



ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ

ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ

ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

**Μελέτη και σύγκριση αλγορίθμων κατηγοριοποίησης
ανθρώπινης κίνησης που χρησιμοποιούνται σε
φορετούς υπολογιστές**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΤΟΥ

Μαρτίδη Χρήστου

Βόλος, 11/07/2014

Η σελίδα αυτή είναι σκόπιμα λευκή.



ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ
ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ
ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ
ΥΠΟΛΟΓΙΣΤΩΝ

**Μελέτη και σύγκριση αλγορίθμων κατηγοριοποίησης
ανθρώπινης κίνησης που χρησιμοποιούνται σε φορητούς
υπολογιστές**

**Study and comparison of algorithms of human motion
classification used in wearable computers**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΤΟΥ

Μαρτίδη Χρήστου

Επιβλέποντες :

Τσομπανοπούλου Παναγιώτα Επίκουρος Καθηγήτρια	Χούστης Ηλίας Ομότιμος Καθηγητής
Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών, Πανεπιστήμιο Θεσσαλίας	Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών, Πανεπιστήμιο Θεσσαλίας

Εγκρίθηκε από την διμελή εξεταστική επιτροπή την 11/07/2014

(Υπογραφή)

.....

ΚΥΡΙΟΣ ΕΠΙΒΛΕΠΩΝ
ΙΔΙΟΤΗΤΑ - ΠΑΝΕΠΙΣΤΗΜΙΟ

Τσομπανοπούλου Παναγιώτα

Επίκουρος Καθηγήτρια

Τμήμα Ηλεκτρολόγων Μηχανικών
και Μηχανικών Υπολογιστών,
Πανεπιστήμιο Θεσσαλίας

(Υπογραφή)

.....

ΔΕΥΤΕΡΕΥΩΝ ΕΠΙΒΛΕΠΩΝ
ΙΔΙΟΤΗΤΑ - ΠΑΝΕΠΙΣΤΗΜΙΟ

Χούστης Ηλίας

Ομότιμος Καθηγητής

Τμήμα Ηλεκτρολόγων Μηχανικών
και Μηχανικών Υπολογιστών,
Πανεπιστήμιο Θεσσαλίας

(Υπογραφή)

.....

ΜΑΡΤΙΔΗΣ ΧΡΗΣΤΟΣ

Διπλωματούχος Μηχανικός Ηλεκτρονικών Υπολογιστών, Τηλεπικοινωνιών και
Δικτύων του Τμήματος Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών,
Πανεπιστημίου Θεσσαλίας

© 2014 – All rights reserved

Η σελίδα αυτή είναι σκόπιμα λευκή.

Πίνακας περιεχομένων

1	Εισαγωγή.....	1
2	Fall Detection	3
2.1	ΑΦΗΡΗΜΕΝΑ.....	3
2.2	ΑΡΧΕΣ ΚΑΙ ΜΕΘΟΔΟΙ	4
2.3	ΑΞΙΟΛΟΓΗΣΗ ΤΩΝ ΑΙΣΘΗΤΗΡΩΝ ΠΤΩΣΗΣ.....	7
2.4	ΑΛΓΟΡΙΘΜΟΙ ΑΝΙΧΝΕΥΣΗΣ ΠΤΩΣΗΣ	11
	2.4.1 ΑΝΙΧΝΕΥΤΗΣ ΠΤΩΣΗΣ ΒΑΣΙΣΜΕΝΟΣ ΣΕ ΤΡΙΑΞΟΝΙΚΟ ΕΠΙΤΑΧΥΝΣΙΟΜΕΤΡΟ ΑΠΟ ΤΟΝ GARRETT BROWN.....	12
	2.4.2 ΣΥΓΚΡΙΣΗ ΑΛΓΟΡΙΘΜΩΝ ΓΙΑ ΑΝΙΧΝΕΥΣΗ ΠΤΩΣΗΣ ΧΑΜΗΛΗΣ ΠΟΛΥΠΛΟΚΟΤΗΤΑΣ.....	28
	2.4.3 ΕΝΑΣ ΑΛΓΟΡΙΘΜΟΣ ΑΝΙΧΝΕΥΣΗΣ ΠΤΩΣΗΣ ΜΕ ΧΡΗΣΗ ΓΥΡΟΣΚΟΠΙΟΥ.....	39
2.5	ΣΥΜΠΕΡΑΣΜΑΤΑ	50
3	Motion Capture.....	51
3.1	ΣΥΣΤΗΜΑΤΑ ΚΑΤΑΓΡΑΦΗΣ ΤΗΣ ΚΙΝΗΣΗΣ	52
3.2	ΠΕΔΙΑ ΕΦΑΡΜΟΓΗΣ.....	53
3.3	ΗΔΗ ΑΡΧΕΙΩΝ ΚΙΝΗΣΗΣ.....	55
	3.3.1 ΑΡΧΕΙΑ BIOVISION HIERARCHY (BHV).....	55
	3.3.2 ΑΡΧΕΙΑ ASF/AMC.....	56
	3.3.3 ΑΡΧΕΙΑ C3D.....	56
3.4	ΜΙΑ ΠΕΡΙΕΚΤΙΚΗ MULTIMODAL HUMAN ACTION DATABASE (MHAD) ..	57

3.4.1	ΥΠΑΡΧΟΥΣΕΣ ΒΑΣΕΙΣ ΔΕΔΟΜΕΝΩΝ ΑΝΑΓΝΩΡΙΣΗΣ ΑΝΘΡΩΠΙΝΗΣ ΚΙΝΗΣΗΣ.....	57
3.4.2	MULTIMODAL HUMAN ACTION DATABASE.....	61
3.4.3	ΑΛΓΟΡΙΘΜΟΙ ΠΟΥ ΧΡΗΣΙΜΟΠΟΙΟΥΝΤΑΙ ΓΙΑ ΤΗΝ ΑΝΑΓΝΩΡΙΣΗ ΤΗΣ ΔΡΑΣΗΣ.....	64
3.4.4	ΠΕΙΡΑΜΑΤΑ ΑΝΑΓΝΩΡΙΣΗΣ ΤΗΣ ΔΡΑΣΗΣ.....	68
3.4.5	ΜΟΝΟΤΡΟΠΙΚΑ (UNIMODAL) ΑΠΟΤΕΛΕΣΜΑΤΑ.....	74
3.4.6	ΠΟΛΥΤΡΟΠΙΚΑ (MULTIMODAL) ΑΠΟΤΕΛΕΣΜΑΤΑ.....	74
3.5	ΑΝΑΚΑΤΑΣΚΕΥΗ ΤΗΣ ΚΙΝΗΣΗΣ ΜΕ ΧΡΗΣΗ ΑΡΑΙΩΝ ΔΕΔΟΜΕΝΩΝ ΕΠΙΤΑΧΥΝΣΙΟΜΕΤΡΩΝ.....	76
3.5.1	ΕΙΣΑΓΩΓΗ.....	76
3.5.2	ΕΛΕΓΧΟΣ ΕΙΣΟΔΟΥ ΚΑΙ ΒΑΣΗ ΓΝΩΣΗΣ.....	78
3.5.3	ONLINE LAZY NEIGHBORHOOD GRAPH (OLNG).....	82
3.5.4	ΑΝΑΚΑΤΑΣΚΕΥΗ ΤΗΣ ΚΙΝΗΣΗΣ.....	89
3.5.5	ΑΠΟΤΕΛΕΣΜΑΤΑ.....	95
4	Επίλογος.....	106
	Βιβλιογραφία.....	108
	ΠΑΡΑΡΤΗΜΑ Α.....	113

ABSTRACT : In this work we will study and compare algorithms for classification of human movement using wearable computers.

We will divide the work into two distinct chapters together :

1. **Fall Detection:** Using wearable computers will study and compare different algorithms that try to detect human falls.
2. **Motion Capture:** We will study the mocap systems that are used to create animation, interactive games and visual effects in movies. We will also study some recognition algorithms of human motion and a reconstruction algorithms of motion (for creating animation) , based on low-dimensional control signal.

Ο Επιάς το 1945, ήταν ο πρώτος μεγάλης κλίμακας ηλεκτρονικός ψηφιακός υπολογιστής ικανός να λύσει ένα πλήρες εύρος υπολογιστικών προβλημάτων . Ο Επιάς είχε περισσότερες από 18.000 λυχνίες κενού και 1500 ηλεκτρονόμους. Ζύγιζε 30 τόνους καταλάμβανε 63 τετραγωνικά μέτρα χώρο και κόστισε 500.000 \$. Από εκείνη την εποχή και μέχρι σήμερα, η ανάπτυξη της τεχνολογίας και η εξέλιξη των υπολογιστικών συστημάτων ήταν ραγδαία. Πλέον με λίγα μόνο ευρώ μπορεί κάποιος να αγοράσει υπολογιστές, smartphones , laptops κτλ. , όπου έχουν μεγάλες υπολογιστικές ικανότητες και μικρό όγκο.

Η εξέλιξη αυτή των υπολογιστικών συστημάτων , η δραματική μείωση του όγκου τους και η μείωση του κόστους τους, έδωσε την ευκαιρία για την χρήση των υπολογιστών σε ευρεία κλίμακα και άνοιξε τον δρόμο για την εφαρμογή των υπολογιστών σε πολλούς τομείς. Πλέον με το μέγεθος που έχουν οι υπολογιστές , μπορούν να μεταφερθούν από έναν άνθρωπο ή να “φορευθούν” πάνω σε αυτόν. Ένας τομέας , πάνω στον οποίο έχει εφαρμογή η φορητότητα και η “φορετότητα” των υπολογιστών, είναι η μελέτη της ανθρώπινης κίνησης.

Στην παρούσα εργασία θα ασχοληθούμε με την ανθρώπινη κίνηση. Πιο συγκεκριμένα θα μελετήσουμε και θα συγκρίνουμε αλγορίθμους που χρησιμοποιούνται για την κατηγοριοποίηση της ανθρώπινης κίνησης με χρήση φορετών υπολογιστών. Θα χωρίσουμε την εργασία σε 2 διακριτά μεταξύ τους κεφάλαια :

- **Την ανίχνευση μιας πτώσης (Fall Detection):** Με την χρήση φορετών υπολογιστών θα μελετήσουμε και θα συγκρίνουμε διάφορους αλγορίθμους που προσπαθούν να ανιχνεύσουν τις πτώσεις ενός ανθρώπου . Εμείς θα επικεντρωθούμε στις πτώσεις των ηλικιωμένων.
- **Την καταγραφή της κίνησης (Motion Capture):** Θα μελετήσουμε τα συστήματα καταγραφής της κίνησης που χρησιμοποιούνται για την δημιουργία animation , σε διαδραστικά παιχνίδια , για την δημιουργία οπτικών εφέ σε ταινίες κτλ. Επίσης θα μελετήσουμε κάποιους αλγορίθμους αναγνώρισης της ανθρώπινης κίνησης , αλλά και έναν αλγόριθμο που προσπαθεί να ανακατασκευάσει την ανθρώπινη κίνηση βασισμένο σε ένα χαμηλής διάστασης σήμα ελέγχου.

2

Fall Detection

Με τον όρο «Fall Detector» (Ανιχνευτής Πτώσης) εννοούμε μία φορητή συσκευή η οποία μπορεί να αναγνωρίσει την ύπαρξη πτώσης του ατόμου που φοράει την συσκευή.

Η ανίχνευση πτώσης είναι ένα ενδιαφέρον επιστημονικό πρόβλημα , δεδομένου ότι είναι μία ασαφής διαδικασία που μπορεί κανείς να προσεγγίσει χρησιμοποιώντας διαφορετικές μεθόδους.

Στην συγκεκριμένη ενότητα θα μελετήσουμε και θα αναλύσουμε την τεχνολογία και τους αλγόριθμους που χρησιμοποιούνται από τις συσκευές ανίχνευσης πτώσης .

2.1 ΑΦΗΡΗΜΕΝΑ

Η εξέλιξη της ιατρικής και το συνεχώς αυξανόμενο , κατά μέσο όρο , βιωτικό επίπεδο των ανθρώπων έχει ως αποτέλεσμα το αναμενόμενο προσδόκιμο ζωής να έχει αυξηθεί τα τελευταία χρόνια, ιδικά στις αναπτυγμένες χώρες.

Ένα σημαντικό πρόβλημα των ηλικιωμένων είναι οι τραυματισμοί από πτώσεις και ένας από τους λόγους όπου πολλά υγιή άτομα αναγκάζονται να αφήσουν την άνεση

και ιδιωτικότητα του σπιτιού τους , για να ζήσουν σε ένα περιβάλλον με συνεχή φροντίδα. Εκτιμάται ότι πάνω από το ένα τρίτο των ενηλίκων , ηλικίας άνω των 65 χρονών , έχουν τουλάχιστον μία πτώση μέσα σε χρονικό διάστημα ενός χρόνου [1,2]. Το 55% των τραυματισμών από πτώση συμβαίνουν μέσα στο σπίτι , ενώ ένα 23% συμβαίνει έξω από το σπίτι , αλλά κοντά σε αυτό.

Για να διευκολυνθεί η ασφάλεια και η συνεχής εποπτεία (κυρίως των ηλικιωμένων) , ενώ επιτρέπει στο άτομο να διατηρήσει την ικανότητα του να ζεί στο σπίτι του , έχουν αναπτυχθεί διάφοροι μέθοδοι παρακολούθησης του χρήστη απομακρυσμένα. Μία από αυτές τις μεθόδους είναι η ανίχνευση πτώσης.

2.2 ΑΡΧΕΣ ΚΑΙ ΜΕΘΟΔΟΙ

Κατά καιρούς έχουν χρησιμοποιηθεί διάφοροι μέθοδοι , στην προσπάθεια ανάπτυξης αλγορίθμων για την ανίχνευση πτώσης :

- Με την βοήθεια επιταχυνσιόμετρων από τους Lord και Colvin το 1991 [3].
- Με επιταχυνσιόμετρα από τον Williams το 1998 [4], όπου μία συσκευή τοποθετημένη στην ζώνη ανίχνευε την πρόσκρουση με το έδαφος , και με την βοήθεια μιας βέργας με υδράργυρο , ανίχνευε αν το άτομο ήταν ξαπλωμένο στο έδαφος.
- Ο Hwang [5], τοποθέτησε ένα τριαξονικό επιταχυνσιόμετρο και ένα γυροσκόπιο στο στήθος.

- Ο Mihailidis [6], από το πανεπιστήμιο του Toronto , τοποθέτησε μία βιντεοκάμερα στο ταβάνι και ανέπτυξε αλγορίθμους για την επεξεργασία των frames από το video.
- Και πολλές άλλες προσπάθειες που βασίζονταν κυρίως σε τριαξονικά επιταχυνσιόμετρα , σε γυροσκόπια ή στην επεξεργασία της εικόνας με χρήση video.

Κατά κόρον χρησιμοποιούνται επιταχυνσιόμετρα , γυροσκόπια , επεξεργασία εικόνας από video , με την βοήθεια της τεχνολογίας GPS , BlueTooth , Internet , Δίκτυα κινητής τηλεφωνίας κτλ. (ΠΑΡΑΡΤΗΜΑ Α) .

Οι μέθοδοι για την ανίχνευση πτώσης βασίζονται κυρίως σε αναλυτικά μοντέλα ή με χρησιμοποίηση τεχνικών μάθησης μηχανής.

