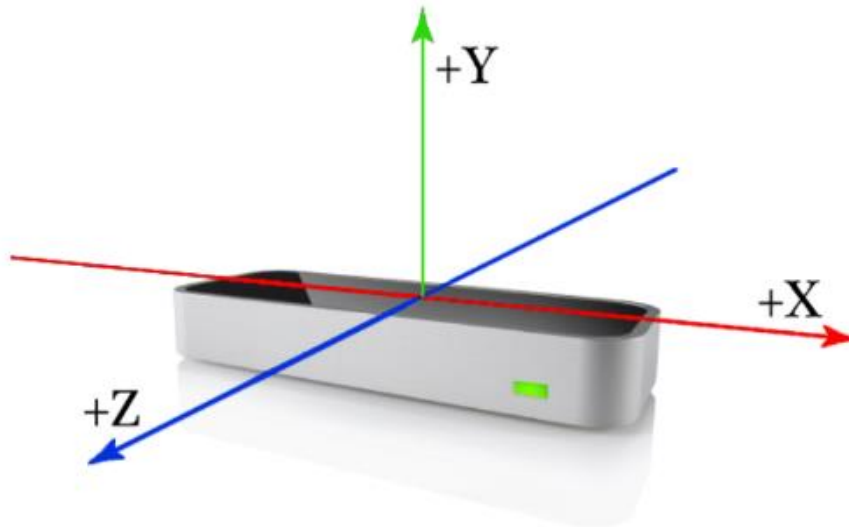
Interaction Area

2 feet above the controller, by 2 feet wide on each side (150° angle), by 2 feet deep on each side (120° angle)

Εικόνα 2.7: Το πεδίο αλληλεπίδρασης της συσκευής Leap Motion

Σύστημα Συντεταγμένων

Η τεχνολογία Leap Motion χρησιμοποιεί ένα καρτεσιανό σύστημα συντεταγμένων. Η προέλευση επικεντρώνεται στην κορυφή του Leap Motion Controller. Οι άξονες x και z βρίσκονται στο οριζόντιο επίπεδο, με τον άξονα x παράλληλο προς τη μεγάλη πλευρά της συσκευής. Ο άξονας y είναι κάθετος, με θετικές τιμές αυξανόμενες προς τα πάνω. Το Leap Motion Controller παρέχει συντεταγμένες σε χιλιοστά (mm). Δηλαδή, εάν η θέση ενός άκρου του δακτύλου δίνεται ως $(x, y, z) = [100, 100, -100]$, οι αριθμοί αυτοί είναι x= 10 εκατοστά, y = 10 εκατοστά, z = -10 εκατοστά.



Εικόνα 2.8: Το σύστημα συντεταγμένων της συσκευής Leap Motion

Μοντέλο χεριού

Το μοντέλο χεριού παρέχει πληροφορίες σχετικά με την ταυτότητα, τη θέση και τα άλλα χαρακτηριστικά του χεριού κατά την ανίχνευση της χειρονομίας. Το λογισμικό Leap Motion παρέχει την παρακολούθηση, ακόμη και όταν τα μέρη του χεριού δεν είναι ορατά. Το μοντέλο χεριού παρέχει πάντα θέσεις για πέντε δάχτυλα, η παρακολούθηση είναι η βέλτιστη και όλα τα δάχτυλά είναι ορατά. Καλό θα ήταν να σημειωθεί εδώ ότι οι λεπτές κινήσεις των δακτύλων δεν είναι συνήθως ανιχνεύσιμες από τους αισθητήρες Leap Motion. Επίσης συνιστάται η διατήρηση το πολύ δύο χεριών στο πεδίο Leap Motion για βέλτιστη ποιότητα παρακολούθησης κίνησης.

Μοντέλο δακτύλων

Ο αισθητήρας Leap Motion παρέχει πληροφορίες για κάθε δάκτυλο σε ένα χέρι. Εάν το σύνολο ή μέρος των δακτύλων δεν είναι ορατά, τότε τα χαρακτηριστικά των δακτύλων υπολογίζονται με βάση τις πιο πρόσφατες παρατηρήσεις. Τα δάχτυλα προσδιορίζονται με βάση το όνομα του τύπου, δηλαδή αντίχειρας, δείκτης, μεσαίος, παράμεσος και μικρός.

Σύστημα οστών των δακτύλων

Για κάθε δάχτυλο αντιστοιχεί το σύστημα οστών με τα παρακάτω στοιχεία:

- Μετακάρπιο (Metacarpal) το οστό στο εσωτερικό του χεριού που συνδέει το δάχτυλο στον καρπό (εκτός του αντίχειρα)
- Εγγύς φάλαγγα (Proximal Phalanges) το οστό στη βάση του δακτύλου, που συνδέεται με την παλάμη
- Ενδιάμεση Φάλαγγα (Intermediate Phalanges) το μεσαίο οστό του δακτύλου μεταξύ της κορυφής και της βάσης
- Άπω Φάλαγγα (Distal Phalanges) το άκρο του δακτύλου

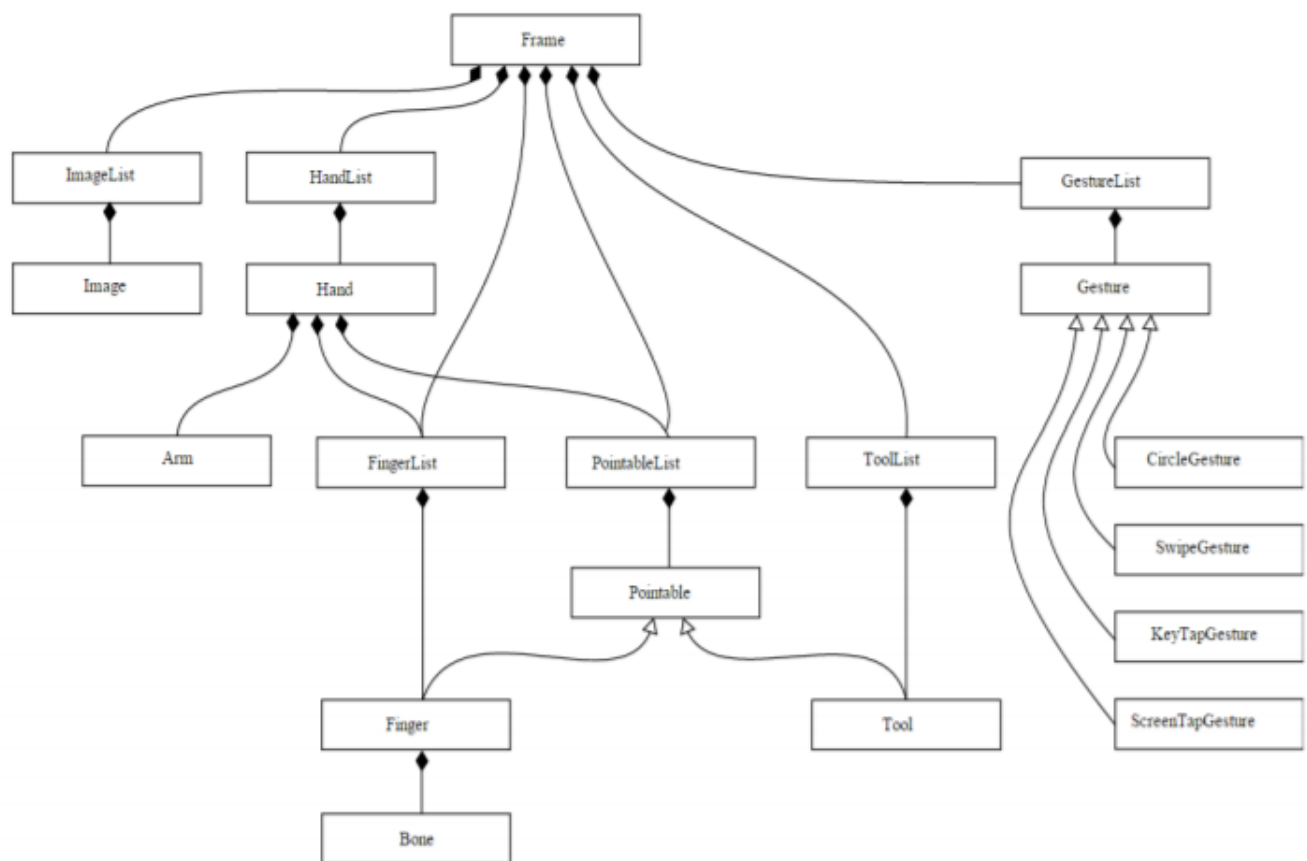
Στην πραγματικότητα ο αντίχειρας έχει ένα λιγότερο οστό από τα άλλα δάχτυλα (δεν έχει μετακάρπιο). Ωστόσο, για την ευκολία του προγραμματισμού, το μοντέλο δακτύλων για τον αντίχειρα περιλαμβάνει ένα μηδενικού μήκους μετακάρπιο οστό, έτσι ώστε ο αντίχειρας να έχει τον ίδιο αριθμό οστών με τα υπόλοιπα δάχτυλα.

2.2.2 Software

Όπως αναλύθηκε και παραπάνω, το Application Programming Interface (API) αναφέρεται στο σύνολο των λειτουργιών που παρέχει μία βιβλιοθήκη προγραμματισμού για να προσφέρει την επικοινωνία με το λειτουργικό σύστημα ή πρόγραμμα ελέγχου ή συσκευή, χωρίς όμως να επιτρέπει την πρόσβαση στον κώδικα. Η τεχνολογία του Leap Motion παρέχει ένα API το οποίο είναι το Leap SDK Api. Η τελευταία έκδοση του Leap Api εισάγει ένα νέο σκελετικό μοντέλο παρακολούθησης που παρέχει πρόσθετες πληροφορίες σχετικά με τα χέρια και τα δάχτυλα, καθώς επίσης βελτιώνει τα δεδομένα παρακολούθησης. Με την μοντελοποίηση ενός ανθρώπινου χεριού το λογισμικό του Leap Motion μπορεί να προβλέψει και να εστιάσει καλύτερα στην θέση των δακτύλων. Κατά τη διάρκεια της ανίχνευσης του χεριού υπάρχουν πάντα πέντε δάχτυλα τα οποία μπορεί να είναι είτε εκτεταμένα (τεντωμένα) είτε όχι. Επιπλέον τα δάχτυλα κατά την παρακολούθησή τους μπορεί να διασταυρωθούν, γι' αυτό το λόγο πρέπει να δοθεί ιδιαίτερη προσοχή προκειμένου να εντοπιστεί με ακρίβεια η θέση του δακτύλου. Η Leap SDK Api περιλαμβάνει:

- Αναγνώριση δεξιού ή αριστερού χεριού
- Αναφορά της θέσης και του προσανατολισμού του κάθε οστού του δακτύλου
- Αναφορά των παραγόντων εάν ένας χρήστης επιδιώξει μια λαβή, ένα πιάσιμο
- Αναφορά της θέσης των πέντε δακτύλων σε κάθε χέρι

Επιπλέον το Leap SDK χρησιμοποιεί κλάσεις οι οποίες αντιπροσωπεύουν τα αντικείμενα που τις ακολουθούν, όπως είναι οι χειρονομίες, το χέρι και τα δάχτυλα του χεριού. Στην Εικόνα 2.9 που φαίνεται παρακάτω έχουμε ένα παράδειγμα της ιεραρχίας των κλάσεων του API της συσκευής.



Εικόνα 2.9: Ιεραρχία των κλάσεων του API της συσκευής

2.3 Σύγκριση Microsoft Kinect - Leap Motion

Μελετώντας τις δύο τεχνολογίες του Microsoft Kinect και του Leap Motion, τα χαρακτηριστικά, τις βιβλιοθήκες και τις δυνατότητες που παρέχονται μπορούν να εξαχθούν στα εξής συμπεράσματα. Το βασικό χαρακτηριστικό και η βασική διαφορά του Microsoft Kinect από το Leap Motion είναι ότι το Microsoft Kinect κάνει ανίχνευση και αναγνωρίζει κίνηση και χειρονομίες σκελετού, αλλά όχι δακτύλων. Αναγνωρίζει την έκφραση του προσώπου και ανιχνεύει την κίνηση και τη στάση του σώματος. Καμία βιβλιοθήκη όμως δεν αναγνωρίζει και δεν ξεχωρίζει την κίνηση στα δάκτυλα, η οποία είναι σημαντική καθώς και καθοριστική παράμετρος για την υλοποίηση της εφαρμογής που ακολουθεί. Από την άλλη πλευρά το Leap Motion κάνει μόνο αναγνώριση χειρονομιών των δακτύλων και της παλάμης. Το Leap Motion εντοπίζει την κίνηση των δακτύλων με ακρίβεια 0.01mm.

Ο προγραμματισμός του Microsoft Kinect είναι πολύ πιο δύσκολος διότι υπάρχουν πολλές βιβλιοθήκες, ανιχνεύει πολλές κινήσεις, ολόκληρο τον ανθρώπινο σκελετό και διαχειρίζεται δεδομένα από διαφορετικές πηγές καμερών. Το γεγονός ότι ανιχνεύει ολόκληρο τον ανθρώπινο σκελετό το καθιστά πιο πολύπλοκο καθώς πρέπει να απομονωθεί η επιθυμητή κίνηση κάθε φορά. Το Leap Motion είναι μία μικρή και εύχρηστη συσκευή σε σχέση πάντα με το Microsoft Kinect και συνδέεται στον υπολογιστή με τη χρήση ενός USB. Επιπλέον μεγάλη διαφορά υπάρχει και στο κόστος των δύο τεχνολογιών καθώς το Leap Motion έχει δύο φορές χαμηλότερο κοστολόγιο από το Microsoft Kinect.

Το μικρότερο μέγεθος της συσκευής Leap Motion και η μεγαλύτερη ανάλυση είναι αυτά που την διαφοροποιούν από το Microsoft Kinect. Το Kinect θεωρείται καλύτερο για παρακολούθηση ολόκληρου του σώματος σε μεγαλύτερο χώρο ενώ το Leap Motion controller είναι πολύ αποτελεσματικό στην παρακολούθηση των κινήσεων των χεριών, ακόμα και των δακτύλων. Η ανίχνευση κινήσεων γίνεται με ακρίβεια εκατοστού του χιλιοστού. Η συσκευή είναι συμβατή με τα λειτουργικά συστήματα Windows και OSX.

ΚΕΦΑΛΑΙΟ 3

3. Ανασκόπηση Βιβλιογραφίας (State of art)

3.1 Τεχνολογία Leap motion

Η τεχνολογία Leap motion είναι μια σχετικά νέα τεχνολογία η οποία αναπτύχθηκε αρχικά για να αντικαταστήσει το υπολογιστικό ποντίκι και το πληκτρολόγιο κάτι το οποίο εγκαταλείφθηκε στην πορεία. Μετά την ανάλυση των επιδόσεων του Leap Motion, έγινε η σύγκρισή του με μια κοινή συσκευή ποντικιού και τα αποτελέσματα δε ήταν ευνοϊκά για τον αισθητήρα. Για την αξιολόγηση χρησιμοποιήθηκε ο νόμος του Fitts, ο οποίος είναι ένα μοντέλο πρόβλεψης της ανθρώπινης κίνησης που χρησιμοποιείται κυρίως στην αλληλεπίδραση ανθρώπου-υπολογιστή και στην εργονομία. Οι επιδόσεις του αισθητήρα σε γενικές εργασίες στον υπολογιστή, όπως το άνοιγμα ενός παραθύρου, ήταν περιορισμένες σε σύγκριση με μια συσκευή ποντικιού δεδομένου του ποσοστού σφάλματος το οποίο ανερχόταν σε 7,8% για τον αισθητήρα και 2,8% για τη συσκευή ποντικιού (Bachmann et al. 2015).

Τόσο το γεγονός ότι το Leap Motion είναι μια νέα τεχνολογία όσο και το ότι ο αρχικός σκοπός δεν ευδοκίμησε είναι οι λόγοι που δεν είναι τόσο διαδεδομένη και δεν έχει δημοσιευθεί ακόμη ικανοποιητικό επιστημονικό έργο γύρω από τις δυνατότητες, την ακρίβεια και την αποτελεσματικότητά του. Όμως αποτελεί μια συσκευή, έναν αισθητήρα φυσικής αλληλεπίδρασης χρήστη που υποστηρίζει την είσοδο δεδομένων στον υπολογιστή χωρίς άγγιγμα ή επαφή, δηλαδή απελευθερώνει τα χέρια του χρήστη. Συνεπώς η τεχνολογία του μπορεί να εφαρμοστεί σε διάφορους τομείς. Πέρα από την βιομηχανία του gaming μπορεί να εφαρμοστεί στην χειρουργική, στην αρχιτεκτονική, στην μηχανική και στον σχεδιασμό. Παρακάτω παρατίθενται κάποιες εφαρμογές που χρησιμοποίησαν αυτήν την τεχνολογία καθώς και η αξιολόγησή τους.