❖ *ΑΝΑΛΥΤΙΚΟΙ ΜΕΘΟΔΟΙ*

1. Η απλούστερη προσέγγιση είναι να ανιχνεύει τη θέση του ατόμου οριζόντια με το έδαφος. Αυτή η μέθοδος είναι κατάλληλη για την παρακολούθηση ενός εν ώρα εργασίας αλλά όχι για την ανίχνευση πτώσεων ενός ηλικιωμένου ατόμου στο περιβάλλον του σπιτιού του , αφού οι ώρες ύπνου είναι άγνωστες. Η μέθοδος αυτή είναι επιρρεπής σε πολλά “FALSE-POSITIVES” (λανθασμένα δείχνει πτώση). Μία λύση είναι η χρήση αισθητήρων στα πλακάκια. Το κόστος όμως είναι μεγάλο και υπάρχει η πιθανότητα να έχουμε πτώση σε σημείο όπου δεν υπάρχουν αισθητήρες.

2. Η έλλειψη κίνησης μπορεί να χρησιμοποιηθεί για την ανίχνευση πτώσης με την βοήθεια ενός αισθητήρα κίνησης/δόνησης τοποθετημένο σε ένα από τα κινητά άκρα του σώματος του χρήστη. Ένα αρνητικό αυτής της προσέγγισης είναι ότι πρέπει η απόφαση για την πτώση να καθυστερήσει λόγω του ότι μπορεί να υπάρχει έλλειψη κίνησης από φυσιολογικά αίτια.
3. Κατά την περίοδο μιας πτώσης , υπάρχει ένα διάστημα όπου έχουμε “ελεύθερη πτώση” . Η κατακόρυφη ταχύτητα αυξάνει γραμμικά με τον χρόνο. Αν μετρήσουμε την κατακόρυφη ταχύτητα σε ελεγχόμενες κινήσεις (να καθίσει , να λυγίσει κτλ.) , τότε μπορούμε να διακρίνουμε τις περιπτώσεις όπου έχουμε πτώση , αν η ταχύτητα ξεπεράσει ένα όριο. Η δυσκολία έγκειται στην επιλογή του ορίου αυτού.

Αν είναι χαμηλό το όριο τότε έχουμε πολλά “FALSE-POSITIVES”.

Αν είναι υψηλό το όριο τότε έχουμε πολλά “FALSE-NEGATIVES”.

Επίσης εξαρτάται από άτομο σε άτομο , και τις συνήθειες αυτού.
4. Μέθοδος με επεξεργασία της εικόνας από κάποιο video , ανιχνεύοντας την ξαπλωμένη θέση του ατόμου με scene-analysis , ή ανιχνεύοντας απότομες κινήσεις χρησιμοποιώντας vector-analysis.

❖ ΜΕΘΟΔΟΙ ΜΗΧΑΝΙΚΗΣ ΜΑΘΗΣΗΣ

Οι μέθοδοι αυτοί απαιτούν μάθηση (χρήση νευρωνικών δικτύων) , αλλά μπορεί να πάρει πάρα πολύ χρόνο η μάθηση ανάλογα και με την συχνότητα των πτώσεων.

2.3 ΑΞΙΟΛΟΓΗΣΗ ΤΩΝ ΑΙΣΘΗΤΗΡΩΝ ΠΤΩΣΗΣ

Δυστυχώς μέχρι τώρα δεν υπάρχει ένα κοινό πλαίσιο δημιουργίας και μία κοινή διαδικασία για την αξιολόγηση των αλγορίθμων ανίχνευσης πτώσης. Πρέπει να θεσμιστούν κάποια κριτήρια για την μελλοντική αξιολόγηση των αισθητήρων πτώσης.

Σε αυτήν την παράγραφο θα παρουσιάσουμε τα κριτήρια ποιότητας και ένα κοινό πλαίσιο που πρότείνει ο N. Noury [7] για τα πειράματα που πρέπει να γίνουν.

A. ΚΡΙΤΗΡΙΑ ΠΟΙΟΤΗΤΑΣ

Η έξοδος ενός ανιχνευτή πτώσης είναι δυαδική. Έχουμε ΠΤΩΣΗ ή ΜΗ ΠΤΩΣΗ. Δεν αρκεί μόνο μία δοκιμή . Πρέπει να διεξαχθεί μία στατιστική ανάλυση σε μία σειρά δοκιμών.

Υπάρχουν 4 πιθανές περιπτώσεις :

1. True Positive (TP) : Αληθώς δείχνει πτώση.
2. False Positive (FP) : Ψευδώς δείχνει πτώση.
3. True Negative (TN) : Αληθώς δείχνει Μή πτώση.
4. False Negative (FN) : Ψευδώς δείχνει Μή πτώση.

Χρησιμοποιούνται 2 κριτήρια :

- Sensitivity (ευαισθησία) → Η ικανότητα να ανιχνεύει ορθά τις πτώσεις.

$$\text{Sensitivity} = \frac{TP}{TP+FN} \quad (1)$$

- Specificity (εξειδίκευση) → Η ικανότητα να ανιχνεύει ορθά τις Μή πτώσεις.

$$\text{Specificity} = \frac{TN}{TN+FP} \quad (2)$$

B. ΣΥΝΙΣΤΩΜΕΝΑ ΠΕΙΡΑΜΑΤΑ

Γενικά τα σενάρια πτώσεων και μή πτώσεων είναι πάρα πολλά. Σύμφωνα με μία μελέτη από την SignalQuest [8] οι πτώσεις των ατόμων τρίτης ηλικίας προέρχονται από :

- 9% από απώλεια συνείδησης.
- 39% ενώ περπατούσαν.
- 20% ενώ ανέβαιναν ή κατέβαιναν σκάλες.
- 24% κατά την μεταφορά προς ή από το κρεβάτι ή την καρέκλα.
- 13% σε προσπάθεια στροφής ή έκτασης του σώματος.
- 12% ενώ έκανε κάποια υψηλού κινδύνου δραστηριότητα (όρθιος πάνω στην καρέκλα , τρέξιμο κτλ.).

Οι πτώσεις χωρίστηκαν σε τρεις μεγάλες κατηγορίες :

- Πτώσεις προς τα εμπρός (πχ. σκοντάφτει σε ένα εμπόδιο κατά το βάδισμα).

- Πτώσεις προς τα πίσω (μετά από γλίστρημα).
- Πτώσεις πλαγίως (μετά από προσπάθεια να σηκωθεί από στάση καθίσματος).

Υπάρχουν επίσης πολλές καθημερινές δραστηριότητες , όπου ή ένταση της κίνησης είναι παρόμοια με εκείνες τις καταστάσεις όπου έχουμε πτώση (η προσπάθεια να ξαπλώσουμε ή να καθίσουμε γρήγορα , ή περιπτώσεις πτώσεων με ανάκτηση κτλ.).

Εν τέλει προτάθηκαν 20 διαφορετικά σενάρια τα οποία φαίνονται στον παρακάτω πίνακα (ΠΙΝΑΚΑΣ 1).

Κατηγορία	Κατάσταση μετά την πτώση - Δραστηριότητα	Αποτέλεσμα
Πτώση προς τα πίσω	Καθιστός	POSITIVE
Πτώση προς τα πίσω	Ξαπλωμένος	POSITIVE
Πτώση προς τα πίσω	Πλάγια θέση	POSITIVE
Πτώση προς τα πίσω	Με ανάκτηση	NEGATIVE
Πτώση προς τα εμπρός	Στα γόνατα	POSITIVE
Πτώση προς τα εμπρός	Με προστασία του βραχίονα	POSITIVE
Πτώση προς τα εμπρός	Ξαπλωμένος	POSITIVE
Πτώση προς τα εμπρός	Σε πλευρική δεξιά θέση	POSITIVE

Πτώση προς τα εμπρός	Σε πλευρική αριστερή θέση	POSITIVE
Πτώση προς τα εμπρός	Με ανάκτηση	NEGATIVE
Πλάγια πτώση προς τα δεξιά	Ξαπλωμένος	POSITIVE
Πλάγια πτώση προς τα δεξιά	Με ανάκτηση	NEGATIVE
Πλάγια πτώση προς τα αριστερά	Ξαπλωμένος	POSITIVE
Πλάγια πτώση προς τα αριστερά	Με ανάκτηση	NEGATIVE
Συγκοπή	Κάθετη ολίσθηση σε έναν τοίχο , σε καθιστή θέση στο τέλος	NEGATIVE
Φυσιολογική	Κάθεται σε καρέκλα και μετά σηκώνεται	NEGATIVE
Φυσιολογική	Ξαπλώσει σε κρεβάτι και μετά σηκώνεται	NEGATIVE
Φυσιολογική	Περπατάει	NEGATIVE
Φυσιολογική	Σκύβει για να πιάσει κάτι από το έδαφος και σηκώνεται	NEGATIVE

Φυσιολογική	Βήχει ή φτερνίζεται	NEGATIVE
-------------	---------------------	----------

ΠΙΝΑΚΑΣ 1 – Σενάρια για την αξιολόγηση των ανιχνευτών πτώσης

Το 50% των σεναρίων είναι Negative και το άλλο 50% Positive.

Για κάθε σενάριο πρέπει να γίνουν περισσότερες από μία δοκιμές. Καλό είναι να μεταβάλλεται η σειρά των δοκιμών , ώστε να μην έχουμε εξοικείωση με μία συγκεκριμένη κίνηση με αποτέλεσμα να γίνει η κίνηση λιγότερο φυσικά.

Θα πρέπει να γίνει διαχωρισμός ανδρών / γυναικών για καλύτερα αποτελέσματα.

Οι δοκιμές πρέπει να γίνουν κυρίως σε ηλικιωμένους αλλά μπορεί να γίνει και προσομοίωση από νεαρά άτομα ή αθλητές.

Παρατηρούμε πώς με πολλές δοκιμές ανά σενάριο και δοκιμάζοντας πολλές συσκευές σε διαφορετικά άτομα , θα χρειαστούν πάρα πολλές δοκιμές συνολικά.

Στο τέλος τα αποτελέσματα των κριτηρίων Sensitivity και Specificity θα υποδείξουν τον καλύτερο αλγόριθμο ανά περίπτωση. Το ιδανικό σύστημα ανίχνευσης θα πρέπει να παρουσιάζει Sensitivity και Specificity = 100% .

2.4 ΑΛΓΟΡΙΘΜΟΙ ΑΝΙΧΝΕΥΣΗΣ ΠΤΩΣΗΣ

Στην παράγραφο αυτή θα μελετήσουμε και θα αναλύσουμε συγκεκριμένους αλγόριθμους και τα πειραματικά αποτελέσματα αυτών , που έχουν προταθεί από ακαδημαϊκούς και επιστημονικές ομάδες.

2.4.1 ΑΝΙΧΝΕΥΤΗΣ ΠΤΩΣΗΣ ΒΑΣΙΣΜΕΝΟΣ ΣΕ ΤΡΙΑΞΟΝΙΚΟ

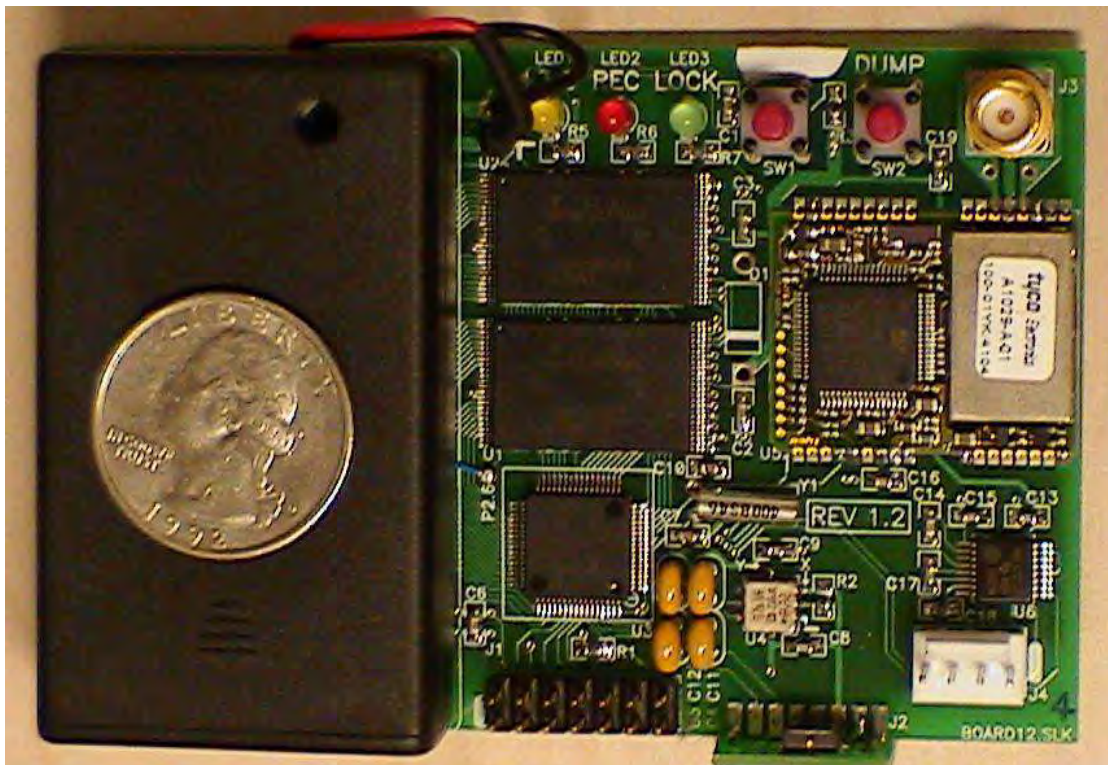
ΕΠΙΤΑΧΥΝΣΙΟΜΕΤΡΟ ΑΠΟ ΤΟΝ GARRETT BROWN.

Θα αναλύσουμε έναν αισθητήρα ανίχνευσης πτώσης, που προτάθηκε από τον Garrett Brown, με την χρήση επιταχυνσιομέτρων [9], με τιμές επιταχύνσεων που αντιστοιχούν στην κίνηση του σώματος του χρήστη της συσκευής.

➤ ΕΞΟΠΛΙΣΜΟΣ

Η συσκευή ανίχνευσης πτώσης θα τοποθετηθεί στο ισχίο του χρήστη με την μορφή μιας ζώνης μέσης, θα λειτουργεί ασύρματα (Bluetooth), και μέσω κινητού τηλεφώνου και της βοήθειας του Internet θα στέλνει το σήμα πτώσης σε κάποιο νοσοκομείο. Η μέση είναι το πιο σίγουρο σημείο για την παρακολούθηση της κίνησης σε περίπτωση πτώσης.

Ο αισθητήρας πτώσης χρησιμοποιεί επιταχυνσιόμετρα δεύτερης γενιάς “GPSADXL” , που βασίζονται σε μικροελεγκτή που καταγράφει τα δεδομένα και ενσωματώνει 2 +/- 10g MEMS ADXL210 επιταχυνσιόμετρα με μία μονάδα GPS και 4 MB στατικής μνήμης RAM. Η συσκευή τροφοδοτείται από τρεις AAA 1,2 volt NiMH επαναφορτιζόμενες μπαταρίες. Η μπαταρία διαρκεί περίπου 4 ώρες. Η συσκευή φαίνεται στην εικόνα 1.