Μια εφαρμογή που χρησιμοποιήθηκε ο αισθητήρας Leap Motion Controller ήταν με σκοπό τον άμεσο χειρισμό ενός βιομηχανικού βραχίονα ρομπότ με έξι βαθμούς ελευθερίας. Η συσκευή χρησιμοποιήθηκε για την καταγραφή της θέσης των δακτύλων. Για να αυξηθεί η ακρίβεια της καταγραφής, πραγματοποιήθηκε παρεμβολή των αποκτηθέντων δεδομένων χρησιμοποιώντας πολυωνυμικές καμπύλες (spline interpolation). Σκοπός της έρευνας ήταν η αναπαραγωγή σύνθετων τρισδιάστατων εργασιών χωρίς περιορισμούς στο χειριστή. Αυτός ο στόχος αντικατοπτρίζει τη σημασία των διεπαφών που είναι βασισμένες στις χειρονομίες και χρησιμοποιούν συσκευές χαμηλού κόστους και αισθητήρα εισόδου για βιομηχανική χρήση. (Hernoux et al. 2013)

Μια άλλη εφαρμογή του Leap Motion Controller δείχνει τις δυνατότητές του σε εφαρμογές αναγνώρισης χειρονομιών και χειρογράφου. Τα αποκτηθέντα δεδομένα εισόδου αντιμετωπίζονται ως χρονολογική σειρά τρισδιάστατων θέσεων και επεξεργάζονται χρησιμοποιώντας τον αλγόριθμο Dynamic Time Warping algorithm. Τα αποτελέσματα

των πειραμάτων δείχνουν πολλά υποσχόμενη ακρίβεια αναγνώρισης και αποτελέσματα απόδοσης. (Regenbrecht et al. 2013)

Παραμένοντας στο ίδιο αντικείμενο, αλλά χρησιμοποιώντας όχι μόνο συσκευές Leap Motion Controller αλλά και αισθητήρες Kinect, πραγματοποιείται μια εφαρμογή για την ακριβέστερη αναγνώριση της κίνησης. Τα χαρακτηριστικά στοιχεία από τον χάρτη βάθους του αισθητήρα Kinect συνδυάζονται με τις δυνατότητες του Leap Motion Controller για βελτίωση της αναγνώρισης των χειρονομιών. Τα πειραματικά δεδομένα τα οποία εστιάζουν μόνο σε στατικές χειρονομίες και όχι σε δυναμικές καταλήγουν σε αρκετά μεγάλη ακρίβεια σε πραγματικό χρόνο. (Marin et al. 2015)

Επιπλέον, παρουσιάζεται μια νέα προσέγγιση διεπαφής που συνδυάζει την εικονική πραγματικότητα με ένα μερικό μοντέλο voxel για να επιτρέψει πιο πειστικές αλληλεπιδράσεις με τρισδιάστατα αντικείμενα. Η διασύνδεση επιτρέπει στους χρήστες να αλληλοεπιδρούν με ένα εικονικό περιβάλλον μέσω της διεπαφής που ελέγχεται από τις κινήσεις των χεριών και να επιτρέψει σωστές αλληλεπικαλύψεις μεταξύ των χεριών, των δακτύλων και του εικονικού περιβάλλοντος. Ο συνδυασμός του Leap Motion Controller και μιας κάμερας web χρησιμοποιείται για την παρακολούθηση των δακτύλων των χρηστών και την επικάλυψη του κατάλληλου βίντεο για μια επαυξημένη προβολή. (Regenbrecht et al. 2013)

Ακόμη έχει χρησιμοποιηθεί σε ένα project για την αποκατάσταση ατόμων μετά από εγκεφαλικό επεισόδιο, έτσι ώστε να ελεγχθεί η μνήμη τους μέσα από εικονικά περιβάλλοντα. Σε ένα τμήμα της εφαρμογής διερευνήθηκε αν η συσκευή μπορεί να χρησιμοποιηθεί σε περιβάλλοντα εικονικής πραγματικότητας και τα αποτελέσματα ήταν ενθαρρυντικά. (Bracegirdle et al. 2014)

Πέραν των προαναφερθεισών εφαρμογών ο αισθητήρας Leap Motion διαθέτει ένα ευρύ φάσμα εμπορεύσιμων εφαρμογών. Διαθέτει ένα App Store όπου πωλούνται εφαρμογές από προγραμματιστές. Τον Νοέμβριο του 2013 είχε αναφερθεί ότι οι εφαρμογές είχαν φτάσει τις 150. Το Leap Motion App Store μπορεί να χρησιμοποιηθεί για περιήγηση και αποθήκευση εφαρμογών, ιστοσελίδων ή άλλων προϊόντων για χρήση με το Leap Motion Controller και τον υπολογιστή. Το App Store παρέχεται ως υπηρεσία για τους χρήστες ώστε να ανακαλύψουν εύκολα τις υπηρεσίες, τις δυνατότητες αλλά και το περιεχόμενο του Leap Motion. Μερικές από τις εφαρμογές αυτές μπορεί να προσφέρει η ίδια η εταιρία Leap Motion, ενώ άλλες μπορεί να διατίθενται από τρίτους που δεν ελέγχονται από αυτήν.

Η τεχνολογία του Leap Motion παρέμεινε στη διάθεση των προγραμματιστών δίνοντάς τους ένα μεγάλο κίνητρο καθώς και τον εξοπλισμό. Μερικοί ακόμα από τους κλάδους που χρησιμοποιούν την τεχνολογία του Leap Motion είναι ο κλάδος της Τέχνης (Art), Τεχνολογίας Ηρ, Εκπαίδευσης (Education), Ιατρικής (Medical).

Η εφαρμογή Sculpting του App Store του Leap Motion είναι μία σύνδεση της τέχνης της γλυπτικής με την τεχνολογία. Σε αυτήν την εφαρμογή ο δημιουργός έχει την δυνατότητα να αναπτύξει τις ιδέες του σε ένα τρισδιάστατο περιβάλλον. Έχει στην διάθεσή του μία μεγάλη ποικιλία βοηθητικών εργαλείων που του επιτρέπουν να διαμορφώσει με ακρίβεια την δημιουργία του σε ένα ευχάριστο και εύκολο περιβάλλον. Έχει την δυνατότητα να επιλέξει να επεξεργαστεί πολλά υλικά στα οποία συμπεριλαμβάνονται το γυαλί, το πλαστικό και ο πηλός.

Η εφαρμογή Cordion Conductor είναι άμεσα συνδεδεμένη με την μουσική. Διαθέτει ένα περιβάλλον προσβάσιμο από όλους, όπου εκεί υπάρχει ένα ευρύ φάσμα από νότες και κλίμακες. Σε αυτή την εφαρμογή ο δημιουργός μπορεί να συνθέσει την μελωδία και τους ήχους που θέλει. Χρησιμοποιεί διάφορα μέσα για την λήψη μελωδίας για την ενσωμάτωση ήχου καθώς και τον έλεγχο της έντασης του ήχου.

Η τεχνολογία HP είναι μια τεχνολογία ελέγχου κίνησης και χειρονομίας από το Leap Motion όπου είναι ενσωματωμένη σε πληκτρολόγιο ακόμα και στο πληκτρολόγιο ενός φορητού υπολογιστή. Αυτή η τεχνολογία δίνει στον χρήστη τη δυνατότητα να χειρίζεται τις δυνατότητες του υπολογιστή του χωρίς τη χρήση ποντικιού, απλά με τις κινήσεις των δακτύλων όπως ακριβώς και στον φυσικό κόσμο. Αυτή η αλληλεπίδραση γίνεται με το κύμα του χεριού ή την ανύψωση ενός δακτύλου.

Η εφαρμογή 3D Geometry είναι άμεσα συνδεδεμένη με τον κλάδο των μαθηματικών και της γεωμετρίας. Αυτή η εφαρμογή δίνει την δυνατότητα στο χρήστη να σχεδιάσει με ακρίβεια, να μετρήσει και να κατασκευάσει διάφορα γεωμετρικά σχήματα σε ένα τρισδιάστατο περιβάλλον, καθώς υπάρχουν πολλοί τύποι πρισμάτων και πυραμίδες. Επιπλέον μπορεί να ρυθμίσει την περιστροφή, τον στροβιλισμό ή και το ίσιωμα αυτών των σχημάτων.

Η εφαρμογή Molecules είναι μια εφαρμογή που σχετίζεται με τη βιολογία και τη μοριακή οπτικοποίηση. Επιτρέπει στον χρήστη να κάνει προβολή αρχείων σε τρισδιάστατη απεικόνιση. Είναι δυνατή η απεικόνιση, η περιστροφή και η αλλαγή της κλίμακας της μοριακής δομής σε πολύ γρήγορο χρόνο. (Κωτσίδου, 2015)

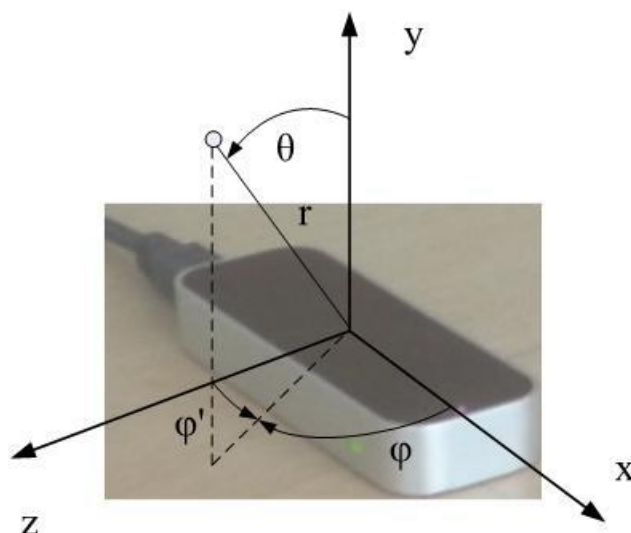
Εκτός από τις εφαρμογές που χρησιμοποιήθηκε ως εργαλείο για την εξαγωγή ενός αποτελέσματος, έχουν πραγματοποιηθεί και κάποιες μελέτες όσον αφορά στην ακρίβεια και στην ανθεκτικότητα του Leap Motion Controller. Παρακάτω παρατίθενται δύο από τις μελέτες αυτές.

Για το πρώτο πείραμα χρησιμοποιήθηκε μία αιχμηρή πένα τοποθετημένη στο ρομποτικό βραχίονα που επέτρεπε την ακρίβεια της θέσης. Τα αποτελέσματα δείχνουν απόκλιση μεταξύ της επιθυμητής θέσης και των μέσων μετρημένων θέσεων κάτω από 0,2 mm για στατικές ρυθμίσεις και 1,2 mm για δυναμικές ρυθμίσεις. Λαμβάνοντας υπόψη ότι η

πλειονότητα των εφαρμογών για τον αισθητήρα Leap Motion είναι διεπαφές χρηστών βασισμένες στις χειρονομίες, η επιτεύξιμη ακρίβεια της μέτρησης της κίνησης ενός ανθρώπινου χεριού είναι ο πιο σημαντικός παράγοντας, ο οποίος επηρεάζεται ουσιαστικά από το λεγόμενο τρέμουλο (tremor). Τρέμουλο ορίζεται η ακούσια και περίπου ρυθμική κίνηση των μυών. Ανάλογα με την ηλικία, το πλάτος του ποικίλλει μεταξύ $0,4\text{mm} \pm 0,2\text{mm}$ για μικρότερες ηλικίες και $1,1\text{mm} \pm 0,6\text{mm}$ για μεγαλύτερες ηλικίες (Sturman et al. 2005). Επομένως, για να αξιολογηθεί ο αισθητήρας Leap Motion σε σχέση με τις διεπαφές χρήστη που βασίζονται σε ανθρώπινη χειρονομία, πρέπει να καθοριστεί ένα σύστημα αναφοράς με ακρίβεια καλύτερη από το τρέμουλο, δηλαδή $0,2\text{mm}$. Για να ικανοποιηθούν αυτές οι απαιτήσεις, επελέγη ένα βιομηχανικό ρομπότ, Kuka Robot KR 125/3, το οποίο παρέχει επαναληπτική ακρίβεια μικρότερη από $0,2\text{mm}$. Η εγκατάσταση (setup) αποτελείται από τον αισθητήρα Motion Leap και ένα βιομηχανικό ρομπότ. (Weichert et al. 2013)

Σε ένα δεύτερο πείραμα η απόδοση του αισθητήρα αξιολογήθηκε μέσω δύο τύπων μετρήσεων, μια σειρά στατικών και μια σειρά δυναμικών μετρήσεων. Μετά το πειραματικό στάδιο, περιγράφηκε η αξιολόγηση της απόδοσης του Leap Motion Controller με τη βοήθεια ενός επαγγελματικού συστήματος καταγραφής της κίνησης ταχείας και υψηλής ακρίβειας. Ο κύριος στόχος της έρευνας ήταν να αναλυθεί συστηματικά το πεδίο του αισθητήρα και να καθοριστεί η εξάρτηση της ακρίβειας και της αξιοπιστίας με το χώρο.

Κατά την πρώτη μέτρηση, το στατικό σενάριο, εντοπίστηκαν και καταγράφηκαν μια σειρά σταθερών στατικών σημείων στο χώρο για μεγάλο χρονικό διάστημα για την αξιολόγηση της συνέπειας και της διασποράς των αποτελεσμάτων. Οι συντεταγμένες των σημείων ήταν επιλεγμένες με τέτοιο τρόπο ώστε να καλυφθεί το μεγαλύτερο τμήμα του πεδίου του αισθητήρα. Η τυπική απόκλιση στις περισσότερες χρονικές στιγμές ήταν λιγότερο από $0,5\text{ mm}$, στις καλύτερες περιπτώσεις μικρότερη από $0,01\text{ mm}$ δείχνοντας υψηλή ακρίβεια (κάτω από $0,2\text{ mm}$), όπως και στο πρώτο πείραμα και αξιολογώντας τον αισθητήρα ως αξιόπιστο και ακριβές σύστημα για την παρακολούθηση στατικών σημείων. Η ανάλυση αποκάλυψε μια σημαντική χωρική εξάρτηση της συνέπειας και της απόδοσης του αισθητήρα. Η γραμμική συσχέτιση έδειξε μια σημαντική αύξηση της τυπικής απόκλισης όταν απομακρύνεται από τον αισθητήρα (ακτίνα) και όταν μετακινείται προς τα άκρα αριστερά ή δεξιά του αισθητήρα (φ'). Η γωνία αυτή φαίνεται στην Εικόνα 3.1.



Εικόνα 3.1: Γωνία που σχηματίζεται μεταξύ των αξόνων του συστήματος συντεταγμένων και του σημείου καταγραφής

Το κύριο κριτήριο για την επιλογή των τελικών 37 χωρικών θέσεων ήταν επομένως η δημιουργία μιας σταθερής θέσης για το βραχίονα παρακολούθησης που επέτρεψε την επιτυχή παρακολούθηση και καταγραφή για μεγαλύτερο χρονικό διάστημα. Σε πολλές περιπτώσεις, δεν μπόρεσε να δημιουργηθεί ένα σταθερό περιβάλλον και ο αισθητήρας κατέγραφε μόνο τα στατικά σημεία για λίγα δευτερόλεπτα και στη συνέχεια σταματούσε. Η πλειονότητα των επιτυχώς επιλεγμένων σημείων βρισκόταν πίσω από τον αισθητήρα ($z < 0$), όταν το χέρι βρισκόταν πάνω από αυτόν και επομένως ήταν πλήρως ορατό μέσα στο πεδίο του αισθητήρα.

Κατά τη δεύτερη μέτρηση, παρέχεται μια σταθερή απόσταση μεταξύ δύο αντικειμένων, τα οποία κατόπιν μετακινήθηκαν ελεύθερα σε όλο το πεδίο του αισθητήρα. Η ακρίβεια καταγραφής του αισθητήρα αξιολογήθηκε με βάση την παραμόρφωση της απόστασης μεταξύ των δύο αντικειμένων. Το σύστημα αναφοράς χρησιμοποιήθηκε για τον προσδιορισμό των ακριβών χωρικών θέσεων των εντοπισμένων αντικειμένων και τις αποστάσεις μεταξύ τους. Το σύνολο των μετρήσεων στο δυναμικό σενάριο αποκάλυψε την ασυνεπή απόδοση του αισθητήρα. Η ακρίβειά του αξιολογήθηκε μέσω της παραμόρφωσης της απόστασης μεταξύ δύο κινούμενων σημείων. Σε αυτή την περίπτωση, η ακρίβεια μειώνεται όταν τα αντικείμενα απομακρυνθούν από τον αισθητήρα. Υπάρχει σημαντική μείωση της ακρίβειας των δειγμάτων που έχουν ληφθεί σε μεγαλύτερη απόσταση από 250 mm πάνω από τη συσκευή. Λόγω αυτού του ενδιαφέροντος και απροσδόκητου φαινομένου, επαναλήφθηκε η μέτρηση για αυτήν την περιοχή και τα αποτελέσματα ήταν τα ίδια. Λόγω αδυναμίας εύρεσης της κύριας αιτίας αυτής της συμπεριφοράς, έγινε η υπόθεση ότι η χρήση αντικειμένων με διαφορετικές αποστάσεις μεταξύ τους θα αποκάλυπτε διαφορετικά αποτελέσματα.