ΕΙΚΟΝΑ 1. Πηγή: (Brown,2005) at [9]

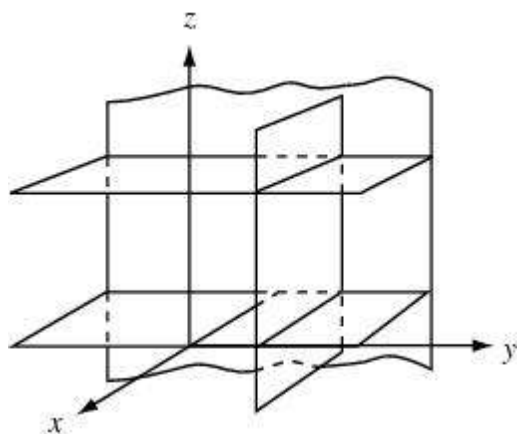
Τα επιταχυνσιόμετρα MEMS τοποθετούνται με γωνία 90° μεταξύ τους , ώστε να επιτρέπουν την μέτρηση της επιτάχυνσης σε τρεις διαστάσεις και παρέχουν δεδομένα με ρυθμό 80 δείγματα ανά sec.

Η μονάδα GPS δίνει πληροφορίες για την ώρα και το γεωγραφικό πλάτος και μήκος , το “lock status” και τον αριθμό των δορυφόρων που βλέπει. Το GPS λειτουργεί μόνο εκτός σπιτιού

Τα δεδομένα στέλνονται σε ένα PC , laptop ή κινητό μέσω της τεχνολογίας Bluetooth.

➤ ΠΑΡΑΜΕΤΡΟΙ ΑΛΓΟΡΙΘΜΩΝ

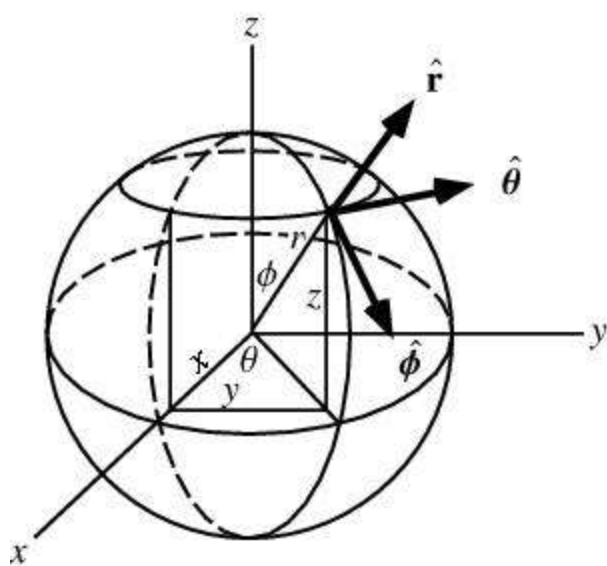
Καρτεσιανές συντεταγμένες σε 3 διαστάσεις-άξονες X , Y , Z που παίρνουν τιμές από $(-\infty , \infty)$, όπως φαίνεται και στην εικόνα 2.



Εικόνα 2. Πηγή: (Brown,2005) at [9]

Χρήση σφαιρικών συντεταγμένων για την περιγραφή της θέσης σε μία σφαίρα.

Ορίζουμε ως Θ , την οριζόντια γωνία στο επίπεδο XY από τον άξονα X με $0 < \Theta < 2\pi$, Φ η κατακόρυφη γωνία από τον άξονα Z με $0 \leq \Phi \leq \pi$ και r είναι η απόσταση (ακτίνα) από ένα σημείο ως την αρχή των αξόνων. (Εικόνα 3)



Εικόνα 3. Πηγή: (Brown,2005) at [9]

Ισχύουν οι εξής σχέσεις :

$$r = \sqrt{x^2 + y^2 + z^2} .$$

$$\Theta = \arctan (y / x) .$$

$$\Phi = \arcsin (\sqrt{x^2 + y^2} / r) = \arccos (z / r) .$$

Από άποψη καρτεσιανών συντεταγμένων ισχύουν οι εξισώσεις :

$$x = r \cos(\Theta) \sin(\Phi) .$$

$$y = r \sin(\Theta) \sin(\Phi) .$$

$$z = r \cos(\Phi) .$$

Dot product [10] μεταξύ 2 διανυσμάτων X και Y είναι το μήκος της προβολής του X πάνω στο μοναδιαίο διάνυσμα Y , όταν τα 2 διανύσματα τοποθετούνται έτσι ώστε οι ουρές τους να συμπίπτουν.

$X \cdot Y = | X | | Y | \cos (\Theta)$, όπου Θ είναι η γωνία μεταξύ των διανυσμάτων X και Y.

➤ Calibration και ανάλυση δεδομένων

Πριν οποιοσδήποτε αλγόριθμος μπορεί να αναλύσει τα δεδομένα , είναι σημαντικό αυτά τα δεδομένα να ερμηνεύονται σωστά. Για να γίνει αυτό είναι απαραίτητο να γίνει calibrate στην συσκευή για να πάρει τις αρχικές τιμές.

Ο αλγόριθμος για το calibration είναι ο εξής :

1. Κρατήστε την συσκευή για 5 sec σε όρθια θέση. Οι τιμές x , y =0.
2. Σε αυτά τα 5 sec το z παίρνει την σωστή τιμή “όρθιας θέσης” , υπολογίζοντας τον μέσο όρο των τιμών του Z.

3. Υπολόγισε τον μέσον όρο των x,y για να πάρεις τις αρχικές μηδενικές τιμές x και y .
4. Γυρίζουμε την συσκευή ανάποδα για 5 sec. $x,y = 0$.
5. Υπολόγισε τον μέσον όρο του z για αυτό το χρονικό διάστημα για να πάρουμε τις downward τιμές.
6. Αφαιρούμε τις 2 τιμές του z που βρήκαμε και διαιρούμε διά 2 για να βρούμε το διάνυσμα βαρύτητας $1G$.
7. Παίρνουμε τον μέσο όρο από τις upright και downward τιμές του z για να πάρουμε την αρχική τιμή 0 για το z . $z = 0$.

Αν ο χρήστης τοποθέτησε σωστά την συσκευή έτσι ώστε $z,y = 0$, τότε σε όρθια στάση η μόνη τιμή που έχει επίδραση είναι το διάνυσμα βαρύτητας G στον άξονα z .

Μετά χρησιμοποιήθηκε η dot product του διανύσματος Z και G για να καθοριστεί πώς ο χρήστης είναι προσανατολισμένος σε σχέση με την αρχική όρθια θέση. Μπορούν να μετατραπούν οι τιμές των δεδομένων επιτάχυνσης σε σφαιρικές συντεταγμένες , για να έχουμε μία εικόνα του στιγμιαίου προσανατολισμού του χρήστη στον πραγματικό κόσμο.

Ένα άλλο σημαντικό ζήτημα είναι αν θα πρέπει οι επιταχύνσεις στους 3 άξονες να μελετηθούν ξεχωριστά ή συνολικά ως $r = \sqrt{x^2 + y^2 + z^2}$.

Παρακάτω θα δούμε 2 αλγορίθμους όπου ο ένας μελετά τις τιμές ξεχωριστά και ο άλλος κάνει συνολική προσέγγιση του r .

➤ ΑΛΓΟΡΙΘΜΟΙ

1. Conjoined Angle Change and Magnitude Detection

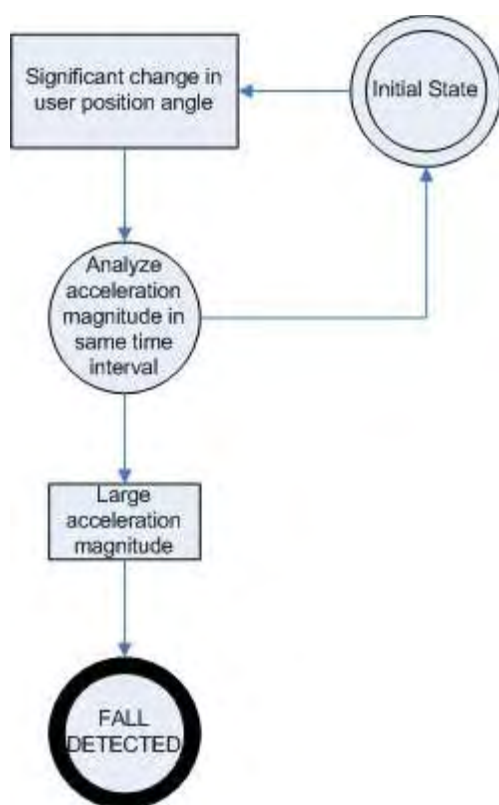
Ο πρώτος αλγόριθμος παρατηρεί τις αλλαγές στις γωνίες των διανυσμάτων, σε συνδυασμό με την απότομη αλλαγή της επιτάχυνσης εντός ενός χρονικού διαστήματος.

Μία σημαντική αλλαγή στην γωνία προσανατολισμού του χρήστη + μεγάλη αλλαγή επιτάχυνσης = ΠΤΩΣΗ .

Σε περιπτώσεις που κάποιος τρέχει , κάνει άλμα , ανεβαίνει σε μία καρέκλα κτλ. , υπάρχει αλλαγή της επιτάχυνσης αλλά η γωνία του σώματος του παραμένει σχετικά αμετάβλητη.

Ο βασικός αλγόριθμος λειτουργεί ως εξής (Εικόνα 4) :

- I. Βρές μία σημαντική αλλαγή της γωνίας μέσα σε ένα καθορισμένο χρονικό διάστημα.
- II. Αν μία σημαντική αλλαγή της γωνίας παρατηρηθεί , τότε δες αν υπάρχει στο ίδιο χρονικό διάστημα και μεγάλη αλλαγή της επιτάχυνσης.
- III. Αν και οι δύο αυτές οι ενέργειες συμβαίνουν στο ίδιο χρονικό διάστημα , τότε θεωρούμε ότι έχουμε πτώση.



Εικόνα 4 - Conjoined Angle Change and Magnitude Detection . Πηγή: (Brown,2005) at [9]

Όμως οι αλλαγές στην γωνία και στην επιτάχυνση , παρατηρούμε ότι είναι μέρος καθημερινών δραστηριοτήτων όπως το τέντωμα , η προσπάθεια να πάρει κάτι από το έδαφος ο χρήστης κτλ. .

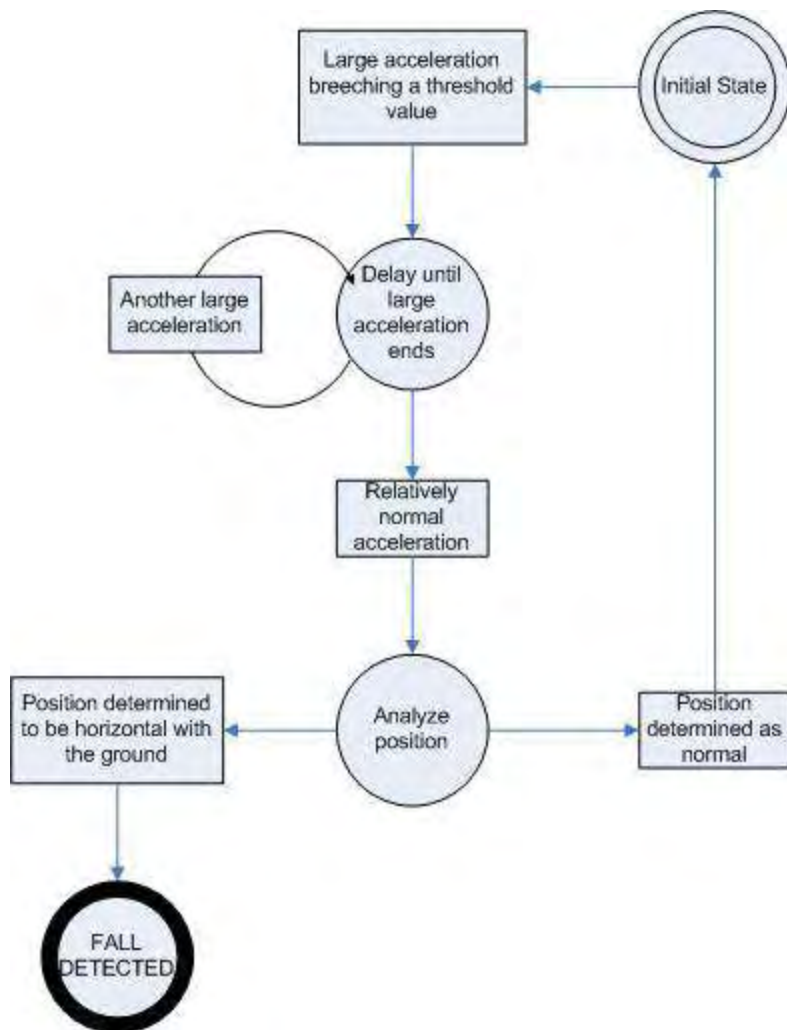
Ο αλγόριθμος θέλει βελτίωση.

2. Simple Magnitude Detection

Ο βασικός αλγόριθμος Simple Magnitude Detection (Εικόνα 5) , δουλεύει με τον εξής τρόπο :

- I. Περίμενε μέχρι η επιτάχυνση να ξεπεράσει ένα μέγιστο όριο φυσιολογικής επιτάχυνσης .

- II. Περίμενε μέχρι να τελειώσει αυτή η μεγάλη αλλαγή της επιτάχυνσης.
- III. Ανάλυση πάλι του προσανατολισμού του ατόμου.
- IV. Εάν η θέση είναι κανονική , τότε δεν έχουμε πτώση.
- V. Εάν η θέση είναι οριζόντια με το έδαφος (το άτομο είναι πεσμένο) , τότε έχουμε ανίχνευση πτώσης.



Εικόνα 5 - Simple Magnitude Detection. Πηγή: (Brown,2005) at [9]

Ο αλγόριθμος δεν θα λειτουργεί σωστά αν ο χρήστης πέσει αλλά δεν είναι παράλληλος με το έδαφος (πιάνει το κρεβάτι ή είναι πάνω σε κάποιο αντικείμενο). Είναι σημαντικό να κάνουμε διαφορετική ανάλυση για τον προσανατολισμό.