Ένας σημαντικός περιορισμός της απόδοσης του αισθητήρα είναι η ασυνεπής συχνότητα δειγματοληψίας του. Η μέση τιμή της, κάτω από 40 Hz, είναι σχετικά χαμηλή και ποικίλλει σημαντικά τόσο υπό στατικές όσο και υπό δυναμικές συνθήκες. Το κύριο μειονέκτημα αυτής της μη ομοιόμορφης δειγματοληψίας είναι η μεγάλη δυσκολία συγχρονισμού του αισθητήρα με άλλα συστήματα πραγματικού χρόνου, δεδομένου ότι απαιτεί δύσκολες μεθόδους επεξεργασίας και συχνά επανάληψη της δειγματοληψίας. (Guna et al. 2014).

Ο αισθητήρας Leap Motion είναι αναμφισβήτητα μια επαναστατική συσκευή εισόδου για αλληλεπίδραση ανθρώπου-ηλεκτρονικού υπολογιστή με χειρονομίες. Σε αυτό το δεύτερο πείραμα που περιγράφηκε, αξιολογήθηκε ως πιθανή αντικατάσταση ενός συστήματος οπτικής κίνησης γρήγορης και υψηλής ακρίβειας σε περιορισμένο χώρο και με περιορισμένο αριθμό αντικειμένων. Με βάση τα τρέχοντα αποτελέσματα και τη συνολική εμπειρία, συμπεραίνεται ότι ο αισθητήρας στην τρέχουσα κατάστασή του δεν θα μπορούσε να χρησιμοποιηθεί ως επαγγελματικό σύστημα παρακολούθησης, κυρίως λόγω του μάλλον περιορισμένου πεδίου καταγραφής αλλά και της ασυνεπής συχνότητας δειγματοληψίας.

Συνεπώς, με βάση τις εκτιμήσεις που αποκτήθηκαν από αυτά τα πειράματα, η περαιτέρω μελέτη του Leap Motion Controller μπορεί να περιλαμβάνει έρευνα σχετικά με την ακρίβεια και την αξιοπιστία της παρακολούθησης πιο πολύπλοκων κινήσεων χειρός, δακτύλων και εργαλείων καθώς και της καταλληλότητάς της για εφαρμογές που εξαρτώνται έντονα από τη μέθοδο εισόδου χειρονομίας.

Λαμβάνοντας υπόψη τις παραπάνω μελέτες θα γίνει και η αξιολόγηση της παρούσας μελέτης που αφορά στην καταγραφή χειρονομιών με δύο συσκευές Leap Motion.

3.2 Τεχνολογία Kinect

Πέραν όμως της συσκευής Leap Motion, αξιοσημείωτα είναι και τα αποτελέσματα ερευνών που χρησιμοποίησαν τη συσκευή Kinect. Ο αισθητήρας αυτός θα μπορούσε να χρησιμοποιηθεί και να αναλυθεί και στην παρούσα μελέτη αλλά λόγω προβλημάτων έλλειψης χώρου και πίεσης χρόνου διερευνήθηκε μόνο το σύστημα του Leap Motion. Μία έρευνα που διαφωτίζει την αρχική επιλογή των συστημάτων καταγραφής κίνησης (motion capture system) είναι αυτή που επικεντρώθηκε στο προσαρμοστικό σύστημα αναγνώρισης χειρονομιών, αναπτύσσοντας ταυτόχρονα μια βάση δεδομένων χειρονομίας για την εξάλειψη των μεμονωμένων παραγόντων που επηρεάζουν την αποτελεσματικότητα του συστήματος αναγνώρισης. Για την απόκτηση δεδομένων εισόδου για το πείραμα χρησιμοποιήθηκε το Qualisys™ Motion Capture System (Aziz et al. 2012).

Εξίσου διαφωτιστική είναι και η μελέτη που αφορά στην λεπτομερή ανάλυση της ακρίβειας και της ανάλυσης των δεδομένων βάθους του αισθητήρα Kinect για εφαρμογές χαρτογράφησης εσωτερικών χώρων. Τα πειραματικά αποτελέσματα δείχνουν ότι το τυχαίο σφάλμα στη μέτρηση βάθους αυξάνεται με την αυξανόμενη απόσταση προς τον αισθητήρα, που κυμαίνεται από μερικά χιλιοστά έως περίπου 4 cm στο μέγιστο εύρος του αισθητήρα. Η ποιότητα των δεδομένων διαπιστώθηκε επίσης ότι επηρεάζεται από τη χαμηλή ανάλυση των μετρήσεων βάθους. Η ακρίβεια που επιτυγχάνεται είναι γενικά επαρκής για την ανίχνευση χειρονομίας του βραχίονα αλλά και ολόκληρου σώματος. Όμως, δεν αρκεί για ακριβείς χειρονομίες δακτύλων, όπως η συγγραφή ενός χειρόγραφου (Khoshelham et al. 2012).

Φυσικά η συσκευή Kinect έχει χρησιμοποιηθεί σε διάφορους κλάδους, στην ιατρική, στην εκπαίδευση, αλλά και στο χώρο της εικονικής πραγματικότητας. Παρακάτω αναφέρονται χαρακτηριστικά δύο από αυτές τις εφαρμογές. Ερευνητές από το Πανεπιστήμιο του Ηνωμένου Βασιλείου, χρησιμοποιούν την τεχνολογία για να βοηθήσουν ασθενείς που αναρρώνουν μετά από εγκεφαλικό επεισόδιο. Έχει σχεδιαστεί ένας ειδικός αλγόριθμος, ο οποίος επιτρέπει στους θεραπευτές να παρακολουθούν από απόσταση τις κινήσεις των χεριών και των δακτύλων των ασθενών. Με αυτό τον τρόπο μπορούν να τους καθοδηγήσουν στην εκτέλεση διαφόρων ασκήσεων που περιλαμβάνονται σε ένα ευρύτερο πρόγραμμα φυσικοθεραπείας.

Μια άλλη εφαρμογή χρησιμοποιεί τα δεδομένα βάθους του Kinect για να δημιουργήσει ένα σύννεφο από σημεία (point cloud) που θα αντιπροσωπεύει την συνολική επιφάνεια ενός αντικείμενου στο τρισδιάστατο χώρο. Για να γίνει αυτό, τα point clouds συγκρίνονται, ενώνονται και τέλος φιλτράρονται για να την ανακατασκευή της επιφάνειας του τρισδιάστατου αντικείμενου. Σκοπός, η δημιουργία ενός προστιτού τρισδιάστατου σαρωτή (3D-scanner) χρησιμοποιώντας ελεύθερο λογισμικό. (Κωτσίδου, 2015)

Έχοντας ανατρέξει σε κάποιες από τις εφαρμογές που έχουν χρησιμοποιηθεί οι συσκευές Kinect και Leap Motion, παρατηρείται ότι κυρίως η συσκευή Kinect επιλέγεται για την καταγραφή της κίνησης. Οι περισσότερες μελέτες αφορούν στον τομέα της ψυχαγωγίας, όπως η ανάπτυξη διαδραστικών εφαρμογών και παιχνιδιών. Συνεπώς, παραμένει ανοιχτό το πεδίο για τη δημιουργία συστήματος καταγραφής κίνησης στον τομέα του πολιτισμού. Μελλοντικές εφαρμογές θα μπορούσαν να αφορούν στην καταγραφή του χορού και της μουσικής. Στο επόμενο κεφάλαιο περιγράφεται η πρακτική εφαρμογή η οποία καταγράφει χειρονομίες με κύριο σκοπό την κατανόηση των εξαγόμενων στοιχείων της συσκευής Leap Motion έτσι ώστε να γίνει για παράδειγμα ευκολότερη η μελλοντική αυτόματη απόδοση της νοηματικής γλώσσας. Πριν την ανάλυση και την αξιολόγηση της εφαρμογής κρίνεται απαραίτητη η επεξήγηση της έννοιας «καταγραφή κίνησης» καθώς και οι μέθοδοι και τα συστήματα που χρησιμοποιούνται.

Motion capture

Η καταγραφή κίνησης (Mo-cap) είναι η διαδικασία καταγραφής της κίνησης αντικείμενων ή ανθρώπων. Χρησιμοποιείται σε στρατιωτικές, ψυχαγωγικές, αθλητικές, ιατρικές εφαρμογές και για την αξιολόγηση της όρασης υπολογιστών και της ρομποτικής. Στην κινηματογραφική παραγωγή και την ανάπτυξη βιντεοπαιχνιδιών, χρησιμοποιείται για την καταγραφή των κινήσεων των ηθοποιών με τις εξαγόμενες πληροφορίες να ζωντανεύουν τα μοντέλα ψηφιακών χαρακτήρων σε κινούμενες εικόνες 2D ή 3D. Όταν η διαδικασία περιλαμβάνει πρόσωπο και δάκτυλα ή καταγράφει λεπτές εκφράσεις, συχνά αναφέρεται ως καταγραφή ενέργειας (performance capture). Σε πολλούς τομείς, η σύλληψη κίνησης ονομάζεται μερικές φορές παρακολούθηση κίνησης (motion tracking), αλλά στις κινηματογραφικές ταινίες και τα παιχνίδια, η παρακολούθηση της κίνησης συνήθως αναφέρεται περισσότερο στην αντιστοίχιση της κίνησης (match moving).

Κατά τη διαδικασία καταγραφής, οι κινήσεις ενός ή περισσότερων ηθοποιών λαμβάνονται ως δείγματα πολλές φορές ανά δευτερόλεπτο. Ενώ οι πρώτες τεχνικές χρησιμοποιούσαν εικόνες από πολλαπλές κάμερες για να υπολογίσουν τις τρισδιάστατες θέσεις, στη συνέχεια ο σκοπός της δέσμευσης κίνησης ήταν να καταγράφει μόνο τις κινήσεις του ηθοποιού και όχι την οπτική του εμφάνιση. Αυτά τα δεδομένα κίνησης αντιστοιχίζονται σε ένα τρισδιάστατο μοντέλο, έτσι ώστε το μοντέλο να εκτελεί τις ίδιες ενέργειες με τον ηθοποιό. Αυτή η διαδικασία μπορεί να αντιπαραβληθεί με την παλαιότερη τεχνική της

ροτοσκόπησης (τεχνική κινούμενης εικόνας για να ανίχνευση κινηματογραφικών πλάνων). Οι κινήσεις χαρακτήρων με αυτήν τη τεχνική μπορούν να επιτευχθούν καταγράφοντας τις κινήσεις του ηθοποιού. Πιο αναλυτικά, ο ηθοποιός εκτελεί μια ενέργεια, και στη συνέχεια η εγγεγραμμένη κίνηση προβάλλεται πάνω σε ένα πίνακα frame-by-frame. Έπειτα εντοπίζονται τα ζωντανά πλάνα δράσης, καταγράφοντας το περίγραμμα του ηθοποιού και τις κινήσεις ανά πλαίσιο, και στη συνέχεια συμπληρώνονται τα εντοπισμένα περιγράμματα με τον κινούμενο χαρακτήρα. Τα τελικά σχέδια φωτογραφίζονται στη συνέχεια ανά πλαίσιο, ταιριάζοντας ακριβώς τις κινήσεις και τις ενέργειες της δράσης. Το τελικό αποτέλεσμα είναι ότι ο κινούμενος χαρακτήρας αντιγράφει ακριβώς τις κινήσεις ζωντανής δράσης του ηθοποιού. Ωστόσο, αυτή η διαδικασία απαιτεί πολύ χρόνο και προσπάθεια.

- Μέθοδοι και συστήματα

Η καταγραφή της κίνησης ή η δέσμευση κίνησης ξεκίνησε ως εργαλείο φωτογραμμετρικής ανάλυσης στη βιομηχανική έρευνα τη δεκαετία του 1970 και τη δεκαετία του 1980 και επεκτάθηκε στην εκπαίδευση, τον αθλητισμό και πρόσφατα στα κινούμενα σχέδια (computer animation) για την τηλεόραση, τον κινηματογράφο και τα βιντεοπαιχνίδια καθώς ωρίμαζε η τεχνολογία. Από τον 20ό αιώνα στο κινούμενο σώμα πρέπει να τοποθετηθούν δείκτες κοντά σε κάθε άρθρωση για να προσδιοριστεί η κίνηση από τις θέσεις ή τις γωνίες μεταξύ των δεικτών. Ακουστικοί, αδρανειακοί, LED, μαγνητικοί ή αντανakλαστικοί δείκτες ή συνδυασμοί οποιονδήποτε από αυτούς, παρακολουθούνται, κατά τουλάχιστον δύο φορές το ποσοστό συχνότητας της επιθυμητής κίνησης. Η ανάλυση του συστήματος είναι σημαντική τόσο στη χωρική ανάλυση όσο και στην προσωρινή ανάλυση, καθώς η θαμπάδα κινήσεων προκαλεί σχεδόν τα ίδια προβλήματα με τη χαμηλή ανάλυση. Από τις αρχές του 21ου αιώνα και εξαιτίας της ταχείας ανάπτυξης της τεχνολογίας αναπτύχθηκαν νέες μέθοδοι. Τα περισσότερα σύγχρονα συστήματα μπορούν να εξαγάγουν τη σιλουέτα του κινητού σώματος από το φόντο. Στη συνέχεια, όλες οι γωνίες άρθρωσης υπολογίζονται με προσαρμογή σε ένα μαθηματικό μοντέλο στην σιλουέτα. Τα συστήματα αυτά χωρίζονται σε οπτικά και μη οπτικά.

- Οπτικά συστήματα

Τα οπτικά συστήματα χρησιμοποιούν δεδομένα που καταγράφονται από τους αισθητήρες εικόνας για τον τριγωνισμό της τρισδιάστατης θέσης ενός υποκειμένου μεταξύ δύο ή περισσότερων καμερών που έχουν βαθμονομηθεί για να παρέχουν επικαλυπτόμενες προβολές. Η απόκτηση δεδομένων γίνεται παραδοσιακά με τη χρήση ειδικών δεικτών που συνδέονται με το αντικείμενο. Ωστόσο, τα πιο πρόσφατα συστήματα είναι σε θέση να παράγουν ακριβή δεδομένα καταγράφοντας χαρακτηριστικά επιφάνειας τα οποία προσδιορίζονται δυναμικά για κάθε συγκεκριμένο θέμα. Η παρακολούθηση μεγάλου αριθμού κινητών αντικειμένων ή η επέκταση της περιοχής λήψης επιτυγχάνεται με την

προσθήκη περισσότερων φωτογραφικών μηχανών. Αυτά τα συστήματα παράγουν δεδομένα με τρεις βαθμούς ελευθερίας για κάθε δείκτη και οι πληροφορίες περιστροφής πρέπει να συναχθούν από τον σχετικό προσανατολισμό τριών ή περισσότερων δεικτών. Για παράδειγμα δείκτες ώμων, αγκώνα και καρπού που παρέχουν τη γωνία του αγκώνα. Τα νεότερα υβριδικά συστήματα συνδυάζουν αδρανειακούς αισθητήρες με οπτικούς αισθητήρες για να αυξήσουν τον αριθμό των χρηστών και να βελτιώσουν την ικανότητα καταγραφής.

➤ Μη οπτικά συστήματα

Αδρανειακά συστήματα

Η τεχνολογία αδρανειακής δέσμευσης κίνησης βασίζεται σε μικροακουστικούς αδρανειακούς αισθητήρες, βιομηχανικά μοντέλα και αλγόριθμους σύντηξης αισθητήρων (sensor fusion). Τα δεδομένα κίνησης των αδρανειακών αισθητήρων (σύστημα αδρανειακής καθοδήγησης) μεταδίδονται συχνά ασύρματα σε έναν υπολογιστή, όπου καταγράφεται ή προβάλλεται η κίνηση. Τα περισσότερα αδρανειακά συστήματα χρησιμοποιούν αδρανειακές μονάδες μέτρησης (IMU) που περιέχουν ένα συνδυασμό γυροσκοπίου, μαγνητόμετρου και επιταχυνσιόμετρου, για τη μέτρηση των ρυθμών περιστροφής. Αυτές οι περιστροφές μεταφράζονται σε ένα σκελετό στο λογισμικό. Όπως συμβαίνει και με τους οπτικούς δείκτες, όσοι περισσότεροι αισθητήρες IMU χρησιμοποιούνται τόσο πιο φυσικά είναι τα δεδομένα. Αξιοσημείωτο είναι ότι δεν απαιτούνται εξωτερικές κάμερες, πομποί ή δείκτες για σχετικές κινήσεις. Τα συστήματα αδρανειακής σύλληψης κίνησης συλλαμβάνουν την πλήρη κίνηση των ανθρώπων σε έξι βαθμούς ελευθερίας σε πραγματικό χρόνο. Εάν περιλαμβάνουν αισθητήρα μαγνητικού ρουλεμάν μπορούν να παρέχουν περιορισμένες πληροφορίες κατεύθυνσης, παρά τη χαμηλότερη ανάλυση και είναι ευαίσθητοι στον ηλεκτρομαγνητικό θόρυβο. Τα οφέλη από τη χρήση αδρανειακών συστημάτων περιλαμβάνουν την καταγραφή σε μια ποικιλία περιβαλλόντων, συμπεριλαμβανομένων στενών χώρων σύλληψης. Τα μειονεκτήματα περιλαμβάνουν χαμηλότερη ακρίβεια θέσης και μετατόπιση θέσης. Αυτά τα συστήματα είναι παρόμοια με τις συσκευές Wii αλλά είναι πιο ευαίσθητα και έχουν μεγαλύτερες ταχύτητες ανάλυσης και ενημέρωσης. Η δημοτικότητα των αδρανειακών συστημάτων αυξάνεται μεταξύ των προγραμματιστών, κυρίως λόγω της γρήγορης και εύκολης εγκατάστασης.

Μηχανική κίνηση

Τα μηχανικά συστήματα καταγραφής κινήσεων παρακολουθούν άμεσα τις αρθρώσεις του σώματος και συχνά αναφέρονται ως εξωσκελετικά συστήματα συλλήψεως κίνησης, λόγω του τρόπου με τον οποίο οι αισθητήρες συνδέονται με το σώμα. Η δομή που μοιάζει με σκελετό συνδέεται με το κινητό σώμα και καθώς αυτό κινείται, κινούνται και τα αρθρωτά μηχανικά μέρη, μετρώντας τη σχετική κίνηση του ερμηνευτή. Τα μηχανικά συστήματα καταγραφής κίνησης είναι συστήματα που λειτουργούν σε πραγματικό χρόνο, σχετικά χαμηλού κόστους, ασύρματα (untethered) που έχουν απεριόριστο όγκο συλλήψεως. Συνήθως, είναι άκαμπτες κατασκευές από αρθρωτές, ευθείες μεταλλικές ή πλαστικές ράβδους συνδεδεμένες μαζί με ποτενσιόμετρα που συνδέονται στις αρθρώσεις του σώματος.

Μαγνητικά συστήματα

Τα μαγνητικά συστήματα υπολογίζουν τη θέση και τον προσανατολισμό από τη σχετική μαγνητική ροή τριών ορθογώνιων πηνίων τόσο στον πομπό όσο και σε κάθε δέκτη. Η σχετική ένταση της τάσης ή του ρεύματος των τριών πηνίων επιτρέπει στα συστήματα αυτά να υπολογίζουν τόσο την εμβέλεια όσο και τον προσανατολισμό, χαρτογραφώντας σχολαστικά τον όγκο παρακολούθησης. Η έξοδος του αισθητήρα είναι 6DOF και παρέχει χρήσιμα αποτελέσματα μόλις με τα δύο τρίτα του αριθμού των στόχων που απαιτούνται για τα οπτικά συστήματα, δηλαδή ένα δείγμα στον άνω βραχίονα και ένα στον κάτω βραχίονα για τη θέση και τη γωνία του αγκώνα. Οι στόχοι δεν επηρεάζονται από μη μεταλλικά αντικείμενα αλλά είναι επιρρεπείς σε μαγνητικές και ηλεκτρικές παρεμβολές από μεταλλικά αντικείμενα στο περιβάλλον, όπως σπλισμός (χάλυβες σπλισμού σκυροδέματος) ή καλωδίωση, που επηρεάζουν το μαγνητικό πεδίο, καθώς και ηλεκτρικές πηγές όπως οθόνες, φώτα, καλώδια και υπολογιστές. Η απόκριση του αισθητήρα είναι μη γραμμική, ειδικά προς τα άκρα της περιοχής σύλληψης. Η καλωδίωση από τους αισθητήρες τείνει να αποκλείει τις ακραίες κινήσεις επιδόσεων. Με τα μαγνητικά συστήματα, είναι δυνατή η παρακολούθηση των αποτελεσμάτων μιας περιόδου καταγραφής κίνησης σε πραγματικό χρόνο. Οι όγκοι σύλληψης για μαγνητικά συστήματα είναι δραματικά μικρότεροι από ό, τι για οπτικά συστήματα. Με τα μαγνητικά συστήματα, υπάρχει διάκριση μεταξύ συστημάτων "AC" και "DC": ο ένας χρησιμοποιεί τετραγωνικούς παλμούς, ο άλλος χρησιμοποιεί παλμό ημιτονοειδούς κύματος.

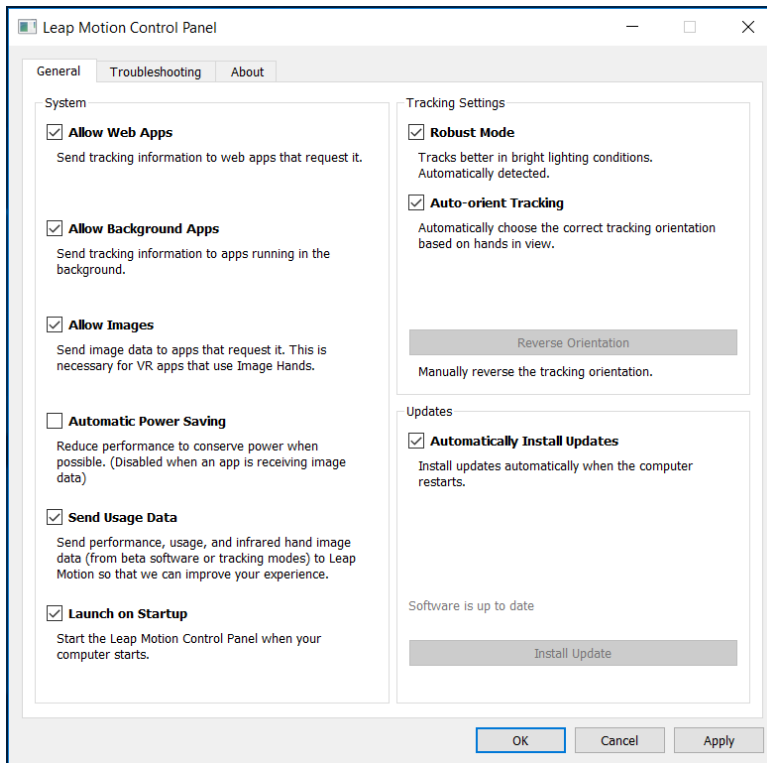
ΚΕΦΑΛΑΙΟ 4

4. Εφαρμογή

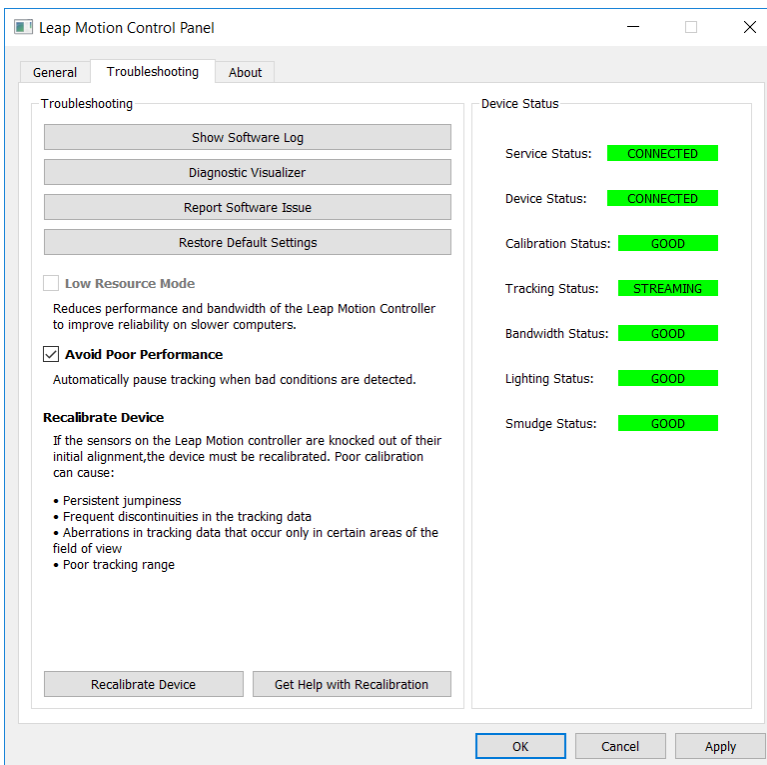
Σκοπός της εφαρμογής που πραγματοποιήθηκε ήταν η καταγραφή, η ανάλυση και η δυνατότητα επανεκτέλεσης της κίνησης μέσω δύο συσκευών ανίχνευσης κίνησης. Με τον όρο επανεκτέλεση που χρησιμοποιείται και παρακάτω εννοείται η εκ νέου οπτικοποίηση των καταγεγραμμένων αρχείων. Για να υλοποιηθεί η εφαρμογή απαραίτητα στοιχεία ήταν δύο συσκευές Leap motion, δύο υπολογιστές και το κατάλληλο λογισμικό για την εκτέλεση. Η συσκευή Leap motion έχει τις εξής απαιτήσεις όσον αφορά στο hardware.

- Windows 7 ή 8 ή Mac OS X 10.7 Lion
- Επεξεργαστή i3 ή i5 ή i7 AMD Phenom ή Intel Core
- Μνήμη 2 GB RAM
- Θύρα USB 2.0
- Σύνδεση στο Internet

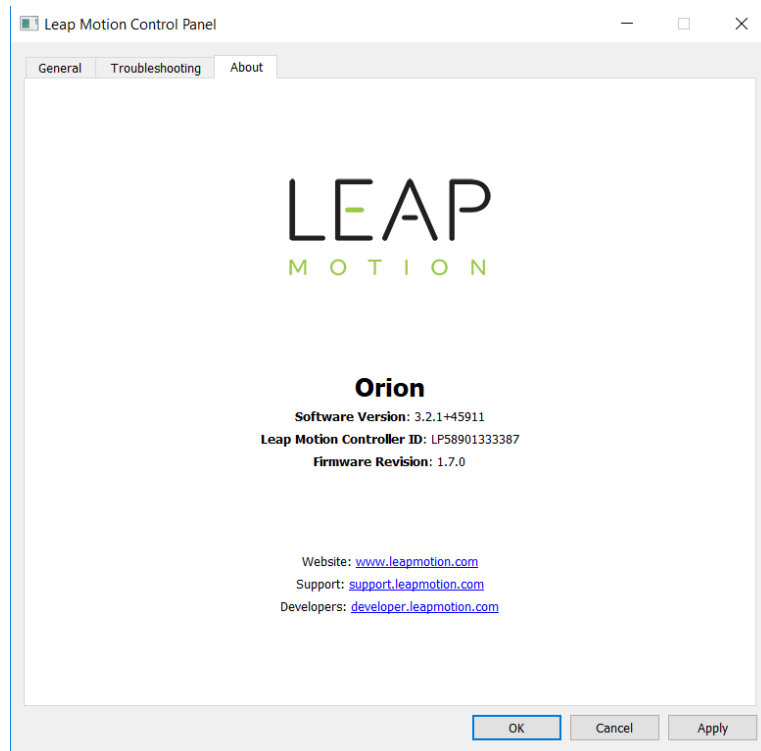
Λαμβάνοντας υπόψη τα προαναφερθέντα χαρακτηριστικά οι δύο υπολογιστές που χρησιμοποιήθηκαν ήταν ένα macOS Sierra Version 10.12.6 και ένα HP Pavillion Gaming 15-ak000nv. Τα χαρακτηριστικά του πρώτου υπολογιστή ήταν οθόνη 15 ιντσών, επεξεργαστής 2,5 Ghz Intel Core i7, μνήμη 16 GB και κάρτα γραφικών AMD Radeon M370X. Τα χαρακτηριστικά του δεύτερου υπολογιστή ήταν οθόνη 15,6 ιντσών, επεξεργαστής 2,6 Ghz Intel Core i7 6700HQ, μνήμη 8 GB και κάρτα γραφικών Nvidia GeForce GTX 950 M. Αρχικά μέσω της επίσημης ιστοσελίδας της συσκευής έγινε η εγκατάσταση του λογισμικού για τη ρύθμιση των συσκευών Leap motion και στους δύο υπολογιστές. Στη συνέχεια έγινε βαθμονόμηση των δύο συσκευών και ο έλεγχος ότι δεν υπάρχει σφάλμα στις ενδείξεις της κατάστασης των δεδομένων. Το πρόβλημα που εμφανίζεται κυρίως αφορά στο φωτισμό, αλλά κατά τη διάρκεια της καταγραφής, η οποία έγινε σε εσωτερικό χώρο, όλα τα φώτα ήταν κλειστά και λόγω της ώρας δεν υπήρχαν εξωγενείς παράγοντες που να επηρεάσουν την καταγραφή. Παρακάτω φαίνεται ότι όλες οι ενδείξεις ήταν θετικές και δεν υπήρχε πρόβλημα κατά την καταγραφή.



Εικόνα 4.1: Πίνακας ρυθμίσεων της συσκευής



Εικόνα 4.2: Πίνακας ελέγχου της κατάστασης της συσκευής



Εικόνα 4.3: Πληροφορίες για την επίσημη ιστοσελίδα και το λογισμικό

Το λογισμικό αυτό ονομάζεται Leap motion control panel και επιτρέπει τη ρύθμιση του αισθητήρα, ελέγχει τις συνθήκες λειτουργίας του και δίνει τη δυνατότητα για οπτικοποίηση των κινήσεων. Για την καταγραφή των δεδομένων βελτιώθηκε ο ήδη υπάρχων κώδικας της υπηρεσίας Leap Recorder με χρήση του επεξεργαστή κειμένου brackets. Η επιλογή της υπηρεσίας Leap Recorder από την επίσημη ιστοσελίδα της συσκευής έγινε διότι δίνει τη δυνατότητα της online καταγραφής της κίνησης αλλά και επανεκτέλεσής της μέσω της αντίστοιχης ιστοσελίδας. Τα αρχεία που αποθηκεύονται είναι τύπου JSON, η μορφή των οποίων αναλύεται παρακάτω.

Το λογισμικό bracket αποτελεί έναν ανοιχτού κώδικα επεξεργαστή κειμένου ο οποίος επιτρέπει την επανάληψη της καταγεγραμμένης κίνησης με το εργαλείο live preview όπου ενημερώνεται κάθε φορά από τις αλλαγές που γίνονται στον κώδικα. Επιπλέον εγκαταστάθηκε γιατί μπορεί και προσομοιώνει ένα server και δίνει τη δυνατότητα να διαβαστούν τα τοπικά αρχεία. Ένα ακόμα πλεονέκτημα του προγράμματος αυτού είναι ότι χρησιμοποιούνται ανανεωμένες βιβλιοθήκες σε σχέση με το demo που παρέχει η υπηρεσία Leap Recorder.

Το πλεονέκτημα της υπηρεσίας Leap Recorder είναι ότι με το ίδιο πρόγραμμα γίνεται και η καταγραφή και η επανεκτέλεση της κίνησης. Όμως ενέκυψαν αρκετά προβλήματα και για αυτό το λόγο επιλέχθηκαν τα παραπάνω λογισμικά και έγιναν οι αντίστοιχες ενέργειες. Η διαδικασία της καταγραφής και επανεκτέλεσης της κίνησης μπορούσε να γίνει μόνο online. Για το λόγο αυτό έγινε «κλωνοποίηση», δηλαδή αντιγραφή του αρχικού κώδικα έτσι ώστε να αποθηκευτεί στον υπολογιστή και να δοθεί η δυνατότητα επανεκτέλεσης του εκάστοτε αρχείου. Μετά την αντιγραφή του προγράμματος παρατηρήθηκε ότι δεν λειτουργούσε καλά και αυτό οφειλόταν σε προβλήματα συμβατότητας. Πιο συγκεκριμένα τα προβλήματα ήταν σχετικά με τα εσωτερικά script του προγράμματος. Παρακάτω παρατίθεται ένα τμήμα του διορθωμένου script. Στην σειρά 43 που φαίνεται στην Εικόνα 4.5 γίνεται η εισαγωγή του εκάστοτε αρχείου που έχει καταγραφεί και χρήζει επανεκτέλεσης. Ένα ακόμα πρόβλημα που παρουσιάστηκε ήταν ότι το πρόγραμμα είναι μια εφαρμογή web, οπότε λόγω των πρωτοκόλλων τα τοπικά αρχεία δεν μπορούν να διαβαστούν. Όλα αυτά συντέλεσαν στην τελική επιλογή.