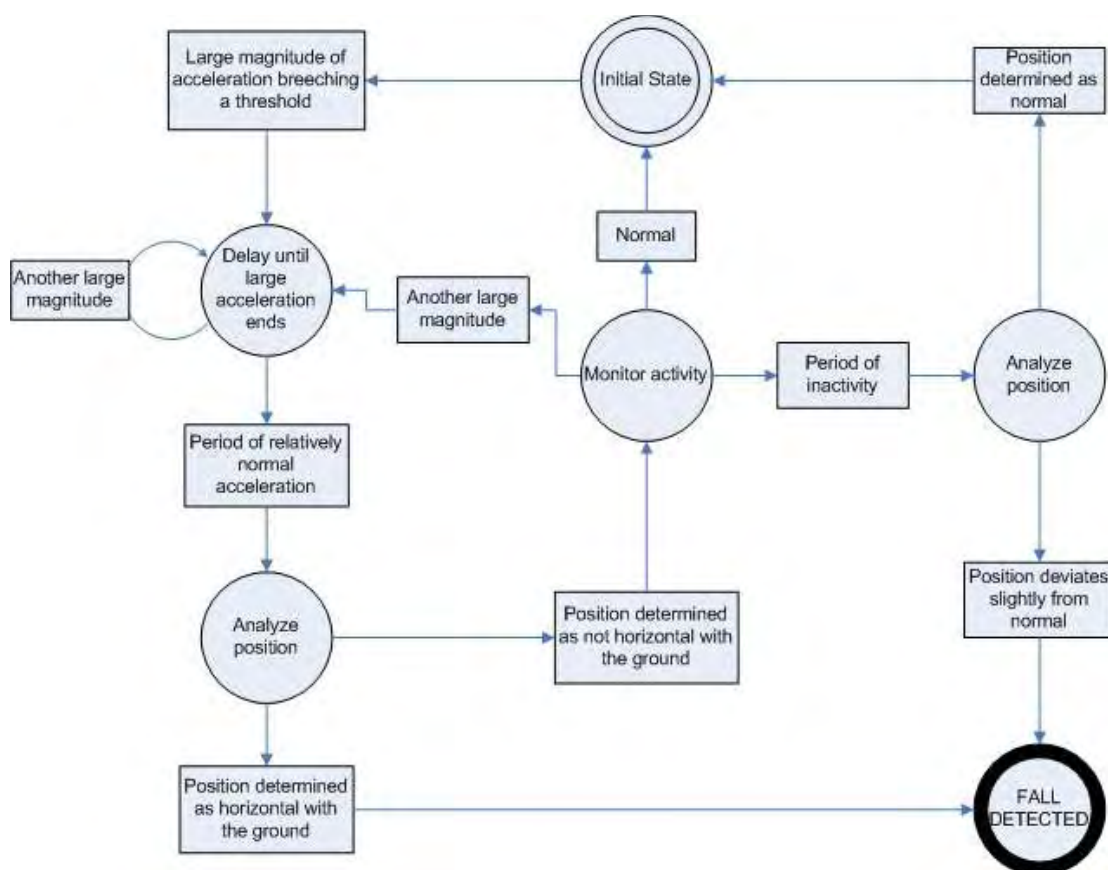
3. Advanced Magnitude Detection

Έγιναν αλλαγές στον προηγούμενο αλγόριθμο , για τις περιπτώσεις όπου το σώμα δεν είναι παράλληλο με το έδαφος μετά από την πτώση.

Αν ο χρήστης μετά από μεγάλη αλλαγή της επιτάχυνσης δεν είναι σε όρθια στάση και είναι ανενεργός για μεγάλο χρονικό διάστημα (δεν υπάρχει μεγάλη επιτάχυνση) , τότε θα αναλυθεί ξανά ο προσανατολισμός.

Ο βασικός αλγόριθμος λειτουργεί ως εξής (Εικόνα 6) :

- I. Δές για κάποια μεγάλη αλλαγή της επιτάχυνσης.
- II. Περίμενε μέχρι η επιτάχυνση να φτάσει στα κανονικά επίπεδα.
- III. Δώστε ένα σύντομο χρονικό διάστημα , περίπου 12 sec , μέχρι ο χρήστης να εγκλιματιστεί.
- IV. Ανέλυσε τον προσανατολισμό του χρήστη.
- V. Εάν είναι οριζόντια με το έδαφος τότε έχουμε πτώση.
- VI. Αν δεν είναι οριζόντια με το έδαφος αλλά έχει απόκλιση από την όρθια στάση , τότε δώσε και άλλο χρόνο στον χρήστη για να εγκλιματιστεί.
- VII. Αν μετά και από αυτόν τον χρόνο εξακολουθεί να έχει απόκλιση από την όρθια στάση , τότε έχουμε πτώση.



Εικόνα 6 - Advanced Magnitude Detection. Πηγή: (Brown,2005) at [9]

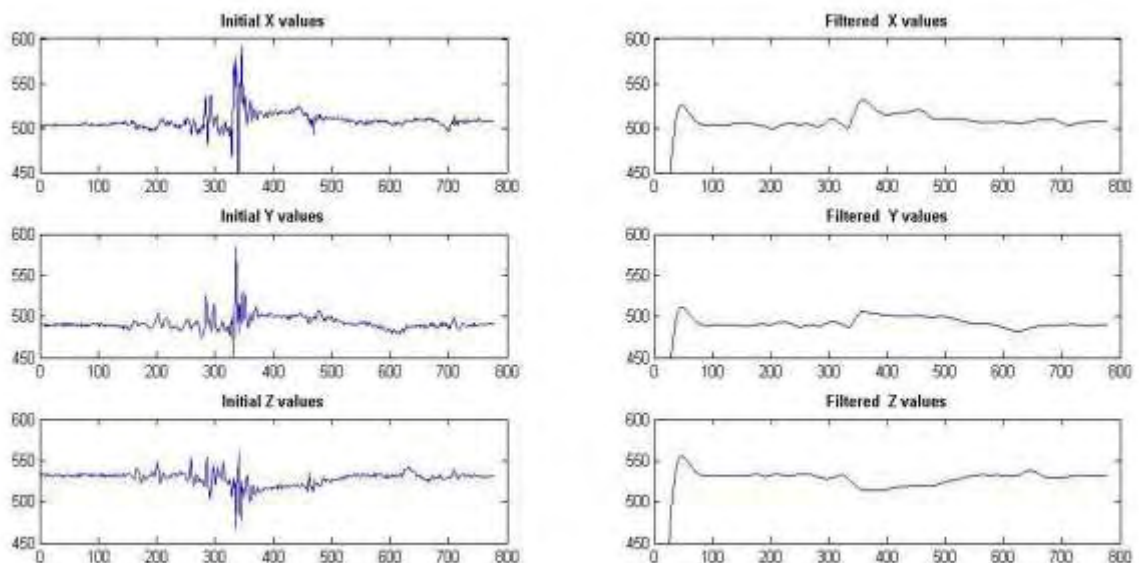
➤ ΔΟΚΙΜΕΣ ΚΑΙ ΑΠΟΤΕΛΕΣΜΑΤΑ

Θα γίνει χρήση κοινών σεναρίων για να δοκιμαστούν οι αλγόριθμοι.

Ταξινομήθηκαν οι πτώσεις σε τρεις γενικές κατηγορίες :

1. *Απλές πτώσεις* : υπάρχει πτώση και το άτομο είναι οριζόντια με το έδαφος.
2. *Σύνθετες πτώσεις* : η συμπεριφορά του επιταχυνσιομέτρου είναι περίπλοκη και το άτομο δεν είναι οριζόντιο με το έδαφος.
3. *Μή πτώσεις* : False-Positives.

Για την παρατήρηση των αλλαγών στον προσανατολισμό του χρήστη , μπορεί να γίνει χρήση των σφαιρικών τιμών. Για να χρησιμοποιηθούν σωστά οι τιμές του επιταχυνσιόμετρου , οι οποίες συνεχώς μετατρέπονται στον πίνακα παρακολούθησης του προσανατολισμού του ατόμου , είναι απαραίτητο πρώτα να εξαλειφθεί ο θόρυβος από τα δεδομένα που παίρνουμε από το επιταχυνσιόμετρο. Μετά από αρκετά πειράματα και βασιζόμενοι στο να εξαλειφθεί όσο το δυνατόν περισσότερο ο θόρυβος στα δεδομένα διατηρώντας τις ενδιαφέρουσες ιδιότητες της επιτάχυνσης , χρησιμοποιήθηκε ένα low-pass ψηφιακό φίλτρο Butterworth [11], με συχνότητα αποκοπής 0.04. Όπως βλέπουμε και στην εικόνα 7 , μολονότι ο θόρυβος και οι αιχμηρές κορυφές έχουν αφαιρεθεί , η συμπεριφορά των δεδομένων επιτάχυνσης παραμένει η ίδια.



Εικόνα 7 – Οι επιπτώσεις της εφαρμογής φίλτρου στα δεδομένα. Πηγή: (Brown,2005) at [9]

ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΛΓΟΡΙΘΜΩΝ

1. Conjoined Angle Change and Magnitude Detection

Το καθορισμένο χρονικό διάστημα που χρησιμοποιήθηκε για αυτόν τον αλγόριθμο ήταν 20 κύκλοι (1/2 sec).

Συνολικά έγιναν 77 πειράματα , από τα οποία 59 ήταν πτώσεις , και 18 Μη-πτώσεις. Από τις 59 πτώσεις οι 51 ήταν απλές πτώσεις και οι 8 ήταν σύνθετες πτώσεις.

Για διαφορετικές τιμές κατωφλίου επιτάχυνσης και αλλαγής της γωνίας , τα αποτελέσματα φαίνονται στον παρακάτω πίνακα , με τις αντίστοιχες τιμές κατωφλίου.

Conjoined Angle and Magnitude Detection Time Interval = 20 cycles(1/2 sec)	Simple Falls 51	Complex Falls 8	Total Falls 59	Non-Falls 18
Large Acceleration Threshold: 2G Significant Angle Change: $\pi/3$	31	6	37	11
Large Acceleration Threshold: 1.6G Significant Angle Change: $\pi/3$	39	8	47	12
Large Acceleration Threshold: 2.6G Significant Angle Change: $\pi/3$	29	4	33	9
Large Acceleration Threshold: 2G Significant Angle Change: $\pi/6$	31	6	37	11
Large Acceleration Threshold: 1.6G Significant Angle Change: $\pi/6$	39	8	47	12
Large Acceleration Threshold: 2.6G Significant Angle Change: $\pi/6$	29	4	33	9

Πηγή: (Brown,2005) at [9]

Τα καλύτερα αποτελέσματα παρατηρήθηκαν για τιμή κατωφλίου επιτάχυνσης 1.6G – τιμή κατωφλίου αλλαγής γωνίας $\pi/3$ και τιμή κατωφλίου επιτάχυνσης 1.6G – τιμή κατωφλίου αλλαγής γωνίας $\pi/6$. Και στις 2 περιπτώσεις αναγνωρίστηκαν ακριβώς ο ίδιος αριθμός πτώσεων και μή-πτώσεων.

Για αυτήν την καλύτερη περίπτωση έχουμε:

- Sensitivity_Total= $47/59 = 0.796 = 79.6\%$
- Sensitivity_Simple_Falls= $39/51 = 0.764 = 76.4\%$
- Sensitivity_Complex_Falls= $8/8 = 1 = 100\%$
- Specificity= $6/18 = 0.333 = 33.3\%$

Ο συγκεκριμένος αλγόριθμος αν και τα πήγε πολύ καλά στις σύνθετες πτώσεις , είχε πολύ άσχημα αποτελέσματα στις απλές πτώσεις και στις μή-πτώσεις.

2. Simple Magnitude Detection

Για αυτόν τον αλγόριθμο χρησιμοποιήθηκαν οι εξής τιμές στις παραμέτρους :

- Κανονική επιτάχυνση= αρχική τιμή από το Calibration.
- Οριζόντιος προσανατολισμός με το έδαφος = InitialZ - %G .

Συνολικά για αυτόν τον αλγόριθμο έγιναν 159 πειράματα από τα οποία 123 ήταν για πτώσεις και 36 για μή-πτώσεις. Από τα 123 πειράματα για πτώσεις τα 103 ήταν απλές πτώσεις ενώ τα 20 ήταν σύνθετες πτώσεις.

Για διαφορετικές τιμές κατωφλίου επιτάχυνσης και τιμής προσανατολισμού από το έδαφος , τα αποτελέσματα φαίνονται στον παρακάτω πίνακα , με τις αντίστοιχες τιμές κατωφλίου.

Simple Magnitude Detection Normal Acceleration = Initial Calibration Values Orientation Horizontal with the Ground = (InitialZ - %G)	Simple Falls 103	Complex Falls 20	Total Falls 123	Non-Falls 36
Large Acceleration Threshold: 1/2G Ground Orientation Value: .8G	91	1	92	2
Large Acceleration Threshold: 1/2G Ground Orientation Value: .7G	92	3	95	2
Large Acceleration Threshold: 1/2G Ground Orientation Value: .6G	94	7	101	2
Large Acceleration Threshold: 1G Ground Orientation Value: .7G	91	3	94	2
Large Acceleration Threshold: 1G Ground Orientation Value: .6G	94	4	98	2
Large Acceleration Threshold: 1.1G Ground Orientation Value: .6G	94	4	98	2
Large Acceleration Threshold: 1.2G Ground Orientation Value: .6G	88	4	92	2
Large Acceleration Threshold: 1.5G Ground Orientation Value: .6G	76	3	79	2

Πηγή: (Brown,2005) at [9]

Τα καλύτερα αποτελέσματα παρατηρήθηκαν για τιμή κατωφλίου επιτάχυνσης 1/2G και τιμή προσανατολισμού από το έδαφος 0.6G.

Για αυτήν την καλύτερη περίπτωση έχουμε :

- Sensitivity_Total= $101/123 = 0.821 = 82.1\%$
- Sensitivity_Simple_Falls= $94/103 = 0.912 = 91.2\%$
- Sensitivity_Complex_Falls= $7/20 = 0.350 = 35\%$
- Specificity= $34/36 = 0.944 = 94.4\%$

Σε σχέση με τον προηγούμενο αλγόριθμο έχουμε σαφώς καλύτερα αποτελέσματα στην αναγνώριση πτώσεων και μή-πτώσεων και ιδικά στις απλές πτώσεις , αλλά έχουμε σημαντικά προβλήματα στην αναγνώριση των σύνθετων πτώσεων.

3. Advanced Magnitude Detection

Για τον αλγόριθμο αυτό χρησιμοποιήθηκαν οι εξής τιμές στις παραμέτρους :

- Κανονική επιτάχυνση = αρχική τιμή από το Calibration.
- Οριζόντιος προσανατολισμός με το έδαφος = InitialZ - .6G .
- Περίοδος αδράνειας = 160 κύκλοι (4 sec).
- Ο προσανατολισμός που παρεκκλίνει από την όρθια θέση = InitialZ - .%G .

Για αυτόν τον αλγόριθμο έγιναν ακριβώς τα ίδια πειράματα με τον προηγούμενο (159 συνολικά).

Για διαφορετικές τιμές κατωφλίου επιτάχυνσης , κανονικής τιμής επιτάχυνσης και τιμής απόκλισης από την όρθια στάση , τα αποτελέσματα φαίνονται στον παρακάτω πίνακα , με τις αντίστοιχες τιμές κατωφλίου.

Advanced Magnitude Detection Normal Acceleration = Initial Calibration Values Orientation Horizontal with the Ground = (InitialZ - .6G) Period of Inactivity Length = 160 cycles Orientation Designated as Deviating from Uprightness = (InitialZ - .5G)	Simple Falls 103	Complex Falls 20	Total Falls 123	Non-Falls 36
Large Acceleration Threshold: 1.1G Normal Acceleration Threshold: 1.2G Upright Deviation Orientation Value: .1G	96	17	113	2
Large Acceleration Threshold: 1.2G Normal Acceleration Threshold: 1.2G Upright Deviation Orientation Value: .1G	96	20	116	2
Large Acceleration Threshold: 1.24G Normal Acceleration Threshold: 1.1G Upright Deviation Orientation Value: .1G	96	16	112	2
Large Acceleration Threshold: 1.24G Normal Acceleration Threshold: 1.2G Upright Deviation Orientation Value: .1G	96	20	116	2
Large Acceleration Threshold: 1.24G Normal Acceleration Threshold: 1.2G Upright Deviation Orientation Value: .2G	94	20	114	2
Large Acceleration Threshold: 1.24G Normal Acceleration Threshold: 1.3G Upright Deviation Orientation Value: .1G	96	20	116	2
Large Acceleration Threshold: 1.31G Normal Acceleration Threshold: 1.2G Upright Deviation Orientation Value: .1G	96	19	115	2
Large Acceleration Threshold: 1.41G Normal Acceleration Threshold: 1.2G Upright Deviation Orientation Value: .1G	95	19	114	2

Πηγή: (Brown,2005) at [9]

Τα καλύτερα αποτελέσματα παρατηρήθηκαν για τιμή κατωφλίου επιτάχυνσης 1.24G, κανονική τιμή επιτάχυνσης 1.2G και τιμή απόκλισης από την όρθια θέση .1G .

Για αυτήν την καλύτερη περίπτωση έχουμε :

- Sensitivity_Total= 116/123 = 0.943 = 94.3%
- Sensitivity_Simple_Falls= 96/103 = 0.932 = 93.2%
- Sensitivity_Complex_Falls= 20/20 = 1 = 100%

- Specificity= $34/36 = 0.944 = 94.4\%$

Τα αποτελέσματα του αλγορίθμου αυτού είναι σαφώς καλύτερα , ιδικά στην αναγνώριση των σύνθετων πτώσεων.

2.4.2 ΣΥΓΚΡΙΣΗ ΑΛΓΟΡΙΘΜΩΝ ΓΙΑ ΑΝΙΧΝΕΥΣΗ ΠΤΩΣΗΣ ΧΑΜΗΛΗΣ

ΠΟΛΥΠΛΟΚΟΤΗΤΑΣ

Σε αυτήν την ενότητα θα παρουσιάσουμε 3 διαφορετικούς αλγορίθμους από τους Marrit Kangas , Antti Konttila , Per Lindgren ,Ilkka Winblad και Timo Jamsa [12], με χρήση τριαξονικών επιταχυνσιομέτρων ,που χρησιμοποιήθηκαν στην μέση , τον καρπό και το κεφάλι.

Κάθε αλγόριθμος χρησιμοποιεί τουλάχιστον δύο από τις παρακάτω φάσεις-παραμέτρους μίας πτώσης :

- Ξεκίνημα της πτώσης.
- Ταχύτητα πτώσης.
- Σύγκρουση κατά την πτώση.
- Στάση του σώματος μετά την πτώση.

ΥΛΙΚΑ ΚΑΙ ΕΠΕΞΕΡΓΑΣΙΑ ΔΕΔΟΜΕΝΩΝ

Οι επιταχύνσεις μετρήθηκαν συγχρόνως στην μέση , τον καρπό και το κεφάλι με χρήση τριαξονικών επιταχυνσιομέτρων.

Κάθε τριαξονικό επιταχυνσιόμετρο (VTI Hamlin SCA CDCV1G, amplitude range +/- 12g) συνδέθηκε με έναν ξεχωριστό καταγραφέα δεδομένων (Tattletale, model 8v2, Onset Computer Corp.) με συχνότητα δειγματοληψίας τα 400Hz.

Τα επιταχυνσιόμετρα τοποθετήθηκαν στον μή κυρίαρχο καρπό , στην μέση και στο μέτωπο του προσώπου.

Το σήμα επιτάχυνσης επεξεργαζόταν με επαναδειγματοληψία στα 50 HZ και διάμεσο φιλτράρισμα ανά 3 δείγματα , για να μειωθεί η ποσότητα των δεδομένων και ο θόρυβος , πριν από οποιαδήποτε άλλη ανάλυση των δεδομένων.

Μετά τα δεδομένα αυτά , φιλτράρονταν ($f_c = 0.25$ Hz) με low-pass (LP) ή high-pass (HP) , με ένα ψηφιακό δεύτερης τάξης Butterworth φίλτρο [11], για ανίχνευση της στάσης ή για δυναμική ανάλυση αντίστοιχα.

- LP φιλτράρισμα → ανίχνευση στάσης.
- HP φιλτράρισμα → δυναμική ανάλυση.

ΠΑΡΑΜΕΤΡΟΙ ΑΛΓΟΡΙΘΜΩΝ

Στους αλγορίθμους που ακολουθούν χρησιμοποιήθηκαν οι εξής παράμετροι :

Το διάνυσμα SV_{TOT} που περιέχει το συνολικό ποσό των δυναμικών και στατικών συνιστωσών επιτάχυνσης.

$$SV_{TOT} = \sqrt{(A_x)^2 + (A_y)^2 + (A_z)^2} \quad (1)$$

Όπου A_x, A_y, A_z είναι η επιτάχυνση στους άξονες x,y,z αντίστοιχα.

Όταν καθόμαστε $SV_{TOT} \cong 1g$ Εικόνα 8 (A)

Η αρχή της πτώσης προσδιορίστηκε όταν το $SV_{TOT} \leq 0.6g$.

Το άθροισμα των δυναμικών διανυσμάτων (SV_D) υπολογίστηκε από τον ίδιο τύπο μετά από HP φιλτράρισμα των δεδομένων.

$$SV_D = \sqrt{(A_x)^2 + (A_y)^2 + (A_z)^2} \quad (2)$$

$SV_D = 0g$ όταν καθόμαστε Εικόνα 8 (A) .

Το SV_D χρησιμοποιήθηκε για να ανιχνευτεί η πρόσκρουση μετά από πτώση.

Οι γρήγορες αλλαγές της επιτάχυνσης ελέγχονται με το αθροιστικό διάνυσμα SV_{MAXMIN} , όπου υπολογίζεται χρησιμοποιώντας την διαφορά μεταξύ της maximum και της minimum τιμής , σε χρονικό διάστημα 0.1 sec για κάθε άξονα ξεχωριστά.

Όταν καθόμαστε $SV_{MAXMIN} = 0g$.

Η κατακόρυφη επιτάχυνση Z_2 υπολογίζεται ως :

$$Z_2 = \frac{SV_{TOT}^2 - SV_D^2 - G^2}{2G} \quad (3) \quad , \text{ όπου } G \text{ η βαρυτική συνιστώσα , με } G= 1g.$$

Η ταχύτητα U_0 , υπολογίζεται με ολοκλήρωση της SV_{TOT} , από την αρχή της πτώσης έως την σύγκρουση , όπου η τιμή της SV_{TOT} είναι μικρότερη από 1g , $SV_{TOT} < 1g$.

Η στάση του σώματος ανιχνεύεται , 2 sec μετά την πρόσκρουση , από το low-pass φιλτραρισμένο κατακόρυφο σήμα , με βάση την μέση επιτάχυνση σε ένα χρονικό διάστημα 0.4 sec. Με τιμή $\leq 0.5g$ θεωρείται ότι ο χρήστης βρίσκεται σε ξαπλωμένη-οριζόντια θέση.

Για σύγκριση με μία προηγούμενη μελέτη, θα αναφέρουμε τον δείκτη πτώσης (falling index (FI)) , που είχε περιγραφεί από τον Yoshida [13], χρησιμοποιώντας ένα χρονικό περιθώριο 0.4 sec για μία συσκευή τοποθετημένη στην μέση :

$$FI_i = \sqrt{\sum_{i-19}^i ((A_x)_i - (A_x)_{i-1})^2 + \sum_{i-19}^i ((A_y)_i - (A_y)_{i-1})^2 + \sum_{i-19}^i ((A_z)_i - (A_z)_{i-1})^2}$$

(4) ,

όπου $(A_x)_i$, $(A_y)_i$, $(A_z)_i$, είναι οι επιταχύνσεις (σε g) , του αριθμού του δείγματος i , στους άξονες x,y,z αντίστοιχα.

ΑΛΓΟΡΙΘΜΟΙ

Θα μελετήσουμε 3 διαφορετικούς αλγόριθμους με αυξανόμενη πολυπλοκότητα όπου ο κάθενας χρησιμοποιεί 2 ή περισσότερες από τις παρακάτω φάσεις :

- Αρχή της πτώσης → START OF FALL .
- Ταχύτητα πτώσης → VELOCITY.
- Πρόσκρουση → IMPACT.
- Στάση του σώματος μετά την πτώση → POSTURE.

Τα όρια κατωφλίου που χρησιμοποιήθηκαν για κάθε παράμετρο φαίνονται στον παρακάτω πίνακα , και επιλέχθηκαν για μέγιστο δυνατό Sensitivity και Specificity [14].

Παράμετρος	Μέση	Κεφάλι	Καρπός
SVTOT (g)	2.0	2.0	5.2
SVD (g)	1.7	1.2	5.1
Z2 (g)	1.5	1.8	3.9

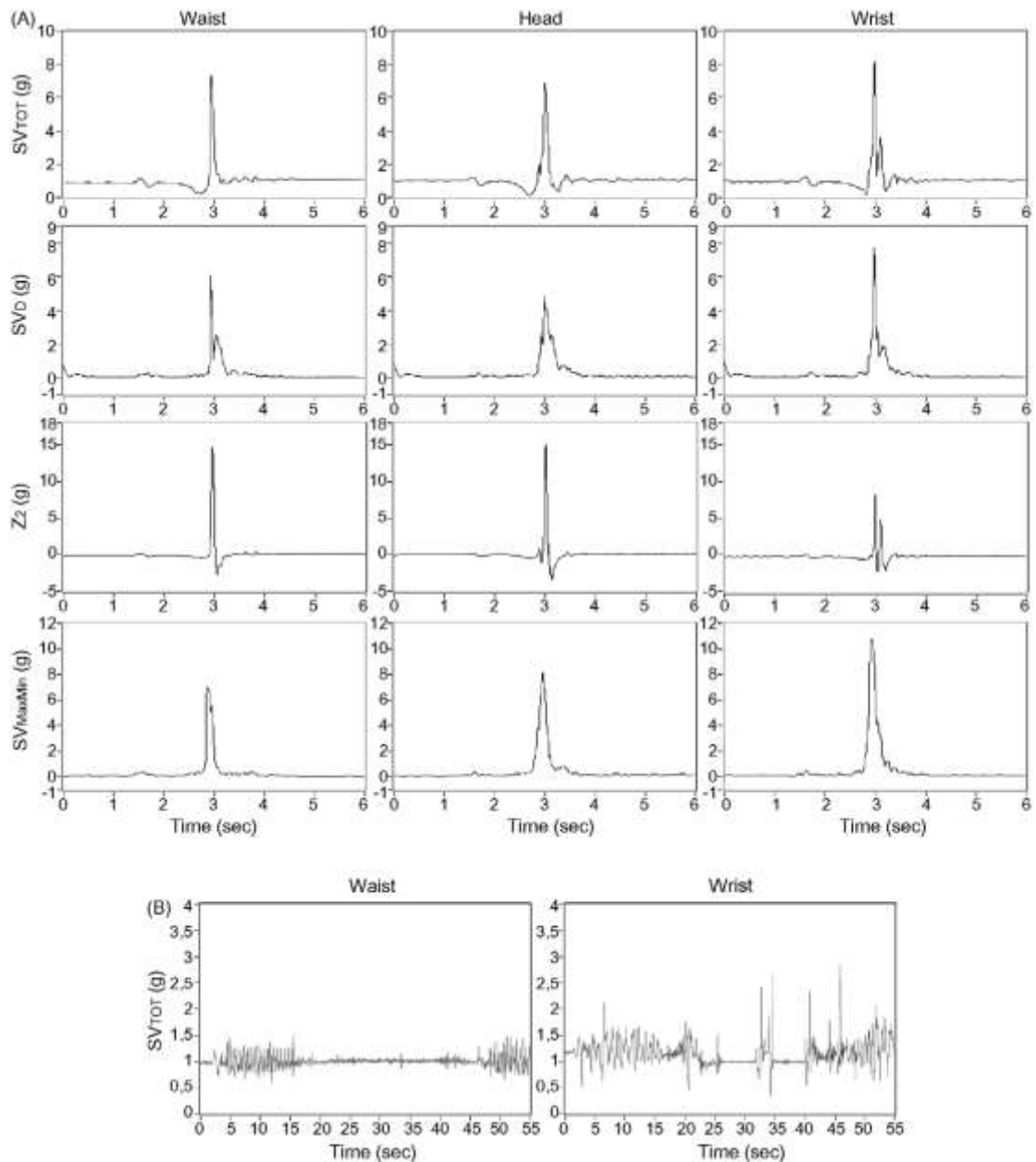
SVMaxMin (g)	2.0	1.7	6.5
v0 (m / s)	0.7	1.0	0.9
SVDS pit (g)	0.6	0.6	0.6

Η ανίχνευση της στάσης του σώματος μετά την πτώση , δεν έγινε για την περίπτωση του καρπού.

Στην Εικόνα 8 (A) και (B) φαίνονται μερικά δείγματα για διαφορετικές παράμετρους επιτάχυνσης , για μία πτώση προς τα εμπρός και για καθημερινές δραστηριότητες αντίστοιχα.

Αλγόριθμοι :

1. Αλγόριθμος 1 (IMPACT + POSTURE) : βασίστηκε στην ανίχνευση της σύγκρουσης από μία τιμή κατωφλίου του SVTOT , SVD ,SVMaxMin ή Z2 , που ακολουθείται από παρακολούθηση της στάσης του σώματος του ατόμου.
2. Αλγόριθμος 2 (START OF FALL + IMPACT + POSTURE) : Ανιχνεύει την έναρξη της πτώσης με παρακολούθηση της SVTOT , όταν $SVTOT \leq 0.6g$, ακολουθείται ανίχνευση σύγκρουσης εντός χρονικού πλαισίου 1 sec , με τιμή κατωφλίου SVTOT ή Z2 και στο τέλος παρακολούθηση της στάσης του σώματος.
3. Αλγόριθμος 3 (START OF FALL + VELOCITY + IMPACT + POSTURE) : ανιχνεύει την έναρξη της πτώσης με τον ίδιο τρόπο όπως ο αλγόριθμος 2, $SVTOT \leq 0.6g$, μετρά την ταχύτητα V0 εάν υπερβαίνει ένα όριο , ανίχνευση της σύγκρουσης από παρατήρηση των τιμών SVTOT ή Z2 και στο τέλος έχουμε παρακολούθηση της στάσης του σώματος.



ΕΙΚΟΝΑ 8 (Α) Παραδείγματα των διάφορων παραμέτρων της επιτάχυνσης , από την μέση, τον καρπό και το κεφάλι , κατά την διάρκεια μιας πτώσης προς τα εμπρός , με την σύγκρουση στο χρονικό σημείο 3 sec περίπου. (Β) Ένα παράδειγμα φυσιολογικής δραστηριότητας , που μετρήθηκε από την μέση και τον καρπό. Κάθεται σε μία καρέκλα (0 sec μέχρι 4 sec) , περπατάει (4 sec μέχρι 16 sec) , ξαπλώνει στο κρεβάτι (23 sec μέχρι 44 sec) και περπατάει (από το 49sec και μετά) . Πηγή (Kangas *et al.* , 2008) at [12]

ΑΠΟΤΕΛΕΣΜΑΤΑ

Για να παρθούν τα δεδομένα επιτάχυνσης από περιπτώσεις πτώσεων ή φυσιολογικών καθημερινών δραστηριοτήτων , έγιναν σκόπιμες πτώσεις από 3 υγιείς εθελοντές , μία γυναίκα και δύο άνδρες. Οι πτώσεις έγιναν σε στρώμα πάχους 20 cm.

Έγιναν τουλάχιστον 2 πτώσεις από τον κάθενα προς τρεις διαφορετικές κατευθύνσεις (εμπρός – πίσω - πλευρικά) όπως φαίνεται και στον επόμενο πίνακα.