Όσον αφορά στα αρχεία που αποθηκεύονται, είναι τύπου JSON και είναι δυσκολότερα διαχειρίσιμα από τα αρχεία BVH. Για το λόγο αυτό αρχικά είχε γίνει η προσπάθεια καταγραφής μέσω ενός λογισμικού το οποίο ονομάζεται brekel hands και αποθηκεύει την καταγεγραμμένη κίνηση ως αρχεία BVH, η μορφή των οποίων αναλύεται παρακάτω. Το λογισμικό αυτό επιλέχθηκε ακριβώς για τον τύπο αρχείων που αποθηκεύει. Τα αρχεία αυτά είναι εύκολα διαχειρίσιμα και μπορούν να εισαχθούν και σε αντίστοιχο πρόγραμμα όπως το Blender για την επανεκτέλεσή τους. Το πρόβλημα που παρουσιάστηκε ήταν ότι το συγκεκριμένο λογισμικό δεν ήταν ελεύθερο (open source) και επέτρεπε την καταγραφή μόνο τεσσάρων δευτερολέπτων, ικανά μόνο για δοκιμή και όχι για ανάλυση των δεδομένων, οπότε και δεν χρησιμοποιήθηκε.

```

1 ▼ (function() {
2 ▼   window.recorder.controller('DataCollection', [
3 ▼     '$scope', function($scope) {
4 ▼       var dropArea;
5 ▼       $scope.recordings = [
6 ▼         {
7 ▼           url: 'recordings/gun.json.lz'
8 ▼         }, {
9 ▼           url: 'recordings/HandSplay-110fps.json.lz'
10 ▼        }, {
11 ▼          url: 'recordings/Waiting-110fps.json.lz'
12 ▼        }
13 ▼      ];
14 ▼      $scope.title = function() {
15 ▼        if ($scope.currentRecording().metadata) {
16 ▼          return $scope.currentRecording().metadata.title;
17 ▼        }
18 ▼      };
19 ▼      $scope.safeApply = function(fn) {
20 ▼        var phase;
21 ▼        phase = this.$root.$$phase;
22 ▼        if (phase === '$apply' || phase === '$digest') {
23 ▼          if (fn && (typeof fn === 'function')) {
24 ▼            return fn();
25 ▼          }
26 ▼        } else {
27 ▼          return this.$apply(fn);
28 ▼        }
29 ▼      };
30 ▼      Object.defineProperty($scope, 'mode', {
31 ▼        get: function() {
32 ▼          return $scope._mode;
33 ▼        },
34 ▼        set: function(value) {
35 ▼          if (!(['intro', 'outro', 'recording', 'off'].indexOf(value) > -1)) {
36 ▼            throw "Invalid mode: " + value;
37 ▼          }
38 ▼          $scope._mode = value;
39 ▼          return $scope.safeApply();
40 ▼        }
41 ▼      });
42 ▼      $scope.mode = 'off';
43 ▼      $scope.next = function(e) {
44 ▼        $(e.originalEvent.target).closest('button').get(0).blur();
45 ▼        $scope.currentRecordingIndex++;
46 ▼        $scope.currentRecordingIndex = $scope.currentRecordingIndex % $scope.recordings.length;
47 ▼        return $scope.setCurrentRecording();
48 ▼      };
49 ▼      $scope.previous = function(e) {
50 ▼        $(e.originalEvent.target).closest('button').get(0).blur();
51 ▼        $scope.currentRecordingIndex--;
52 ▼        if ($scope.currentRecordingIndex < 0) {
53 ▼          $scope.currentRecordingIndex = $scope.recordings.length - 1;
54 ▼        }
55 ▼        return $scope.setCurrentRecording();
56 ▼      };
57 ▼      $scope.setCurrentRecording = function() {
58 ▼        return $scope.safeApply(function() {
59 ▼          return player().setRecording($scope.currentRecording()).play();
60 ▼        });
61 ▼      };
62 ▼      $scope.currentRecording = function() {

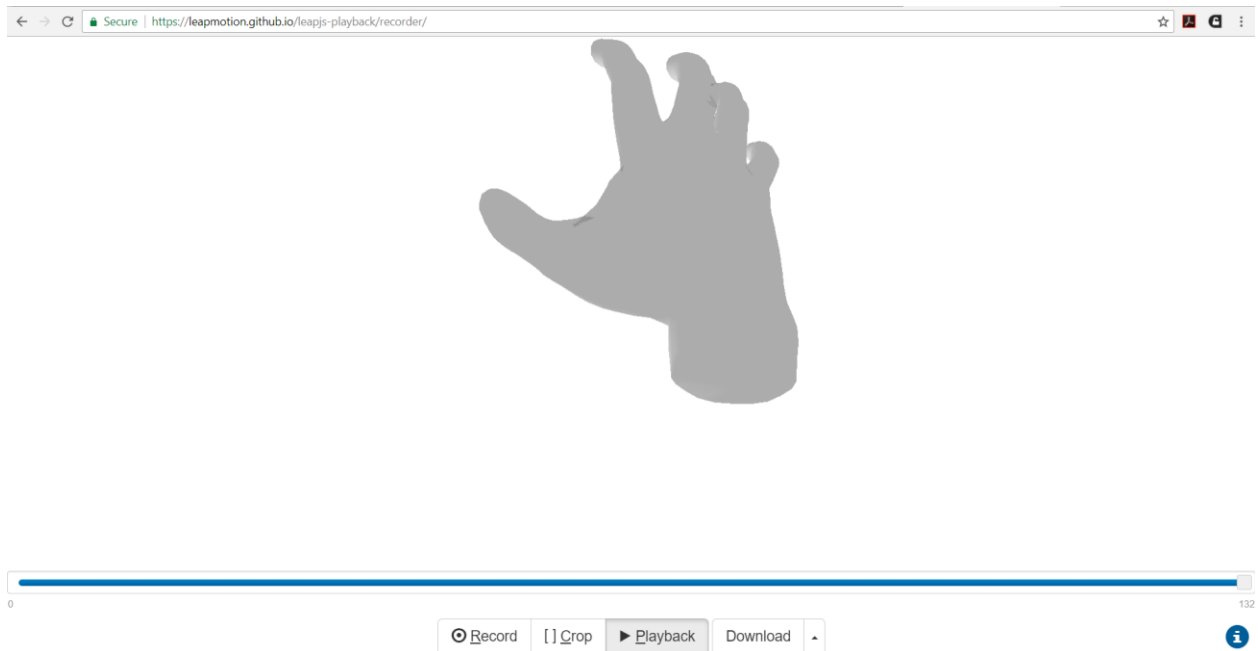
```

Εικόνα 4.4: Τμήμα εσωτερικού script που διαβάζει τα τοπικά αρχεία

```
1 <html>
2 <head>
3
4 <script src="recorder/javascripts/lib/three.min.js"></script>
5
6 <script src="http://js.leapmotion.com/leap-0.6.4.js"></script>
7 <script src="http://js.leapmotion.com/leap-plugins-0.1.11.js"></script>
8 <script src="http://js.leapmotion.com/leap.rigged-hand-0.1.4.js"></script>
9 <!-- override the bundled playback plugin: -->
10 <script src="build/leap.playback-0.2.1.js"></script>
11
12 <style>
13   body{
14     font-family: Helvetica;
15   }
16   canvas {
17     position: fixed;
18     width: 100%;
19     height: 100%;
20     top: 0;
21     left: 0;
22   }
23   #debugger{
24     top: 0;
25     left: 0;
26     position: fixed;
27     background: black;
28     color: #eee;
29     padding: 4px;
30     display: none;
31   }
32 </style>
33
34 </head>
35 <body>
36
37 <div id="debugger"></div>
38
39
40 <script>
41   Leap.loop({})
42   .use('playback', {
43     recording: 'recorder/recordings/numbersrandom.json',
44     requiredProtocolVersion: 6,
45     pauseOnHand: true,
46     loop: true
47   })
48   .use('riggedHand');
49
50   window.controller = Leap.loopController;
51 </script>
52
53 </body>
54 </html>
55
```

Line 43, Column 52 — 55 Lines

Εικόνα 4.5: Λογισμικό Bracket



Εικόνα 4.6: Demo της ιστοσελίδας Leap motion (Leap recorder)

Αρχεία BVH και αρχεία JSON

Οι συσκευές καταγραφής κίνησης επιτρέπουν την καταγραφή ζωντανών κινήσεων εντοπίζοντας διάφορα βασικά σημεία στο χώρο με την πάροδο του χρόνου, τα οποία μεταφράζονται σε μια τρισδιάστατη ψηφιακή αναπαράσταση. Το αντικείμενο μπορεί να είναι οτιδήποτε υπάρχει στον πραγματικό κόσμο, με τα βασικά σημεία τοποθετημένα σε αυτό έτσι ώστε να αντιπροσωπεύουν καλύτερα τους προσανατολισμούς των κινούμενων τμημάτων του αντικειμένου, για παράδειγμα τις ενώσεις των δακτύλων ή τα σημεία περιστροφής (pivot points). Για να γίνει η ακριβής τριγωνική τοποθέτηση των σημείων σήμανσης χρησιμοποιούνται τουλάχιστον 4 κάμερες.

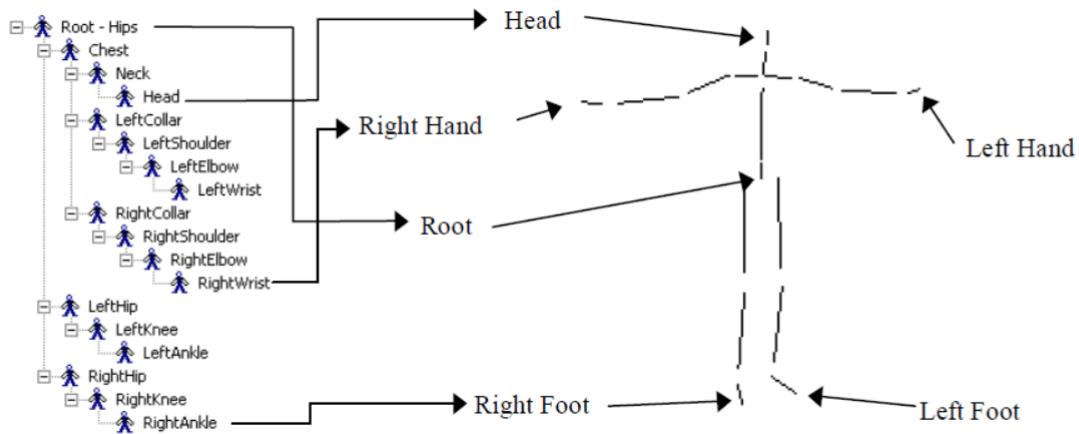
Η χρήση των δεδομένων καταγραφής κίνησης έχει ήδη χρησιμοποιηθεί σε πολλά είδη εφαρμογών. Η επιτυχία των αρχείων αυτών καταγραφής της κίνησης οδήγησε στην αύξηση του αριθμού παραγωγής προϊόντων που μπορούν να καταγράφουν και να παρέχουν δεδομένα κίνησης. Ωστόσο, πολλές εταιρείες ανέπτυξαν τη δική τους μορφή αρχείου. Αυτό σημαίνει ότι οι μορφές αρχείων των δεδομένων δέσμευσης κίνησης δεν είναι καθόλου δεδομένες. Όμως, η φύση ASCII πολλών μορφών καθιστά ευλόγως εύκολη την αποκωδικοποίηση και κατανόηση με απλή επιθεώρηση των δεδομένων.

Παρακάτω παρουσιάζεται μια σύντομη επισκόπηση της ορολογίας και του συμβολισμού που χρησιμοποιείται για να περιγράψει τις μορφές αρχείων, ακολουθούμενη από μια λίστα με πολλές από τις μορφές σύλληψης κίνησης που χρησιμοποιούνται σήμερα. Στη συνέχεια εξηγούνται δύο από τις μορφές αρχείων όσον αφορά τη δομή και τη διαδικασία που απαιτούνται για τη σωστή ερμηνεία των δεδομένων για αναπαραγωγή. Οι μορφές αυτές είναι τα αρχεία BVH και τα αρχεία JSON. Η πρώτη μορφή ήταν η πρώτη προσέγγιση για την καταγραφή και αναπαράσταση της κίνησης μέσω της συσκευής Leap motion και η δεύτερη ήταν αυτή που τελικά χρησιμοποιήθηκε.

Ορολογία

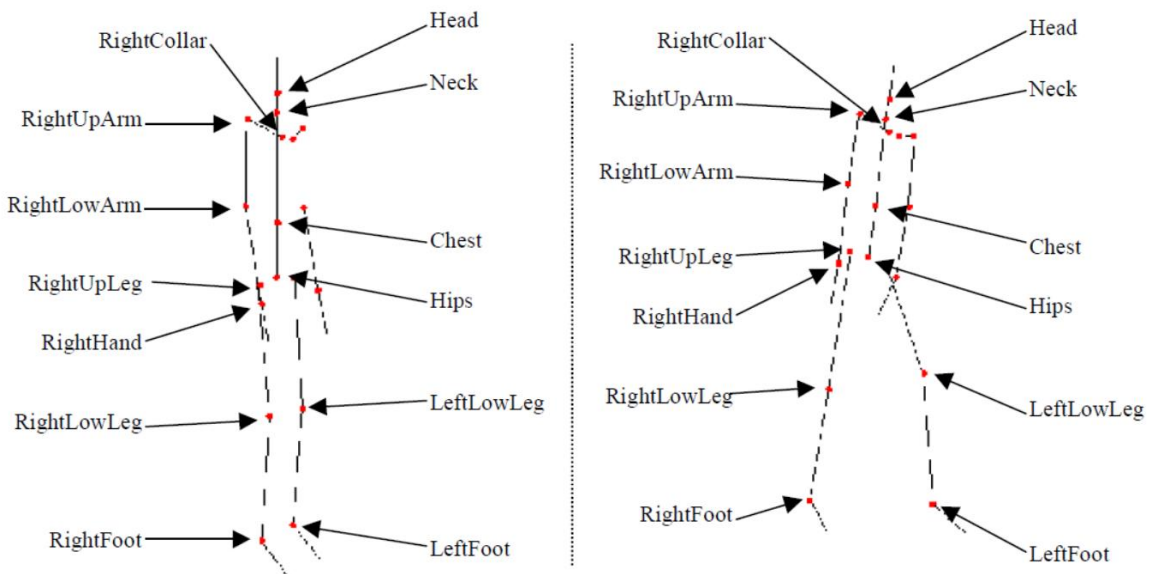
Αυτή η ενότητα παρέχει μια σύντομη επισκόπηση της ορολογίας και του συμβολισμού που θα χρησιμοποιηθεί για να περιγραφούν οι διαδικασίες που σχετίζονται με την ανάγνωση και επεξεργασία δεδομένων κίνησης. Η παρακάτω λίστα περιγράφει μερικές από τις πιο σημαντικές λέξεις-κλειδιά που θα χρησιμοποιηθούν για την αναγνώριση και περιγραφή διαφόρων πτυχών της κίνησης.

- **Skeleton:** Ολόκληρος ο χαρακτήρας για τον οποίο παρουσιάζεται η κίνηση.
- **Bone:** Η βασική οντότητα που αντιπροσωπεύει έναν σκελετό. Κάθε οστό αντιπροσωπεύει το μικρότερο τμήμα στο πλαίσιο της κίνησης που υπόκειται σε μεμονωμένες αλλαγές μετατόπισης και προσανατολισμού κατά τη διάρκεια της κίνησης. Ένας σκελετός αποτελείται από έναν αριθμό οστών, συνήθως σε μια ιεραρχική δομή (Εικόνα 7), όπου κάθε οστό μπορεί να συσχετιστεί με μια κορυφή για να αντιπροσωπεύει ένα συγκεκριμένο τμήμα του χαρακτήρα, για παράδειγμα το μηρό ή το βραχίονα.
- **Κανάλι ή βαθμός ελευθερίας (DOF)** - Κάθε οστό μέσα σε ένα σκελετό μπορεί να υποβληθεί σε αλλαγή θέσης, προσανατολισμού και αλλαγές κλίμακας κατά τη διάρκεια της κίνησης, όπου κάθε παράμετρος από τις προαναφερθείσες αναφέρεται ως κανάλι ή βαθμός ελευθερίας. Οι αλλαγές στα δεδομένα του καναλιού με την πάροδο του χρόνου δημιουργούν την κίνηση.
- **Frame** - Κάθε κίνηση αποτελείται από έναν αριθμό frame όπου για κάθε καρέ ορίζεται το κανάλι δεδομένων για κάθε οστό. Τα δεδομένα δέσμησης κίνησης μπορούν να ληφθούν μέχρι και με 240 καρέ ανά δευτερόλεπτο, ωστόσο σε πολλές εφαρμογές ένας ρυθμός 30 ή 60 καρέ ανά δευτερόλεπτο τείνει να είναι ο κανόνας. Οι υψηλοί ρυθμοί καρέ χρησιμοποιούνται για τη λήψη κινήσεων που περιέχουν περιεχόμενο υψηλής συχνότητας, όπως ο συνδυασμός κινήσεων καράτε. Παρόλο που σε πολλές περιπτώσεις δεν είναι δυνατή η εμφάνιση των επιπλέον λεπτομερειών κατά τη διάρκεια μιας αναπαραγωγής σε πραγματικό χρόνο, λόγω των μέγιστων ρυθμών ανανέωσης του υλικού εμφάνισης, μπορεί παρέχουν χρήσιμες πληροφορίες για την ανάλυση της κίνησης



Εικόνα 4.7: Ιεραρχική δομή των δεδομένων της κίνησης

Η μορφή BVH αποτελεί την εξέλιξη της μορφής δεδομένων BVA της BioVision με την αξιόλογη προσθήκη μιας ιεραρχικής δομής δεδομένων που αντιπροσωπεύει τα οστά του σκελετού. Το αρχείο BVH αποτελείται από δύο μέρη όπου η πρώτη ενότητα περιγράφει λεπτομερώς την ιεραρχία και την αρχική πόζα του σκελετού και το δεύτερο τμήμα περιγράφει τα δεδομένα καναλιού για κάθε frame (καρέ), δηλαδή το τμήμα κίνησης. Εικονογραφήσεις της θέσης βάσης και του πρώτου πλαισίου μιας κινούμενης εικόνας δίνονται στην Εικόνα 4.8 α), όπου τα δεδομένα παρατίθενται στην Εικόνα 4.8 β) (Meredith et al. 2014).



Εικόνα 4.8: Σκελετική δομή ενός παραδείγματος αρχείου BVH (α) αρχική θέση (β) πρώτο καρέ της κίνησης

Το JSON (JavaScript Object Notation) είναι ένα πρότυπο ανταλλαγής δεδομένων. Είναι εύκολο τόσο για τους χρήστες να το διαβάσουν και να γράψουν όσο και τις μηχανές να το αναλύσουν (parse) και να το παράγουν (generate). Είναι βασισμένο πάνω σε ένα υποσύνολο της γλώσσας προγραμματισμού JavaScript, Standard ECMA-262 Έκδοση 3η - Δεκέμβριος 1999. Το JSON είναι ένα πρότυπο κειμένου το οποίο είναι τελείως ανεξάρτητο από γλώσσες προγραμματισμού αλλά χρησιμοποιεί πρακτικές (conventions) οι οποίες είναι γνωστές στους προγραμματιστές της οικογένειας προγραμματισμού C, συμπεριλαμβανομένων των C, C++, C#, Java, JavaScript, Perl, Python, και πολλών άλλων. Αυτές οι ιδιότητες κάνουν την JSON μια ιδανική γλώσσα προγραμματισμού ανταλλαγής δεδομένων.

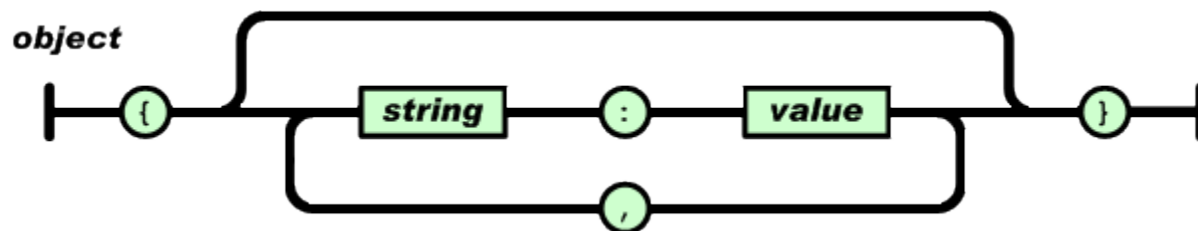
Τα αρχεία JSON είναι χτισμένα σε δύο δομές:

- Μια συλλογή από ζευγάρια ονομάτων/τιμών. Σε πολλές γλώσσες προγραμματισμού, αυτή η δομή χαρακτηρίζεται ως αντικείμενο (object), εγγραφή (record), struct (δομή), dictionary (λεξικό), πίνακας κατακερματισμού (hash table), λίστα κλειδιών (keyed list) ή συσχετιστική διάταξη (associative array).
- Μία ταξινομημένη λίστα τιμών. Στις περισσότερες γλώσσες προγραμματισμού, αυτή η δομή ονομάζεται πίνακας (array), διάνυσμα (vector), λίστα (list) ή ακολουθία (sequence).

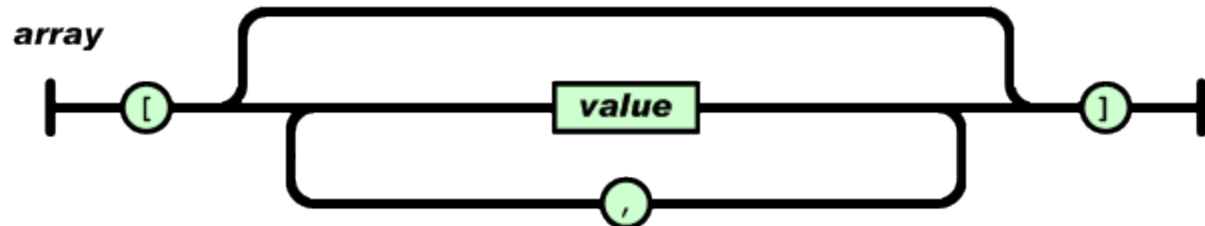
Αυτές είναι οι παγκόσμιες δομές δεδομένων. Ουσιαστικά όλες οι σύγχρονες γλώσσες προγραμματισμού τις υποστηρίζουν με τον έναν ή τον άλλον τρόπο και είναι λογικό πως ένα πρότυπο δεδομένων το οποίο είναι εύκολα μεταβαλλόμενο από γλώσσες προγραμματισμού θα βασίζεται σε αυτές τις δομές.

Τα αρχεία JSON παίρνουν αυτές τις μορφές:

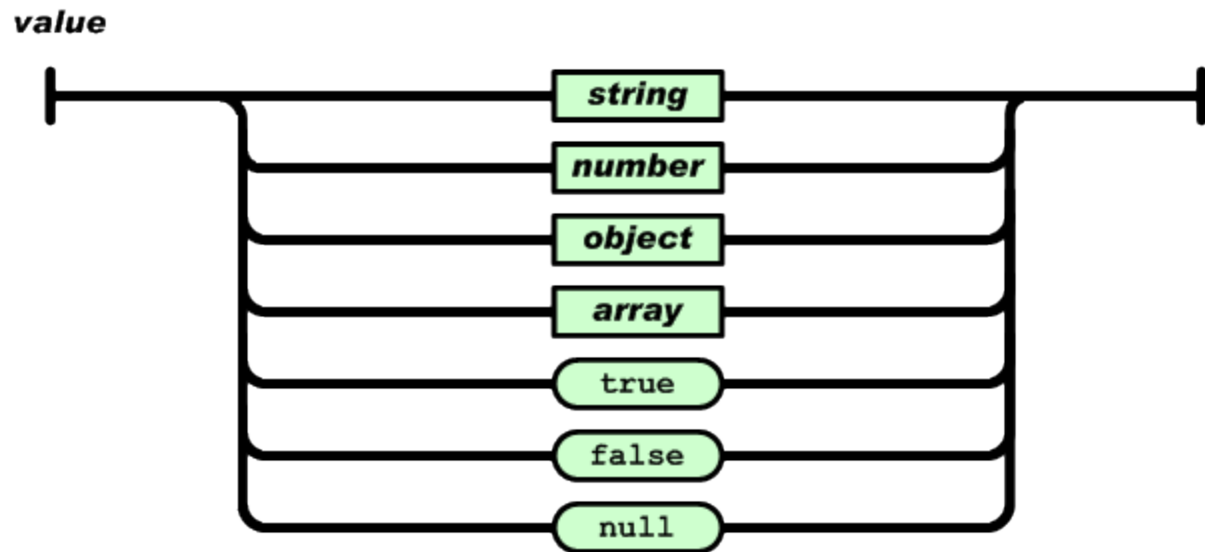
Ένα αντικείμενο (object) είναι ένα σύνολο από ζευγάρια ονομάτων/τιμών. Κάθε αντικείμενο ξεκινάει με { (αριστερό άγκιστρο) και τελειώνει με } (δεξιό άγκιστρο). Κάθε όνομα ακολουθείται από : (άνω-κάτω τελεία) και τα ζευγάρια ονόματος/τιμής χωρίζονται με , (κόμμα).



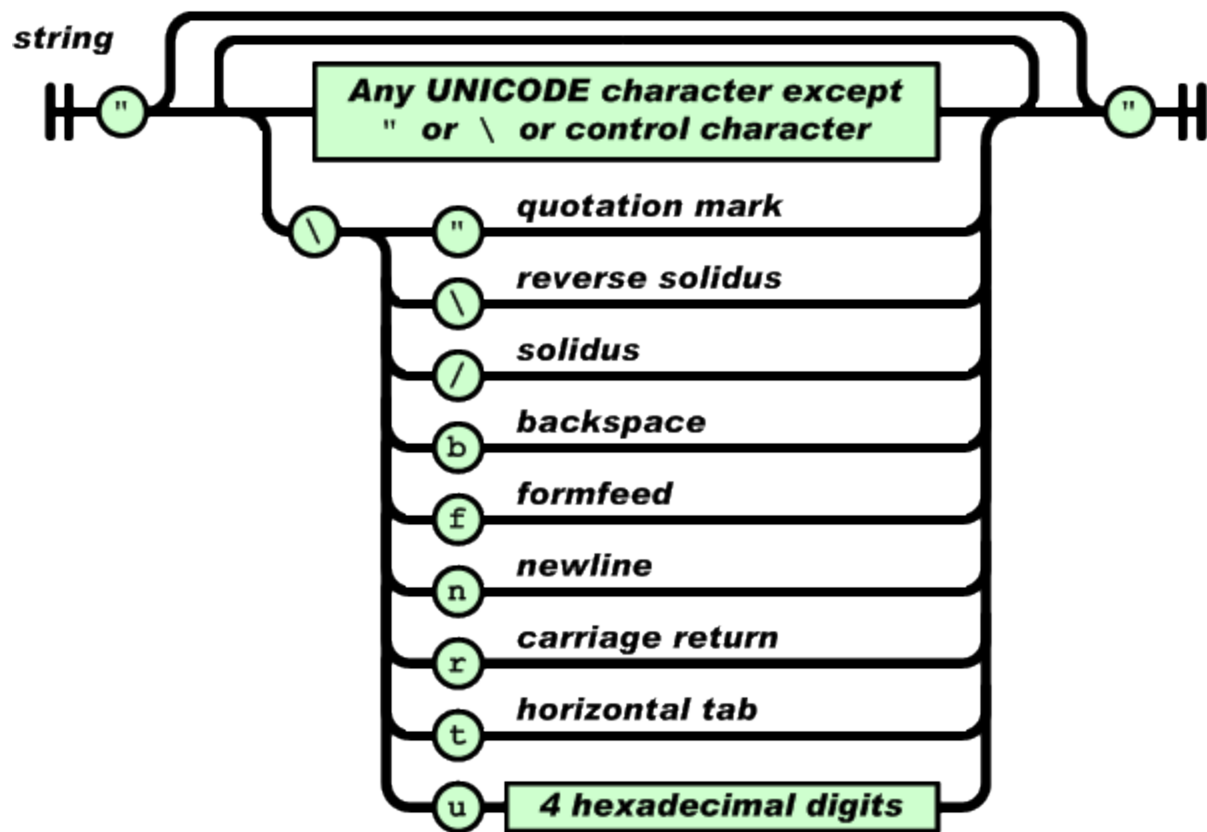
Ένας πίνακας (array) είναι μια συλλογή από τιμές σε σειρά. Κάθε πίνακας (array) ξεκινάει με [(αριστερή αγκύλη) και τελειώνει με] (δεξιά αγκύλη). Οι τιμές χωρίζονται με, (κόμμα).



Μία τιμή μπορεί να είναι χαρακτήρας (string) μέσα σε διπλά quotes, ή αριθμός (number), ή αληθής (true) ή ψευδής (false) ή null, ή αντικείμενο (object) ή πίνακας (array). Αυτές οι τιμές μπορεί να είναι και ανακατεμένες.



Ένα string είναι μια συλλογή από μηδέν ή περισσότερους Unicode χαρακτήρες, μέσα σε διπλά quotes, χρησιμοποιώντας αντίστροφους κάθετους \ (backslash) για escapes. Ένα string μοιάζει πολύ σαν ένα C ή Java string.



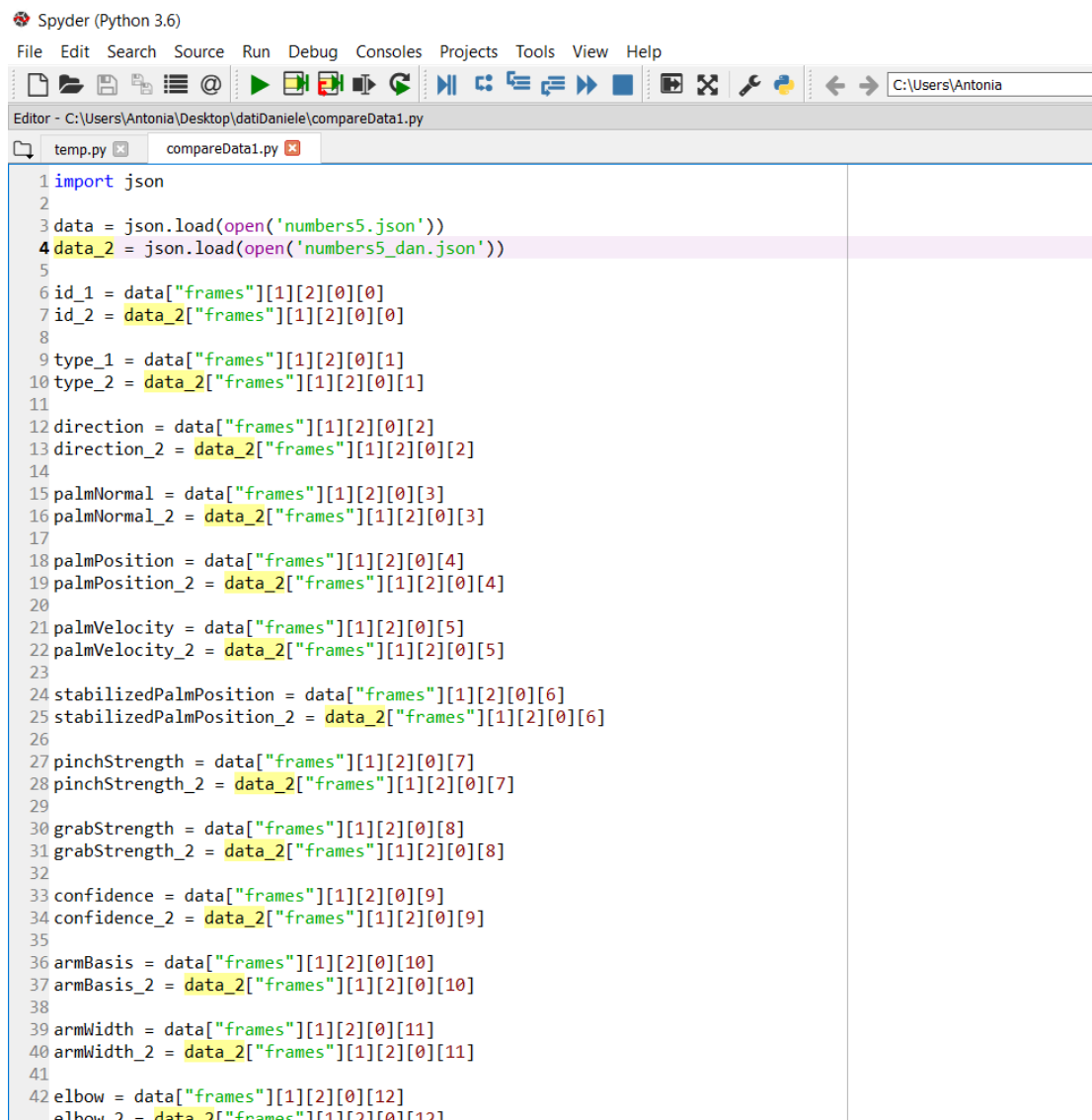
4.1 Καταγραφή και ανάλυση δεδομένων

Με γνωστή τη μορφή των αρχείων επόμενο βήμα είναι η καταγραφή των δεδομένων όπως φαίνεται στην παρακάτω εικόνα. Οι συσκευές τοποθετήθηκαν με τον ίδιο προσανατολισμό η μία δίπλα στην άλλη έτσι ώστε να μπορεί να γίνει η σύγκριση των αποτελεσμάτων και να είναι ευκολότερη η ανάλυσή τους γνωρίζοντας ότι έχουν το ίδιο σύστημα συντεταγμένων. Έγινε καταγραφή του ενός χεριού να αναπαριστά τα νούμερα από το ένα έως το πέντε, καταγραφή και των δύο χεριών με το ένα σχεδόν ακίνητο να αναπαριστούν τα νούμερα από το έξι έως το δέκα και καταγραφή και των δύο χεριών να εκτελούν τυχαίους αριθμούς, για να κινούνται και τα δύο.