No.	Κατεύθυνση	Οδηγίες	Εξαρτήματα
1	Εμπρός	Ευθεία , χωρίς βήμα	-
2	Εμπρός	Βήμα προς τα κάτω από μία πλατφόρμα	Πλατφόρμα
3	Εμπρός	Κάθονται , σηκώνονται , κάνουν ένα βήμα	Καρέκλα
4	Πίσω	Το πόδι αιωρείται προς τα	-

		εμπρός	
5	Πίσω	Στροφή του σώματος	-
6	Πίσω	Κάθονται στον αέρα	-
7	Πλευρικά	Βήμα προς τα κάτω από μία πλατφόρμα	Πλατφόρμα
8	Πλευρικά	Ευθεία χωρίς να γίνει βήμα	-

Τα αποτελέσματα (Sensitivity %) για την μέση , το κεφάλι και τον καρπό ήταν τα εξής :

	n	Αλγόριθμος 1 (a)	“	“	“	Αλγόριθμος 2 (b)	“	Αλγόριθμος 3 (c)	“
Μέση		SVTOT	SVD	SVMaXMin	Z2	SVTOT	Z2	SVTOT	Z2
F	21	100	95	100	100	100	100	71	71
B	21	90	62	76	90	86	86	71	71
L	17	100	82	94	100	100	100	88	88
All	59	97	80	90	97	95	95	76	76
	n	Αλγόριθμος 1 (a)	“	“	“	Αλγόριθμος 2 (b)	“	Αλγόριθμος 3 (c)	“
Κεφάλι		SVTOT	SVD	SVMaXMin	Z2	SVTOT	Z2	SVTOT	Z2

F	15	100	100	100	100	93	93	47	47
B	21	100	100	100	95	81	81	43	43
L	20	95	95	95	95	85	85	50	50
All	56	98	98	98	97	86	86	47	47
	n	Αλγόριθμος 1 (a)	“	“	“	Αλγόριθμος 2 (b)	“	Αλγόριθμος 3 (c)	“
Καρπός (d)		SVTOT	SVD	SVMaxMin	Z2	SVTOT	Z2	SVTOT	Z2
F	22	73	59	73	73	-----	73	-----	41
B	19	63	63	58	84	-----	68	-----	47
L	17	29	6	12	53	-----	41	-----	24
All	57	55	36	50	71	-----	64	-----	37

F = forward, B = backward, L = lateral.

(a) IMPACT + POSTURE.

(b) START OF FALL + IMPACT + POSTURE.

(c) START OF FALL + VELOCITY + IMPACT + POSTURE.

(d) Η ανίχνευση της στάσης του σώματος δεν έγινε.

Παρακάτω θα αναλύσουμε τα αποτελέσματα των αλγορίθμων για την μέση , το κεφάλι και τον καρπό :

I. Μέση :

- Το Sensitivity για όλες τις πτώσεις , κυμαίνεται από 76% έως 97% , ανάλογα τον αλγόριθμο και την παράμετρο.

- Οι πτώσεις αναγνωρίζονταν καλύτερα (97%) με τον αλγόριθμο 1 (IMPACT + POSTURE) , χρησιμοποιώντας το SVTOT ή το Z2 ως δείκτη για την σύγκρουση.
- Το SVMaXMin ανιχνεύει αποτελεσματικά τις πτώσεις εμπρός και πλευρικά (100% και 94% αντίστοιχα) , ενώ μόνο το 76% των πτώσεων προς τα πίσω ανίχνευσε.
- Το Svd είχε το χαμηλότερο sensitivity συνολικά από όλες τις παραμέτρους.
- Για τους αλγορίθμους 2 και 3 χρησιμοποιήθηκαν οι παράμετροι SVTOT και Z2.
- Το Sensitivity των αλγορίθμων 1 και 2 ήταν παρόμοιο.
- Γενικά στους αλγορίθμους 1 και 2 ανιχνεύονταν καλύτερα οι πτώσεις εμπρός και πλευρικά.
- Μέσο Sensitivity στον αλγόριθμο 3 = 76%. Ανιχνεύει κυρίως τις πλάγιες πτώσεις.
- Χρησιμοποιώντας την παράμετρο FI , είχαμε 96% Sensitivity με κατώφλι τα 3.16g , για τις πτώσεις προς τα εμπρός. Αλλά για τις πλάγιες είχαμε 71% και στις προς τα πίσω 56%. Με κατώφλι 2.6g είχαμε 74%. [14]. Τα αποτελέσματα αυτά είναι χειρότερα από τα αποτελέσματα του αλγορίθμου 1 , με παράμετρο την SVTOT ή την Z2.

II. Κεφάλι :

- Αλγόριθμος 1: Sensitivity = 97-98% , ανεξάρτητα από τις παραμέτρους που χρησιμοποιήθηκαν.

- Αλγόριθμος 2: είχε χαμηλότερο sensitivity σε σχέση με τον αλγόριθμο 1 και ανίχνευε καλύτερα τις πτώσεις προς τα εμπρός παρά τις υπόλοιπες.
- Αλγόριθμος 3: Ανίχνευσε λιγότερες από τις μισές πτώσεις.

III. Καρπός :

- Αλγόριθμος 1 : ήταν ο καλύτερος αλγόριθμος (με χρήση στον καρπό) με χρήση της παραμέτρου Z2. Το Sensitivity ήταν μεταξύ 53-84%.

Γενικά η ανίχνευση ήταν χαμηλή στις πλευρικές πτώσεις.

- Αλγόριθμος 2 : Μέσο sensitivity =64% χρησιμοποιώντας το Z2 σαν παράμετρο. Γενικά κυμαίνεται από 41% έως 73% ανάλογα με το είδος των πτώσεων.

Οι περισσότερες πτώσεις (79%) είχαν $SVTOT \leq 0.6g$, ακριβώς λίγο πριν την σύγκρουση αλλά η ανίχνευση της σύγκρουσης δεν πραγματοποιήθηκε σε όλες τις περιπτώσεις.

- Αλγόριθμος 3 : Μέσο sensitivity = 37% με παράμετρο το Z2.

Τα δύο κριτήρια $SVTOT \leq 0.6g$ και $V0 > 0.9 \text{ m/s}$ πριν την σύγκρουση εκπληρώθηκαν στο 56% , 69% και 27% για τις πτώσεις προς τα εμπρός , πίσω ή πλευρικά αντίστοιχα.

Γενικότερα καλύτερα αποτελέσματα έδωσε ο αλγόριθμος 1 (αν και ήταν ο πιο απλός) με χρήση της παραμέτρου Z2 , και καλύτερο σημείο για τοποθέτηση της συσκευής ήταν η μέση σε συνδυασμό με τον αλγόριθμο 1. Η μέση και το κεφάλι

έδωσαν γενικά καλά αποτελέσματα σε σχέση με το σημείο του καρπού , το οποίο φαίνεται ακατάλληλο για ανίχνευση πτώσης.

Οι προς τα πίσω πτώσεις, όπου δεν ανιχνεύονταν αποτελεσματικά στην μέση , οφειλόταν σε μεγάλο βαθμό στο μαλακό στρώμα όπου έπεφταν (με την μέση) οι εθελοντές.

2.4.3 ΕΝΑΣ ΑΛΓΟΡΙΘΜΟΣ ΑΝΙΧΝΕΥΣΗΣ ΠΤΩΣΗΣ ΜΕ ΧΡΗΣΗ

ΓΥΡΟΣΚΟΠΙΟΥ

Ο Najafi [15] , τοποθετώντας αισθητήρες γυροσκοπίου προσπάθησε να αξιολογήσει το ρίσκο πτώσης σε ηλικιωμένα άτομα , με μέτρηση των μεταβάσεων stand-sit και sit-stand. Στην συγκεκριμένη προσπάθεια δεν προσπαθούσε να ανιχνεύσει τις πτώσης σε πραγματικό χρόνο , αλλά μόνο να διαπιστώσει ότι ο ηλικιωμένος βρισκόταν σε κίνδυνο πτώσης.

Έναν παρόμοιο pre-impact ανιχνευτή πτώσης παρουσίασε και ο Nyan [16], τοποθετώντας αισθητήρες γυροσκοπίου σε 3 διαφορετικά σημεία του σώματος , το στέρνο , στο μπροστινό μέρος της μέσης και κάτω από το μπράτσο . Με αυτό το σύστημα κατάφερε να πετύχει sensitivity = 100% , αλλά είχε specificity μόλις 16%.

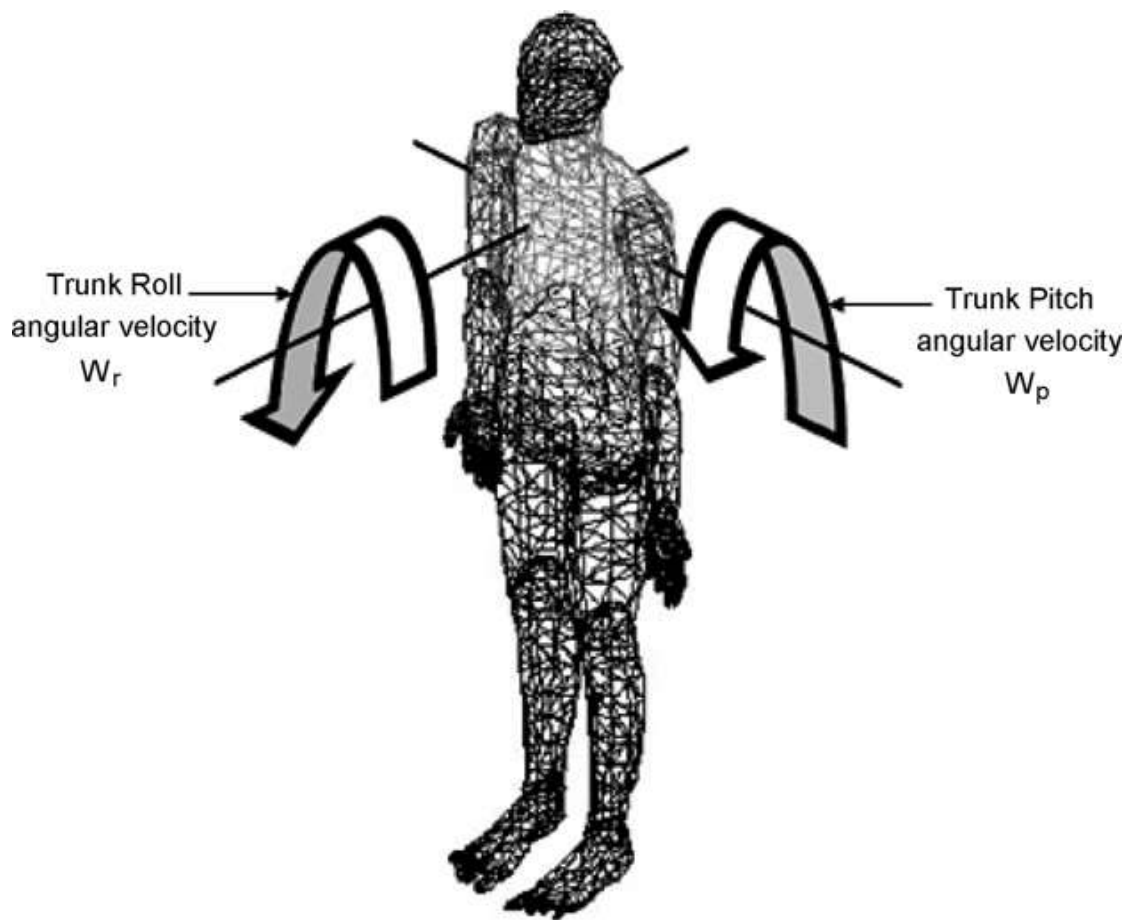
Σε αυτήν την παράγραφο θα παρουσιάσουμε έναν αλγόριθμο ανίχνευσης πτώσης και τα αποτελέσματα του, με χρήση 2D-γυροσκοπίου , από τους Bourke και Lyons [17]. Τα πειράματα για τις πτώσεις έγιναν από νεαρά άτομα , ενώ για τις καθημερινές

δραστηριότητες (ADL → Activities of Daily Living) χρησιμοποιήθηκαν ηλικιωμένα άτομα για να προσομοιώσουν καλύτερα αυτές τις καταστάσεις.

Έγινε η υπόθεση ότι όταν ένα άτομο χτυπά στο έδαφος, οι μεταβολές στην γωνιακή επιτάχυνση , γωνιακή ταχύτητα και γωνία του σώματος του θα είναι διαφορετικές από ότι είναι στις καθημερινές δραστηριότητες του.

ΥΛΙΚΑ ΚΑΙ ΜΕΘΟΔΟΙ

Για να επαληθευτεί η υπόθεση αυτή , καταγράφηκαν οι μετρήσεις του γυροσκοπίου , trunk pitch και roll (Εικόνα 9), κατά την διάρκεια διαφορετικών προσομοιώσεων πτώσης και καθημερινών δραστηριοτήτων.



ΕΙΚΟΝΑ 9 – W_r → trunk roll γωνιακή ταχύτητα , W_p → trunk pitch γωνιακή ταχύτητα . Πηγή : (Bourke και Lyons, 2006) at [17]

- *Μελέτη προσομοίωσης πτώσης* : Για την προσομοίωση των πτώσεων συμμετείχαν 10 υγιείς νεαροί άνδρες. Κάθε ένας εκτέλεσε 8 διαφορετικούς τύπους πτώσης και κάθε τύπος πτώσης επαναλήφθηκε 3 φορές από τον καθένα. Ο καθένας εκτέλεσε 24 πτώσεις.

Οι τύποι των πτώσεων που χρησιμοποιήθηκαν , επιλέχθηκαν με σκοπό να προσομοιώνουν όσο γίνεται καλύτερα τους τύπους των πτώσεων που εμφανίζονται πιο συχνά στους ηλικιωμένους :

- Πτώσεις προς τα εμπρός.
 - Πτώσεις προς τα πίσω.
 - Πλάγιες πτώσεις αριστερά και δεξιά.
 - Όλες οι παραπάνω πτώσεις προσομοιώθηκαν με τα δύο πόδια ίσια αλλά και με τα γόνατα λυγισμένα.
- *Μελέτη καθημερινών δραστηριοτήτων* : Για την προσομοίωση των καθημερινών δραστηριοτήτων συμμετείχαν 10 ηλικιωμένοι , ηλικίας από 70 έως 83 χρονών. Τρεις από αυτούς ήταν γυναίκες και επτά άνδρες.

Οι τύποι των ADL που επιλέχθηκαν ήταν εκείνες που είχαν απότομες μεταβολές στην κίνηση του ατόμου και ενδεχομένως να οδηγήσουν σε fall-positives. Κάθε άτομο επανέλαβε τις παρακάτω ADL , από τρεις φορές και ξεκινούσε και τελείωνε κάθε ADL σε όρθια θέση :

 - Κάθεται και σηκώνεται από μία πολυθρόνα.
 - Κάθεται και σηκώνεται από μία καρέκλα.
 - Κάθεται και σηκώνεται από την τουαλέτα.
 - Κάθεται και σηκώνεται από ένα χαμηλό σκαμνί.
 - Κάθεται και σηκώνεται από το κρεβάτι.
 - Ξαπλώνει και σηκώνεται από το κρεβάτι.
 - Μπαίνοντας και βγαίνοντας από το αυτοκίνητο.
 - Περπατάει για 10 μέτρα.
 - *Υλικά* : Ένα καταγραφικό Biomedical Monitoring BM42 χρησιμοποιήθηκε για την απόκτηση των δεδομένων. Τα σήματα του αισθητήρα καταγράφηκαν σε συχνότητα 1 kHz και ανάλυση 12 bits.