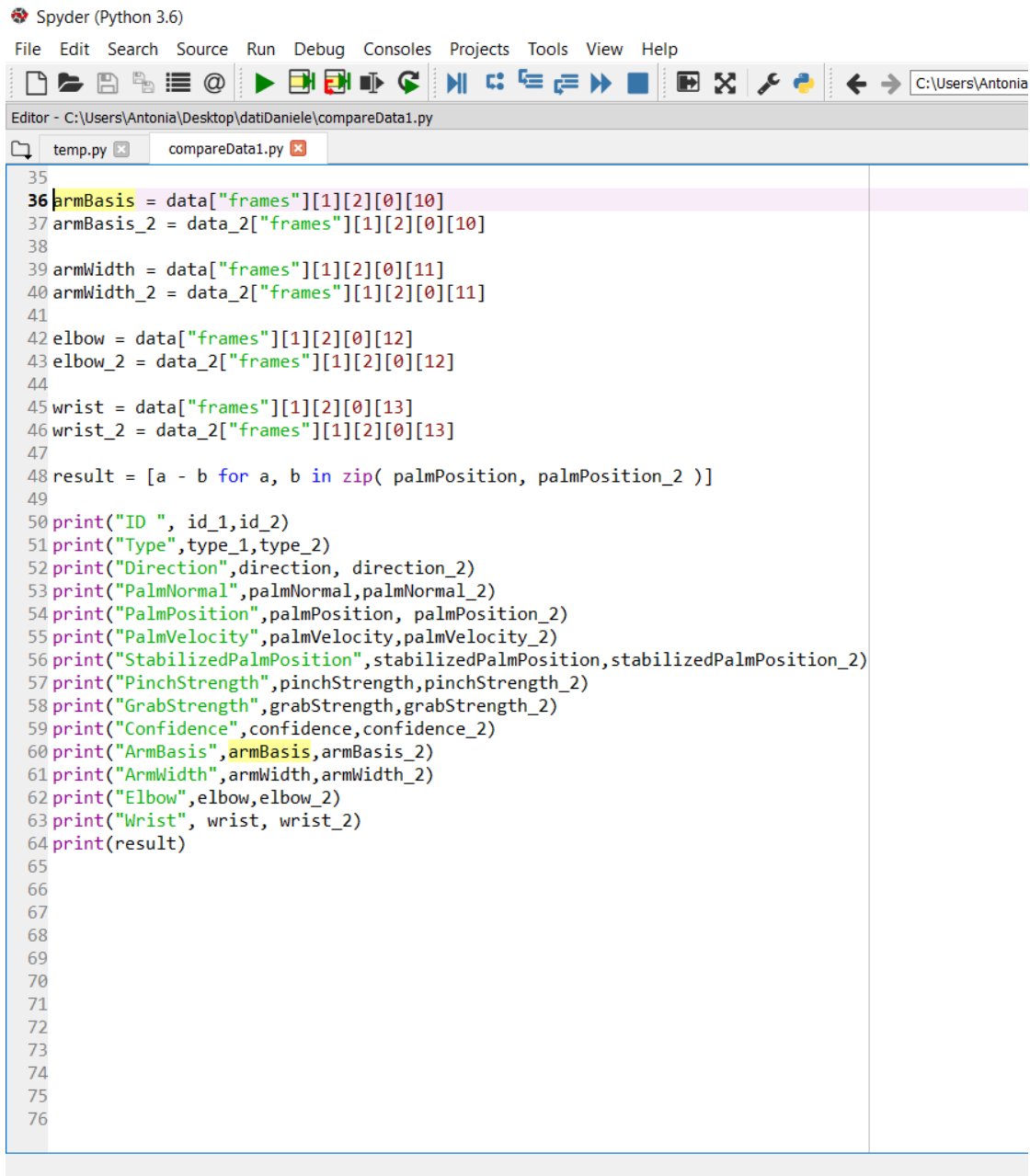
Εικόνα 4.9: Καταγραφή δεδομένων

Αρχικά, όπως αναλύθηκε και παραπάνω, ο τύπος JSON περιέχει διάφορες δομές, λίστες, αριθμούς, βιβλιοθήκες. Οι δομές αυτές για να μπορέσουν να αποθηκευτούν σε ένα αρχείο πρέπει να μετατραπούν σε χαρακτήρες (strings). Μόνο με αυτόν τον τρόπο μπορεί ο τύπος αυτός να διαβαστεί ή να γραφτεί σε ένα αρχείο. Η Python διαθέτει ένα module που επιτρέπει τη μετατροπή αυτή. Για το λόγο αυτό επιλέχθηκε ένα ανοικτού κώδικα ολοκληρωμένο αναπτυξιακό περιβάλλον (IDE) επ' ονόματι Spyder το οποίο επιτρέπει τον προγραμματισμό στη γλώσσα Python. Μετά την καταγραφή των δεδομένων λοιπόν, έγινε η εισαγωγή τους στο πρόγραμμα Spyder για την ανάλυσή τους. Όπως φαίνεται στην Εικόνα 4.10 με την εντολή `import json` εισάγεται το αντίστοιχο module.



```
1 import json
2
3 data = json.load(open('numbers5.json'))
4 data_2 = json.load(open('numbers5_dan.json'))
5
6 id_1 = data["frames"][1][2][0][0]
7 id_2 = data_2["frames"][1][2][0][0]
8
9 type_1 = data["frames"][1][2][0][1]
10 type_2 = data_2["frames"][1][2][0][1]
11
12 direction = data["frames"][1][2][0][2]
13 direction_2 = data_2["frames"][1][2][0][2]
14
15 palmNormal = data["frames"][1][2][0][3]
16 palmNormal_2 = data_2["frames"][1][2][0][3]
17
18 palmPosition = data["frames"][1][2][0][4]
19 palmPosition_2 = data_2["frames"][1][2][0][4]
20
21 palmVelocity = data["frames"][1][2][0][5]
22 palmVelocity_2 = data_2["frames"][1][2][0][5]
23
24 stabilizedPalmPosition = data["frames"][1][2][0][6]
25 stabilizedPalmPosition_2 = data_2["frames"][1][2][0][6]
26
27 pinchStrength = data["frames"][1][2][0][7]
28 pinchStrength_2 = data_2["frames"][1][2][0][7]
29
30 grabStrength = data["frames"][1][2][0][8]
31 grabStrength_2 = data_2["frames"][1][2][0][8]
32
33 confidence = data["frames"][1][2][0][9]
34 confidence_2 = data_2["frames"][1][2][0][9]
35
36 armBasis = data["frames"][1][2][0][10]
37 armBasis_2 = data_2["frames"][1][2][0][10]
38
39 armWidth = data["frames"][1][2][0][11]
40 armWidth_2 = data_2["frames"][1][2][0][11]
41
42 elbow = data["frames"][1][2][0][12]
43 elbow_2 = data_2["frames"][1][2][0][12]
```

Εικόνα 4.10

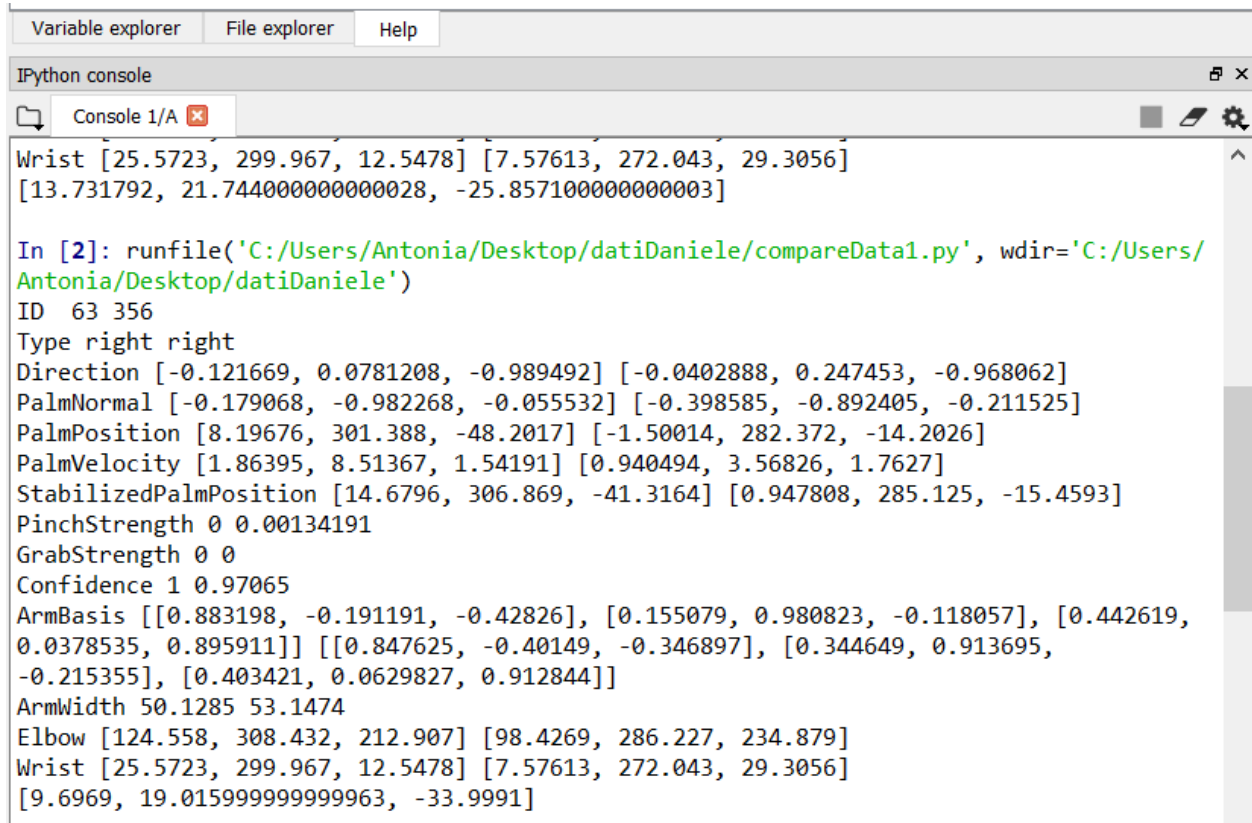


```
35
36 armBasis = data["frames"][1][2][0][10]
37 armBasis_2 = data_2["frames"][1][2][0][10]
38
39 armWidth = data["frames"][1][2][0][11]
40 armWidth_2 = data_2["frames"][1][2][0][11]
41
42 elbow = data["frames"][1][2][0][12]
43 elbow_2 = data_2["frames"][1][2][0][12]
44
45 wrist = data["frames"][1][2][0][13]
46 wrist_2 = data_2["frames"][1][2][0][13]
47
48 result = [a - b for a, b in zip( palmPosition, palmPosition_2 )]
49
50 print("ID ", id_1,id_2)
51 print("Type",type_1,type_2)
52 print("Direction",direction, direction_2)
53 print("PalmNormal",palmNormal,palmNormal_2)
54 print("PalmPosition",palmPosition, palmPosition_2)
55 print("PalmVelocity",palmVelocity,palmVelocity_2)
56 print("StabilizedPalmPosition",stabilizedPalmPosition,stabilizedPalmPosition_2)
57 print("PinchStrength",pinchStrength,pinchStrength_2)
58 print("GrabStrength",grabStrength,grabStrength_2)
59 print("Confidence",confidence,confidence_2)
60 print("ArmBasis",armBasis,armBasis_2)
61 print("ArmWidth",armWidth,armWidth_2)
62 print("Elbow",elbow,elbow_2)
63 print("Wrist",wrist,wrist_2)
64 print(result)
65
66
67
68
69
70
71
72
73
74
75
76
```

Εικόνα 4.11

Στη συνέχεια, γίνεται η εισαγωγή των αρχείων καταγραφής των οποίων θα γίνει η ανάλυση με τις εντολές που φαίνονται στις σειρές 3 και 4 του κώδικα της Εικόνας 4.10. Στην σειρά 3 γίνεται εισαγωγή του αρχείου της πρώτης συσκευής Leap Motion και στην σειρά 4 της δεύτερης συσκευής. Για την εισαγωγή άλλων καταγραφών αρκεί να γίνει αλλαγή του ονόματος και να τοποθετηθεί το αντίστοιχο αρχείο.

Όλα τα αρχεία πρέπει να βρίσκονται στην ίδια τοποθεσία γιατί το πρόγραμμα ακολουθεί συγκεκριμένη διαδρομή (path). Έπειτα, ορίζονται τα επιθυμητά στοιχεία και «τυπώνονται» στην οθόνη όπως φαίνεται στην Εικόνα 4.12.



```
Variable explorer | File explorer | Help
IPython console
Console 1/A
Wrist [25.5723, 299.967, 12.5478] [7.57613, 272.043, 29.3056]
[13.731792, 21.744000000000028, -25.85710000000003]

In [2]: runfile('C:/Users/Antonia/Desktop/datiDaniele/compareData1.py', wdir='C:/Users/Antonia/Desktop/datiDaniele')
ID 63 356
Type right right
Direction [-0.121669, 0.0781208, -0.989492] [-0.0402888, 0.247453, -0.968062]
PalmNormal [-0.179068, -0.982268, -0.055532] [-0.398585, -0.892405, -0.211525]
PalmPosition [8.19676, 301.388, -48.2017] [-1.50014, 282.372, -14.2026]
PalmVelocity [1.86395, 8.51367, 1.54191] [0.940494, 3.56826, 1.7627]
StabilizedPalmPosition [14.6796, 306.869, -41.3164] [0.947808, 285.125, -15.4593]
PinchStrength 0 0.00134191
GrabStrength 0 0
Confidence 1 0.97065
ArmBasis [[0.883198, -0.191191, -0.42826], [0.155079, 0.980823, -0.118057], [0.442619, 0.0378535, 0.895911]] [[0.847625, -0.40149, -0.346897], [0.344649, 0.913695, -0.215355], [0.403421, 0.0629827, 0.912844]]
ArmWidth 50.1285 53.1474
Elbow [124.558, 308.432, 212.907] [98.4269, 286.227, 234.879]
Wrist [25.5723, 299.967, 12.5478] [7.57613, 272.043, 29.3056]
[9.6969, 19.015999999999963, -33.9991]
```

Εικόνα 4.12: «Τυπωμένα» στοιχεία καταγραφής

Τα στοιχεία αυτά ορίζονται με βάση τη δομή των δεδομένων που φαίνεται στην Εικόνα 4.13. Για παράδειγμα αν το ζητούμενο δεδομένο είναι ο τύπος του χεριού στο τέταρτο frame, δηλαδή αν είναι το αριστερό ή το δεξί χέρι, θα πρέπει να γραφτεί : `data["frames"][4][2][0][1]`, όπου το 4 αναφέρεται στο τέταρτο frame, το 2 αναφέρεται στα χέρια, το 0 αναφέρεται στο πρώτο χέρι και το 1 στον τύπο του χεριού. Αντίστοιχα αν το ζητούμενο είναι το δεύτερο χέρι το μηδέν αντικαθίσταται με ένα και αν το ζητούμενο είναι η θέση της παλάμης το ένα αντικαθίσταται με τέσσερα.


```

1 data Structure
2
3 data["frames"][number of frame]
4 data["frames"][1 to the last one][0] = id
5 data["frames"][1 to the last one][1] = timestamp
6 data["frames"][1 to the last one][2] = hands
7 data["frames"][1 to the last one][3] = pointables
8 data["frames"][1 to the last one][4] = interactionBox
9
10 hands
11 data["frames"][1][2][0] = first hand
12 data["frames"][1][2][1] = second hand
13
14 [0] = 'id',
15 [1] = 'type',
16 [2] = 'direction',
17 [3] = 'palmNormal',
18 [4] = 'palmPosition',
19 [5] = 'palmVelocity',
20 [6] = 'stabilizedPalmPosition',
21 [7] = 'pinchStrength',
22 [8] = 'grabStrength',
23 [9] = 'confidence',
24 [10] = 'armBasis',
25 [11] = 'armWidth',
26 [12] = 'elbow',
27 [13] = 'wrist'

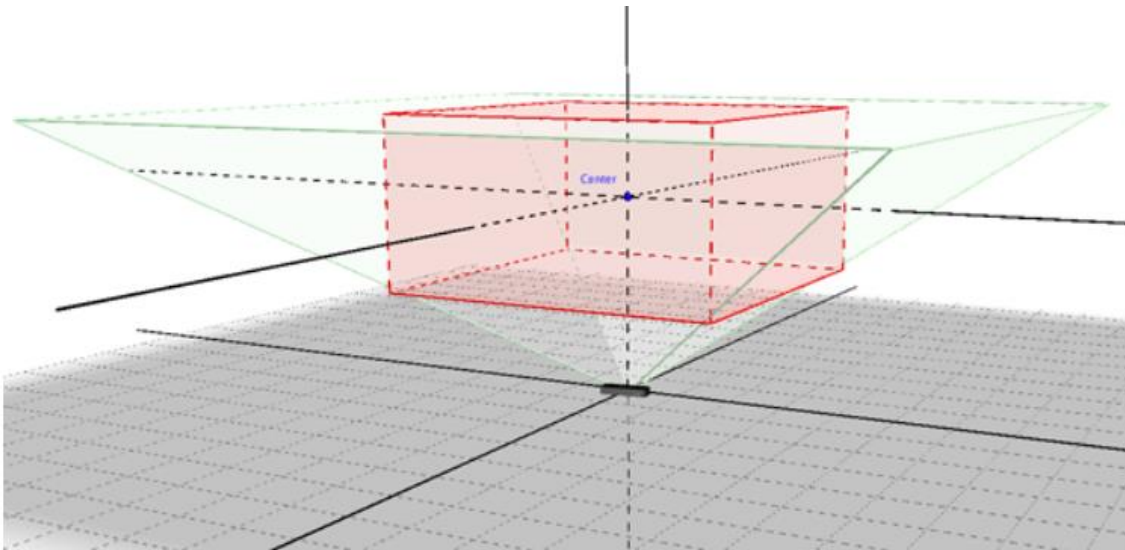
```

Εικόνα 4.13: Δομή δεδομένων αρχείου json

Ακολούθως αναλύονται όλα τα προς εξέταση στοιχεία.