Το διαξονικό γυροσκόπιο κατασκευάστηκε από δύο αναλογικά ADXRS300 iMEMS γυροσκόπια , που τοποθετήθηκαν ορθογώνια μεταξύ τους.

Ο αισθητήρας γυροσκοπίου τοποθετήθηκε στο στέρνο του κάθε ατόμου όπως φαίνεται και στην παρακάτω εικόνα :



Πηγή : (Bourke και Lyons, 2006) at [17]

- *Επεξεργασία σήματος* : Κάθε pitch (Wp) και roll (Wr) γωνιακή ταχύτητα , φιλτραρίστηκε , με ένα low-pass φίλτρο χρησιμοποιώντας μιας δεύτερης τάξης low-pass ButterWorth 2-pass digital φίλτρο , με συχνότητα αποκοπής τα 100Hz.

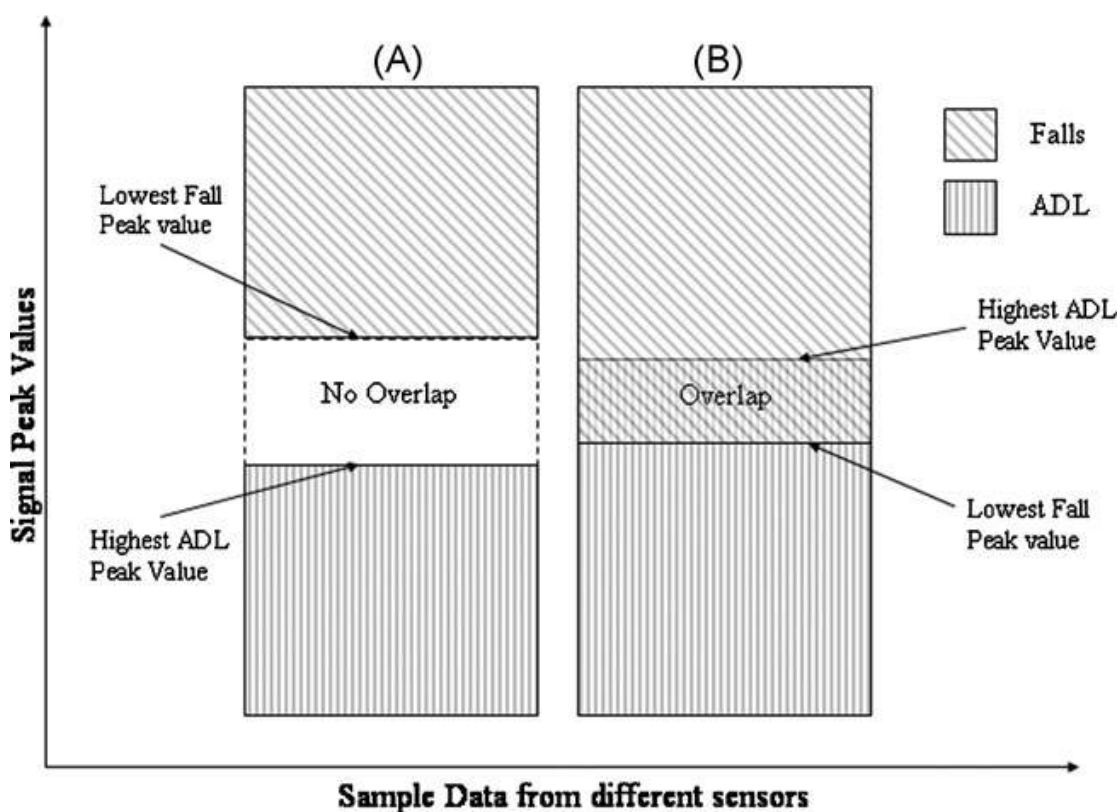
Το προκύπτων διάνυσμα για το σήμα της γωνιακής ταχύτητας W_{res} , ισούται με την ρίζα του αθροίσματος των τετραγώνων του W_p και W_r και θα αποτελέσει το συνδιασμένο μέτρο της γωνιακής ταχύτητας.

$$W_{res} = \sqrt{W_p^2 + W_r^2} .$$

ΑΛΓΟΡΙΘΜΟΣ ΑΝΙΧΝΕΥΣΗΣ ΠΤΩΣΗΣ

Η ανίχνευση της πτώσης με την εφαρμογή ενός κατωφλίου για τις τιμές κορυφής (peak value) , της προκύπτουσας γωνιακής ταχύτητας W_{res} , που καταγράφονται από τα δεδομένα πτώσης ή ADL , θα έχει ως αποτέλεσμα ένα από τα δύο παρακάτω σενάρια :

- I. Οι μέγιστες τιμές W_{res} που καταγράφονται από τις ADL **δεν** θα επικαλύπτονται από τις καταγεγραμμένες ελάχιστες peak value τιμές πτώσης W_{res} (Εικόνα 10). Έτσι ως κατώφλι μπορεί να χρησιμοποιηθεί η χαμηλότερη μέγιστη τιμή πτώσης για να διακριθούν οι πτώσεις από τις ADL(καθημερινές δραστηριότητες).
- II. Οι μέγιστες τιμές W_{res} που καταγράφονται από τις ADL **θα** επικαλύπτονται από τις καταγεγραμμένες ελάχιστες peak value τιμές πτώσης W_{res} (Εικόνα 10). Σε αυτήν την περίπτωση η εφαρμογή ενός κατωφλίου για να διακριθούν οι πτώσεις από τις ADL είναι ανεπαρκής. Χρειάζεται περισσότερη διερεύνηση σε πρόσθετες πτυχές του σήματος.



ΕΙΚΟΝΑ 10 – Επικαλυπτόμενα και μη επικαλυπτόμενα δείγματα δεδομένων πτώσεις και ADL. Πηγή : (Bourke και Lyons, 2006) at [17]

Θέτοντας ως τιμή κατωφλίου την χαμηλότερη προκύπτουσα γωνιακή ταχύτητα W_{res} , διασφαλίστηκε ότι θα ανιχνευτούν όλες οι πτώσεις σωστά αλλά ορισμένες ADL θα αναγνωριστούν λανθασμένα ως πτώσεις. Αυτήν η τιμή κατωφλίου ονομάστηκε Fall Threshold 1 (FT1).

Θα εξεταστεί η προκύπτουσα γωνιακή επιτάχυνση A_{res} , και η αλλαγή του σήματος της γωνίας του στέρνου Θ_{res} , για να διακριθούν οι υπόλοιπες ADL από τις πτώσεις.

Ορίστηκε ως τιμή κατωφλίου Fall Threshold 2 (FT2), η τιμή A_{res} , που είναι ακριβώς κάτω από την χαμηλότερη καταγεγραμμένη μέγιστη τιμή πτώσης A_{res} , που

καταγράφηκε σε μία χρονική περίοδο 1 sec κεντραρισμένο στο t_0 . Όπου ως t_0 , αναφέρεται στον χρόνο $t = 0\text{ms}$, όπου έγινε ακριβώς η υπέρβαση της τιμής FT1, από την προκύπτουσα γωνιακή ταχύτητα W_{res} . Αυτό το 1 sec χρονικό παράθυρο επιλέχθηκε για να συλλάβει την απότομη αλλαγή της περιστροφής του κορμού του ατόμου, που συνδέεται με την πρόσκρουση κατά της πτώση. Δηλαδή t_0 είναι η χρονική στιγμή της πιθανής πρόσκρουσης από την πιθανή πτώση ή t_0 είναι το χρονικό σημείο όπου $W_{res} = FT1$.

Ισχύουν οι σχέσεις :

$$A_{res} = \sqrt{A_p^2 + A_r^2}, \text{ όπου,}$$

$$A_p = \frac{d\{Wp\}}{dt} \text{ από } -0.5\text{s} \text{ έως } 0.5\text{s}$$

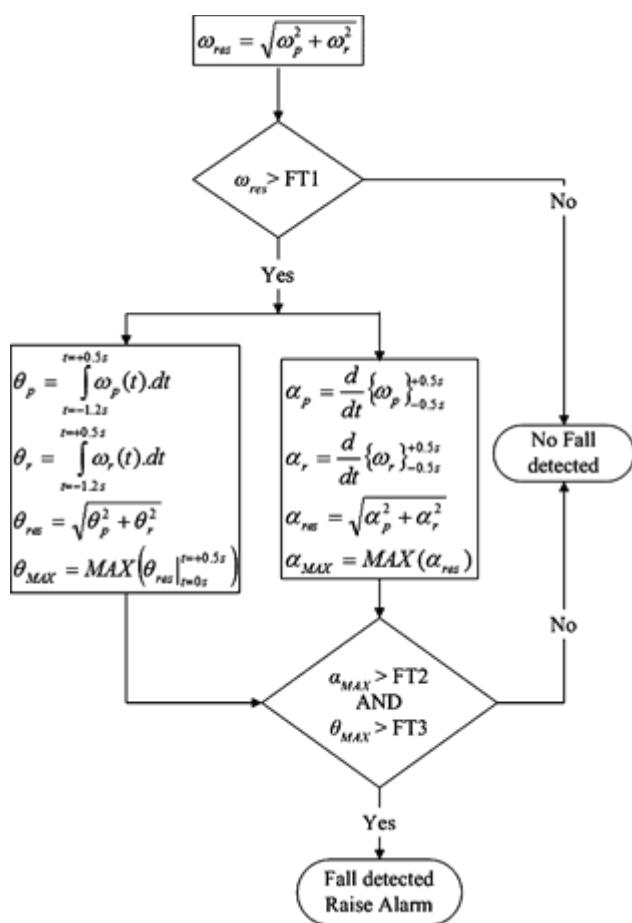
$$A_r = \frac{d\{Wr\}}{dt} \text{ από } -0.5\text{s} \text{ έως } 0.5\text{s}$$

Για την αλλαγή της γωνίας του σήματος Θ_{res} κατά την πτώση, μας ενδιαφέρει η αλλαγή λίγο πριν και λίγο μετά την πτώση. Μετρήθηκε η τιμή της Θ_{res} 1.2 sec πριν την πτώση και 0.5 sec μετά, δηλαδή 1.2 sec πριν το t_0 και 0.5 sec μετά. Η τιμή 1.2 sec επιλέχθηκε λόγω μιας μελέτης των Hsiao and Robinovitch [18], που παρατήρησαν ότι η πρόσκρουση από την πτώση γίνεται περίπου $715\text{ ms} \pm 160$, μετά την αρχική διαταραχή. $1.2\text{ sec} = 715\text{ ms} \pm 3 * 160$.

Για την αλλαγή της γωνίας του σήματος Θ_{res} , που είναι αποτέλεσμα της πρόσκρουσης, μας ενδιαφέρει μόνο το χρονικό διάστημα από t_0 έως 0.5sec.

Ορίστηκε ως τιμή κατωφλίου Fall Threshold 3 (FT3), η χαμηλότερη καταγεγραμμένη μέγιστη τιμή (peak value) πτώσης Θ_{res} .

Το διάγραμμα ροής του αλγορίθμου φαίνεται παρακάτω (όπου MAX αναφέρεται στην μέγιστη τιμή από τις μέγιστες τιμές του επί εξέταση σήματος) :



Πηγή : (Bourke και Lyons, 2006) at [17]

ΑΠΟΤΕΛΕΣΜΑΤΑ

Μέσα από την ανάλυση των 240 πειραμάτων πτώσης που έγιναν , οι 3 τιμές κατωφλίου για τα FT1 , FT2 και FT3 που θα μπορούσαν να χρησιμοποιηθούν ώστε να έχουμε sensitivity 100% φαίνονται στον παρακάτω πίνακα :

	FT1	FT2	FT3
Σύμβολο	Wres	Ares	Θres
Σήμα	Γωνιακή ταχύτητα	Γωνιακή	Αλλαγή της γωνίας

		επιτάχυνση	του θώρακα
Τιμή κατωφλίου	3.1 rad/sec	0.05 rad/sec ²	0.59 rad

Από την στιγμή που τα FT1 ,FT2 και FT3 είχαν οριστεί ως οι χαμηλότερες τιμές από τις καταγεγραμμένες μέγιστες τιμές(peak values) για κάθε σήμα αντίστοιχα , τότε το sensitivity είναι εγγυημένα 100% , τουλάχιστον για τις συγκεκριμένες 240 πτώσεις (Εικόνα 11).

Το όριο FT1 αν το δούμε ατομικά αναγνωρίζει το 97.5% των μή πτώσεων (specificity = 97.5%) . Οι ADL που υπερβαίνουν αυτά τα όρια ήταν οι εξής :

- Κάθεται και σηκώνεται από μία πολυθρόνα.
- Μπαίνοντας και βγαίνοντας από το αυτοκίνητο.
- Κάθεται και σηκώνεται από ένα χαμηλό σκαμνί.
- Κάθεται και σηκώνεται από το κρεβάτι.

Με τον συνδυασμό FT1 και FT2 το specificity ήταν 99.2% . Αναγνώρισε λανθασμένα σαν πτώσεις μόνο τις εξής ADL:

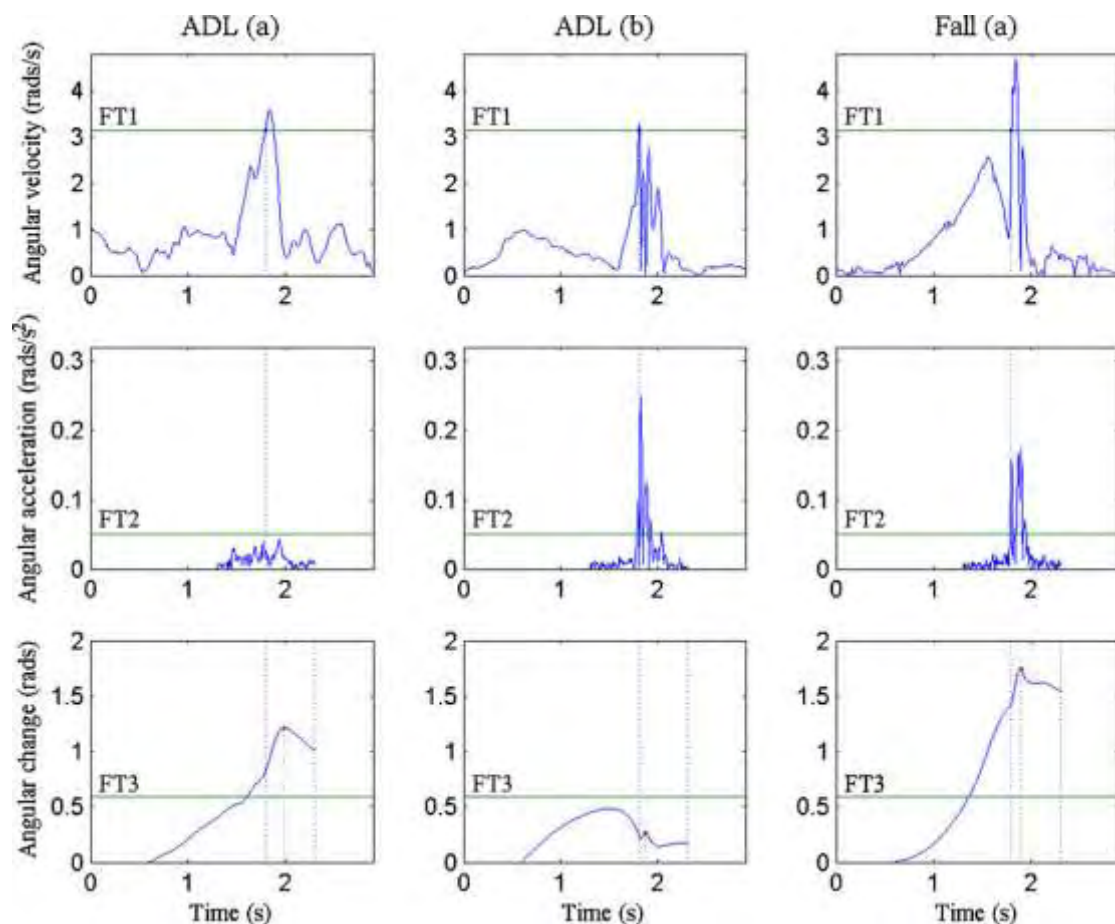
- Κάθεται και σηκώνεται από ένα χαμηλό σκαμνί.
- Κάθεται και σηκώνεται από μία καρέκλα.