- Id : Είναι ο μοναδικός κωδικός για κάθε frame.
- Timestamp: Αποτελεί τον χρόνο σύλληψης του καρτέ (frame) σε microseconds.
- Hands: Αναφέρει τα φυσικά χαρακτηριστικά ενός χεριού που ανιχνεύθηκε.
- Pointables: Το Pointable αναφέρει τα φυσικά χαρακτηριστικά του εκάστοτε ανιχνευμένου δακτύλου.
- InteractionBox: Το InteractionBox αναπαριστά μια χωρική περιοχή μορφής παραλληλογράμμου που βρίσκεται εντελώς μέσα στο οπτικό πεδίο του αισθητήρα Leap Motion. Αποτελεί ένα ορθογώνιο πρίσμα ευθυγραμμισμένο με τους άξονες της συσκευής και παρέχει κανονικοποιημένες συντεταγμένες για τα χέρια και τα δάχτυλα που κινούνται

μέσα σε αυτό το πεδίο. Οι κανονικοποιημένες συντεταγμένες είναι αυτές που μετατρέπονται από χιλιοστά σε ένα εύρος από 0 έως 1 έτσι ώστε η ελάχιστη τιμή του να αντιστοιχεί σε 0 και η μέγιστη τιμή σε 1. Πιο συγκεκριμένα η αριστερή, κάτω θέση έχει την κανονικοποιημένη θέση: $[0,0,0]$, ενώ η δεξιά, πάνω θέση: $[1,1,1]$. Το InteractionBox μπορεί να διευκολύνει την εύρεση των θέσεων στο σύστημα συντεταγμένων του Leap Motion και ορίζεται από ένα κέντρο (center) και τις διαστάσεις κατά μήκος των αξόνων x, y και z (size).

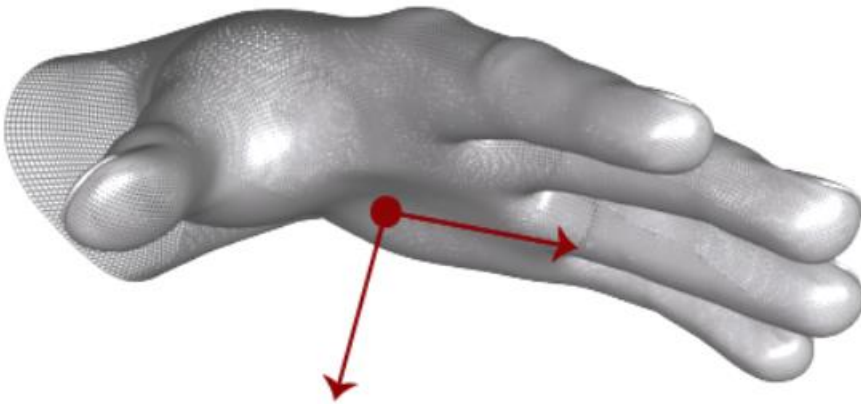


Εικόνα 4.14: Interaction Box

Πιο συγκεκριμένα, όσον αφορά στα χαρακτηριστικά του χεριού τα προς εξέταση στοιχεία είναι τα εξής:

- Id: Ο μοναδικός κωδικός για κάθε χέρι
- Type: Ο τύπος του χεριού, δηλαδή δεξί ή αριστερό
- Direction: Ένα διάνυσμα κατεύθυνσης που δείχνει από το κέντρο της παλάμης προς τα δάχτυλα. Χρησιμοποιείται για να υπολογιστούν γωνίες κλίσης και περιστροφής της παλάμης σε σχέση με το οριζόντιο επίπεδο.

- Palm Normal: Ένα διάνυσμα κάθετο στο επίπεδο που σχηματίζεται από την παλάμη του χεριού.



- Palm Position: Οι συντεταγμένες του κέντρου της παλάμης που μετρώνται σε χιλιοστά.
- Palm Velocity: Η ταχύτητα και η κατεύθυνση της κίνησης της παλάμης σε χιλιοστά ανά δευτερόλεπτο.
- StabilizedPalmPosition: Η φιλτραρισμένη θέση της παλάμης του χεριού. Η εξομάλυνση και η σταθεροποίηση γίνεται για να καταστεί αυτή η τιμή πιο κατάλληλη για αλληλεπίδραση με περιεχόμενο 2D. Η σταθεροποιημένη θέση υστερεί πίσω από τη θέση της παλάμης με μεταβλητή ποσότητα, εξαρτώμενη κυρίως από την ταχύτητα κίνησης.
- GrabStrength: Παίρνει τιμές από μηδέν έως ένα. Η τιμή της παραμέτρου είναι μηδέν όταν το χέρι είναι ανοιχτό και φτάνει στο 1 όταν αναγνωρίζεται κίνηση κρατήματος αντικειμένου.
- PinchStrength: Η τιμή της αντίστοιχα είναι μηδέν όταν το χέρι είναι ανοιχτό και φτάνει στο 1 όταν αναγνωρίζεται κίνηση ένωσης του αντίχειρα με οποιοδήποτε άλλο δάχτυλο.

- Confidence: Η εκτίμηση της καταλληλότητας των δεδομένων. Οι τιμές κυμαίνονται από 0 έως 1.
- ArmBasis: Ο βασικός πίνακας που αντιπροσωπεύει τον προσανατολισμό του βραχίονα.
- ArmWidth: Το εκτιμώμενο μέσο πλάτος του βραχίονα.
- Elbow: Οι συντεταγμένες του αγκώνα που μετρώνται σε χιλιοστά.
- Wrist: Οι συντεταγμένες του καρπού που μετρώνται σε χιλιοστά.

Συγκρίνοντας τα δεδομένα παρατηρείται ότι δεν υπάρχουν μεγάλες διαφορές από τη μία συσκευή στην άλλη. Οι μεγαλύτερες διαφορές είναι στον άξονα των z και αυτό οφείλεται στη διαφορετική θέση των συσκευών. Αφαιρώντας από τις μετρήσεις κάθε φορά την απόσταση των συσκευών, η οποία είναι ίση με το πλάτος της συσκευής, δηλαδή 12,7 mm μπορεί να γίνει η σύγκριση των τιμών. Πριν από τη σύγκριση των αποτελεσμάτων των δύο συσκευών πρέπει το σύστημα των συντεταγμένων να έχει κοινή αρχή. Συνεπώς μετατοπίζεται το σύστημα είτε της μίας είτε της άλλης συσκευής έτσι ώστε να συμπίπτουν. Κατά αυτόν τον τρόπο οι διαφορές στον άξονα των z μειώνονται.

Αντίστοιχη διαδικασία θα γινόταν και με διαφορετική διάταξη στις συσκευές. Η συγκεκριμένη διάταξη επιλέχθηκε διότι, από όρισμού του ο αισθητήρας δημιουργήθηκε για να τοποθετείται σε φυσική επιφάνεια στραμμένος προς τα επάνω. Οποιαδήποτε άλλη διάταξη θα έδινε λύση σε προβλήματα ορατότητας διευρύνοντας το οπτικό πεδίο, όπως το ότι όταν η παλάμη θα στρεφόταν 90 μοίρες θα υπήρχε δυσκολία διαχωρισμού της από τα δάκτυλα. Από την άλλη πλευρά ο προσανατολισμός της συσκευής θα άλλαζε κάτι το οποίο θα επηρέαζε την ακρίβεια και θα δυσκόλευε τη σύγκριση των δύο συσκευών. Φυσικά μόνο με δύο συσκευές δεν είναι δυνατόν να εξαχθούν συμπεράσματα όσον αφορά στην ακρίβεια της συσκευής. Όμως, δεδομένου του ότι οι μεγαλύτερες διαφορές είναι της τάξης του εκατοστού, θα μπορούσε να εξαχθεί το συμπέρασμα ότι η συσκευή μπορεί να χρησιμοποιηθεί για την καταγραφή χειρονομιών. Με την ένωση των δεδομένων όχι μόνο από δύο αλλά και περισσότερες συσκευές τα αποτελέσματα θα είναι ακριβέστερα.

ΚΕΦΑΛΑΙΟ **5**

5.Συμπεράσματα και προτάσεις

Στην παρούσα διπλωματική εργασία έγινε διερεύνηση του συστήματος Leap Motion με σκοπό την καταγραφή χειρονομιών. Μεγαλύτερη έμφαση δόθηκε στον τρόπο λειτουργίας του αισθητήρα όσο και στα αποτελέσματα που δίνει. Έγινε ανάλυση όλων των στοιχείων που μπορούν να εξαχθούν όσο και η επεξήγησή τους. Λόγω του γεγονότος ότι αποτελεί μια νέα τεχνολογία και έχουν διεξαχθεί ελάχιστες έρευνες που μελετούν τα εξαγόμενα αυτά στοιχεία, η εξαγωγή συμπερασμάτων είναι αρκετά δύσκολη. Οι περισσότερες μελέτες αφορούν στη χρήση της συσκευής ως μέσο διεπαφής του ανθρώπου με τον υπολογιστή και όχι ως μετρητικό μέσο. Αμφισβητείται λοιπόν η ακρίβεια και η αξιοπιστία του.

Η δυσκολία αξιολόγησης έγκειται στο γεγονός ότι τα μαθηματικά και οι αλγόριθμοι που χρησιμοποιούνται για την εύρεση των συντεταγμένων παραμένουν άγνωστα. Συγκρίνοντας τις διαφορετικές καταγραφές που έγιναν παρατηρήθηκε ότι δεν υπάρχει συνέπεια στη συχνότητα δειγματοληψίας. Σε κάποιες καταγραφές τα καρέ ήταν πολύ περισσότερα και σε άλλες λιγότερα χωρίς να διαφαίνεται κάποια σταθερά στη συχνότητα. Επιπλέον, η αξιοπιστία του κάθε αισθητήρα μεμονωμένα δεν μπορεί να αξιολογηθεί αφού δεν υπάρχει σύγκριση των μετρήσεων με άλλες συσκευές. Η ακρίβεια όμως είναι αρκετά ικανοποιητική ανάμεσα στους δύο αισθητήρες, αφού οι διαφορές σε όλες τις καταγραφές ανάμεσα τους ήταν της τάξης του εκατοστού, με την ακρίβεια να φτάνει στο ένα χιλιοστό. Κατά την επανεκτέλεση των αρχείων γίνεται αντιληπτό ότι σε μερικά δεν έχουν καταγραφεί όλες οι κινήσεις και υπάρχει απόκλιση από τις αρχικές.

Όπως γίνεται φανερό με όσο περισσότερες συσκευές γίνει η καταγραφή τόσο πιο ακριβή θα είναι τα αποτελέσματα, αφού δίνεται η δυνατότητα ένωσης των δεδομένων. Στοιχεία που δεν έχουν καταγραφεί από τη μία συσκευή θα παίρνονται από την άλλη και όπου υπάρχουν δεδομένα και από τις δύο συσκευές θα επιλέγεται το καταλληλότερο ανάλογα με τη θέση της συσκευής του χεριού αλλά και ένας μέσος όρος από όλες τις συσκευές. Για να κριθεί η καταλληλότητα του αισθητήρα ως ενός οικονομικά ελκυστικού αισθητήρα καταγραφής κίνησης, χειρονομιών και εντοπισμού αντικειμένων θα πρέπει να γίνει και η σύγκρισή του με άλλες δοκιμασμένες συσκευές έτσι ώστε να ελεγχθεί η ακρίβεια και η πιστότητά του.

Εκτός από την ανάλυση του αισθητήρα Leap Motion διερευνήθηκε και η συσκευή Kinect. Σύμφωνα με τα τεχνικά χαρακτηριστικά της συσκευής και των αρχών λειτουργίας της, συμπεραίνεται ότι και αυτή η συσκευή αποτελεί έναν οικονομικά ελκυστικό αισθητήρα καταγραφής κίνησης, χειρονομιών και εντοπισμού αντικειμένων. Η συγκεκριμένη συσκευή έχει χρησιμοποιηθεί πολύ περισσότερο για μετρητικούς σκοπούς και τα αποτελέσματα είναι ορατά σε πολλές μελέτες.

Μια πρόταση αποτελεί ο συνδυασμός των συσκευών. Με το Kinect τα αποτελέσματα για τη θέση του σώματος θα είναι πολύ πιο ακριβή, χωρίς όμως να δίνονται πολλές λεπτομέρειες. Για το λόγο αυτό θα γίνει ο συνδυασμός με τον αισθητήρα Leap Motion ώστε να μπορούν να καταγραφούν ακόμα και οι κινήσεις των δακτύλων. Επιπλέον, η χρήση περισσότερων από δύο συσκευών Leap Motion αλλά και η χρήση περισσότερων από δύο συσκευών Kinect θα διευκόλυνε την καταγραφή. Περισσότερες καταγραφές μπορούν να δώσουν απαντήσεις για την ακρίβεια και την αξιοπιστία των συσκευών όπως και μεγαλύτερη ανάλυση των επιμέρους δεδομένων. Η ένωση των δεδομένων του Leap Motion είναι αρκετά πιο εύκολη, αφού και τα ίδια τα δεδομένα είναι αρκετά διαχειρίσιμα σε σχέση με τα δεδομένα του Kinect, όμως οι βασικές αρχές είναι οι ίδιες και αυτό αποτελεί μια πρόκληση για το μέλλον.

Βιβλιογραφία

ΞΕΝΗ ΒΙΒΛΙΟΓΡΑΦΙΑ

Aziz, A.A., Wan, K., Zaaba, S.K., Shahrman, A.B., Adnan, N.H., Nor, R.M., Ayob, M.N., Ismail, A.H., Ramly, M.F., 2012. Development of a gesture database for an adaptive gesture recognition system. *Int. J. Electr. Comput. Sci.* 12, 38 - 44.

Bracegirdle, A., Mitrovic, S. T., Mathews, M., 2014. Investigating the usability of the leap motion controller: Gesturebased interaction with a 3d virtual environment

Burkhard, P., Langston, J., Tetrud, J., 2002. Voluntarily simulated tremor in normal subjects, *Neurophysiol. Clin.* 32, 119 - 126.

Dimitropoulos, K., Kitsikidis, A., Grammalidis, N., 2017. Information Technologies in Traditional Dance Analysis

Hernoux, F., Béarée, R., Gajny, L., Nyiri, E., Bancalini, J., Gibaru, O., 2013. Leap Motion pour la capture de mouvement 3D par spline L1. Application à la robotique. *Journées du Groupe de Travail en Modélisation Géométrique*

Marin, G., Dominio, F., Zanuttigh, P., 2015. Hand gesture recognition with jointly calibrated leap motion and depth sensor.

Meredith, M., Maddock, S., 2001. Motion Capture File Formats Explained

Moeslund, T., Hilton, A. Krüger, V., 2006. A survey of advances in vision-based human motion capture and analysis.

Poppe, R., 2007. Vision-based human motion analysis: An overview. *Computer Vision and Image Understanding*, 108

Regenbrecht, H., Collins, J., Hoermann, S. A, 2013. Leap-Supported, Hybrid AR Interface Approach. In Proceedings of the OZCHI'13, 25th Australian Computer-Human Interaction Conference: Augmentation, Application, Innovation, Collaboration, Adelaide, Australia

Sturman, M.M., Vaillancourt, D.E., Corcos, D., 2005. Effects of aging on the regularity of physiological tremor. *J. Neurophysiol.* 93, 3064–3074.

Teather, R.J., Pavlovych, A., Stuerzlinger, W., MacKenzie, I.S., 2009. Effects of Tracking Technology, Latency, and Spatial Jitter on Object Movement. In Proceedings of the IEEE Symposium on 3D User Interfaces, Lafayette, LA, USA, pp. 43–50.

Vikram, S., 2013. Handwriting and Gestures in the Air, Recognizing on the Fly. In Proceedings of the CHI 2013 Extended Abstracts, Paris, France

Weichert, F., Bachmann, D., Rudak, B., Fisseler, D., 2013. Analysis of the accuracy and robustness of the leap motion controller. *Sensors*, 13, 6380–6393.

Khoshelham, K., Elberink, S.O., 2012. Accuracy and resolution of kinect depth data for indoor mapping applications. *Sensors*, 12, 1437–1454.

ΕΛΛΗΝΙΚΗ ΒΙΒΛΙΟΓΡΑΦΙΑ

Αρβανίτης, Κ., 2015. Διασύνδεση του Microsoft Kinect με 3D Gaming Engine και υλοποίηση συστήματος Real-Time Dynamic Skeleton Retargeting

Βωβός, Ν., 2014. Πλατφόρμα παιχνιδιών κινητικής αποκατάστασης με την χρήση του Unity 3D σε συνεργασία με το Kinect της Microsoft

Δουληφής, Σ., 2016. Μελέτη Αλγορίθμων Τμηματοποίησης σε 3D Νέφη Σημείων Με Χρήση Αισθητήρα KINECT

Κόντος, Χ., 2014. Ανάπτυξη διαδραστικής εφαρμογής: “αόρατα κρουστά” με την χρήση της συσκευής Kinect και της γλώσσας προγραμματισμού Processing

Κωτσίδου, Δ., 2015. Ανίχνευση προτύπων κίνησης χεριών: Πρότυπη υλοποίηση για κωφάλαλους

Συντυχάκης, Ε., 2012. Ανάπτυξη Εφαρμογής με τη χρήση του Microsoft Kinect Sensor