Με τον συνδυασμό FT1 ,FT2 και FT3 το specificity και το sensitivity είχαν τιμή 100%.

Εδώ όμως θα πρέπει να σημειώσουμε, ότι μπορεί να έχουμε 100% sensitivity και specificity αλλά μόνο για τις 240 + 240 προσομοιώσεις πτώσεων και καθημερινών

δραστηριοτήτων αντίστοιχα που έγιναν. Σε πραγματικές συνθήκες αυτά τα νούμερα δεν ισχύουν για λόγους όπως :

- Η προσομοίωση πτώσης νεαρών ατόμων δεν προσομοιώνει 100% τις πραγματικές πτώσεις ηλικιωμένων ατόμων.
- Έγιναν 240 πτώσεις και άλλα τόσα πειράματα καθημερινών δραστηριοτήτων , αλλά στην πράξη υπάρχουν πολύ περισσότερα είδη πτώσεων και δραστηριοτήτων.



EIKONA 11 –ADL (a)→ ADL όπου δεν ξεπεράστηκε το FT2 ,ADL (b)→ ADL όπου δεν ξεπεράστηκε το FT3, Fall (a) όπου έχουμε πτώση. Πηγή : (Bourke και Lyons, 2006) at [17]

2.5 ΣΥΜΠΕΡΑΣΜΑΤΑ

Δυστυχώς όπως είδαμε , δεν υπάρχει μία κοινή διαδικασία αξιολόγησης των αλγορίθμων ανίχνευσης πτώσης. Η κάθε επιστημονική ομάδα χρησιμοποίησε τα δικά της δεδομένα (επιταχυνσιομέτρων , γυροσκοπίων κτλ) που προήλθαν από δικά τους πειράματα. Για τον λόγο αυτό είναι αδύνατο να βγάλουμε ασφαλή συμπεράσματα για τον καλύτερο ή χειρότερο αλγόριθμο. Επίσης κάποιος αλγόριθμος μπορεί να υπερέχει στην ανίχνευση πτώσεων αλλά να υστερεί στην ανίχνευση των μή-πτώσεων ή το αντίθετο , σε σύγκριση με κάποιον άλλον αλγόριθμο.

Το μόνο σίγουρο είναι ότι ακόμα κανένας αλγόριθμος δεν κατάφερε να έχει επιτυχία 100% (sensitivity και specificity 100%). Τέτοια ποσοστά έχουν επιτευχθεί μόνο σε πειραματικά περιβάλλοντα και όχι στον περιβάλλον του σπιτιού του χρήστη.

Ορισμένα εμπορικά προϊόντα έχουν γίνει διαθέσιμα στην αγορά , παρόλα αυτά οι αισθητήρες πτώσεων , έχουν πολύ μικρή εφαρμογή στην καθημερινότητα. Οι λόγοι που δεν χρησιμοποιούνται είναι αρκετοί όπως :

- Ανεπαρκής λειτουργία.
- Κακή εργονομία (γίνεται ενοχλητικό στους χρήστες).
- Πολλοί ψευδοί συναγερμοί.
- Η υπερηφάνεια των ηλικιωμένων που δεν δέχονται να χρησιμοποιήσουν κάποια συσκευή ανίχνευσης πτώσης.

3

Motion Capture

Ο όρος motion capture (καταγραφή της κίνησης) αναφέρεται στην διαδικασία κατά την οποία , μέσω της χρήσης κατάλληλων συσκευών, γίνεται δυνατή η καταγραφή της κίνησης ενός ανθρώπου ή κάποιου άλλου αντικειμένου με σκοπό την δημιουργία μιας τρισδιάστατης ψηφιακής αναπαράστασης της. Αυτή η αναπαράσταση προκύπτει από την συνεχή παρακολούθηση κάποιων σημείων του σώματος του ανθρώπου (σημεία αρθρώσεων) ή του αντικειμένου για κάποιο χρονικό διάστημα. Τα δεδομένα της αναπαράστασης που συλλέγονται αποθηκεύονται σε συγκεκριμένους τύπους αρχείων, όπως θα δούμε στις επόμενες ενότητες, για να χρησιμοποιηθούν από κάποιο πρόγραμμα του υπολογιστή.

Σε αυτήν την ενότητα θα παρουσιάσουμε τα συστήματα καταγραφής κίνησης και θα δούμε σε ποιές εφαρμογές χρησιμοποιούνται. Θα μελετήσουμε αναλυτικότερα την βάση δεδομένων καταγραφής κίνησης MHAD και τους αλγόριθμους που χρησιμοποιήθηκαν και έναν αλγόριθμο ανακατασκευής της κίνησης με χρήση επιταχυνσιομέτρου.

3.1 ΣΥΣΤΗΜΑΤΑ ΚΑΤΑΓΡΑΦΗΣ ΤΗΣ ΚΙΝΗΣΗΣ

Τα συστήματα καταγραφής κίνησης επιτρέπουν τον υπολογισμό της θέσης κάποιων σημείων του ανθρώπινου σώματος σε κάποιες χρονικές στιγμές. Οι τρεις πιο σημαντικές μέθοδοι σύλληψης δεδομένων της ανθρώπινης κίνησης είναι η μηχανική , η μαγνητική και η οπτική μέθοδος [19] :

➤ Μηχανική μέθοδος :

Αυτή η μέθοδος είναι η πρώτη που εφαρμόστηκε σε συστήματα λήψης δεδομένων κίνησης. Πάνω στον άνθρωπο εφαρμόζεται ένα σύνολο από μηχανικά κομμάτια. Καθώς αυτός κινείται, τα μηχανικά κομμάτια κινούνται και αυτά με αποτέλεσμα οι αισθητήρες που τοποθετούνται πάνω στις αρθρώσεις αυτών να στέλνουν το σήμα σε έναν υπολογιστή.

Η μέθοδος αυτή είναι πολύ ακριβή και αρκετά δύσκολη στην υλοποίηση καθώς πρέπει να δημιουργηθούν με μεγάλη ακρίβεια αυτά τα μηχανικά μέρη. Επίσης εμποδίζουν την ανθρώπινη κίνηση , με αποτέλεσμα να επηρεάζονται οι φυσικές ανθρώπινες κινήσεις.



Πηγή εικόνας <http://www.metamotion.com>

➤ Μαγνητική μέθοδος :

Η μέθοδος αυτή χρησιμοποιεί συσκευές αποστολής σήματος που δημιουργούν ένα μαγνητικό πεδίο σε κάποια περιοχή στην οποία βρίσκεται ο άνθρωπος στον οποίο είναι τοποθετημένοι αισθητήρες ,οι οποίοι μπορούν να μετρήσουν την απόσταση τους από τις συσκευές αποστολής σήματος καθώς και την κατεύθυνση τους. Η μέθοδος αυτή είναι αρκετά φθηνή και αποδοτική και μπορεί να λειτουργήσει σε real-time. Έχει όμως και πολλά αρνητικά , όπως θόρυβο από άλλα μαγνητικά πεδία στην περιοχή. Επίσης λόγω των καλωδίων εμποδίζεται η φυσική κίνηση των ανθρώπων όπως και στην προηγούμενη μέθοδο.

➤ Οπτική μέθοδος:

Αυτή η μέθοδος είναι η πλέον εύκολη και αποδοτική. Τοποθετούνται πάνω στον άνθρωπο ειδικοί δείκτες (markers) που είναι αντανακλαστικοί ή που εκπέμπουν υπέρυθρο σήμα , και με την χρήση καμερών γύρω γύρω καθορίζεται η ακριβής θέση των σημείων-δεικτών σε ένα χρονικό διάστημα εκτέλεσης μιας φυσικής ανθρώπινης κίνησης.

3.2 ΠΕΔΙΑ ΕΦΑΡΜΟΓΗΣ

Η καταγραφή της κίνησης χρησιμοποιείται για ιατρικούς λόγους , animation αλλά και σε παιχνίδια με real-time χειρισμό, που προέρχεται από την ανθρώπινη κίνηση και χωρίς την χρήση κουμπιών.

Χρησιμοποιείται σε ορθοπαιδικές εφαρμογές , στην φυσικοθεραπεία , αλλά και στην πρόληψη των τραυματισμών. Η ανάλυση βάρδισης (gait analysis) [20] έχει αποδειχθεί

πολύ χρήσιμη στην ανίχνευση ανωμαλιών κατά το περπάτημα που οφείλονται σε ορθοπεδικές ανωμαλίες από τραυματισμούς ή μη, εγκεφαλικά επεισόδια κτλ. Η ανάλυση βάδισης μπορεί να συμβάλει στον εντοπισμό των αιτιών του προβλήματος αλλά και στην συνεχή παρακολούθηση και αξιολόγηση της κατάστασης των ασθενών.

Η καταγραφή της κίνησης παίζει πολύ σημαντικό ρόλο στην δημιουργία animation στα video games ή δημιουργίας οπτικών εφέ σε ταινίες. Οι περίπλοκες κινήσεις ή εξειδικευμένες κινήσεις (πχ γνωστών ποδοσφαιριστών σε παιχνίδια ποδοσφαίρου) , εκτελούνται από ανθρώπους και καταγράφονται από συσκευές καταγραφής κίνησης. Αυτά τα δεδομένα χρησιμοποιούνται από τους animators και μπορούν να δώσουν κίνηση σε κάποια μοντέλα που φτιάχνονται σε υπολογιστή.

Μεγάλη εφαρμογή, έχει η καταγραφή της κίνησης, τα τελευταία χρόνια στα διαδραστικά παιχνίδια , όπου ο χειρισμός γίνεται χωρίς την χρήση πληκτρολογίου ή κάποιου gamepad , αλλά με την κίνηση του ανθρώπινου σώματος :

- Wii Remote της Nintendo.
- Το Move για το Playstation 3 και το Ps4 Eye από την sony.
- Το Kinect της Microsoft για το Xbox.

Τέλος χρησιμοποιείται στον αθλητισμό για την κατανόηση και βελτίωση της τεχνικής, στρατιωτικούς σκοπούς κτλ.

3.3 ΗΔΗ ΑΡΧΕΙΩΝ ΚΙΝΗΣΗΣ

Όλα τα δεδομένα που καταγράφονται αποθηκεύονται σε συγκεκριμένου τύπου αρχεία ώστε αργότερα να αποτελέσουν αντικείμενα ανάλυσης και επεξεργασίας σε προγράμματα όπως το MotionBuilder της Autodesk , Maya 3D animation , matlab μετά από επεξεργασία κτλ. Τα αρχεία αυτά περιλαμβάνουν τις συντεταγμένες σε τρισδιάστατο χώρο ή τις τιμές των γωνιών περιστροφής για τις αρθρώσεις , για ένα πλήθος διαδοχικών χρονικών πλαισίων(frames). Παρακάτω θα δούμε μερικά από αυτά τα ήδη αρχείων :

3.3.1 ΑΡΧΕΙΑ BIOVISION HIERARCHY (BHV)

Τα αρχεία BHV [21] δημιουργήθηκαν από την εταιρεία καταγραφής κίνησης Biovision . Τα αρχεία αυτά είναι τα δημοφιλέστερα που υπάρχουν στον τομέα του mocap , και υποστηρίζονται από πολλά πακέτα λογισμικού animation και γραφικών.

Ένα αρχείο BHV χωρίζεται σε 2 μέρη :

- Το πρώτο μέρος που ονομάζεται header :

Σε αυτό το μέρος του αρχείου υπάρχουν πληροφορίες για την ιεραρχία και την αρχική πόζα του σκελετού. Υπάρχει μία άρθρωση ROOT που αποτελεί την ρίζα της σκελετικής ιεραρχίας. Όλες οι άλλες αρθρώσεις αποτελούν παιδιά από υπάρχουσες αρθρώσεις γονέων. Για μία άρθρωση παιδί υπολογίζεται η απόσταση του από τον γονέα καθώς και η κατεύθυνση του στην αρχική πόζα.

- Το δεύτερο μέρος :

Σε αυτό το μέρος του αρχείου υπάρχουν πληροφορίες για την κίνηση του κάθε μέρους της ιεραρχίας σε κάθε frame. Ο υπολογισμός κάθε θέσης της άρθρωσης , γίνεται με την βοήθεια ενός πίνακα τοπικού μετασχηματισμού. Η θέση κάθε άρθρωσης προκύπτει από την συσσώρευση όλων των τοπικών μετασχηματισμών από την συγκεκριμένη άρθρωση έως την άρθρωση ρίζα.

3.3.2 ΑΡΧΕΙΑ ASF/AMC

Τα αρχεία αυτή της μορφής αναπτύχθηκαν από την εταιρεία βιντεοπαχνιδιών Acclaim [22] και πλέον χρησιμοποιούνται και από άλλες εταιρείες και από πολλά λογισμικά γραφικών.

Τα δεδομένα χωρίζονται σε 2 διαφορετικά αρχεία. Τα ASF (Acclaim Skeleton File) έχουν δεδομένα για την σκελετική ιεραρχία και τα AMC (Acclaim Motion Capture data) περιέχουν δεδομένα κίνησης. Αυτός ο διαχωρισμός έγινε για να μπορούν να εφαρμοστούν πιο εύκολα πολλές διαφορετικές κινήσεις σε έναν συγκεκριμένο σκελετό χωρίς να χρειάζεται να υπάρχει η πληροφορία του σκελετού σε πολλά διαφορετικά αρχεία , όπως στην περίπτωση των BHV αρχείων.

3.3.3 ΑΡΧΕΙΑ C3D

Τα αρχεία αυτά [23] είναι μια μορφή δυαδικών αρχείων και αναπτύχθηκαν από την εταιρεία Vicon [24] , που ασχολείται με την ανάπτυξη οπτικών συστημάτων σύλληψης δεδομένων κίνησης. Χρησιμοποιούνται για animation , ανάλυση βηματισμού κτλ.

Πριν την ανάπτυξη αυτής της μορφής αρχείων , τα διάφορα mocap συστήματα , αποθήκευαν τα δεδομένα σε πολλά διαφορετικά αρχεία και σε διαφορετικές μορφές. Η ιδέα της δημιουργίας των C3D αρχείων ήταν να χρησιμοποιείται ένα μόνο αρχείο για την αποθήκευση όλων των δεδομένων μαζί.

3.4 ΜΙΑ ΠΕΡΙΕΚΤΙΚΗ MULTIMODAL HUMAN

ACTION DATABASE (MHAD)

Με τα χρόνια ένας μεγάλος αριθμός μεθόδων έχουν προταθεί για την ανάλυση της ανθρώπινης στάσης και κίνησης από εικόνες , video και πρόσφατα από πιο λεπτομερείς δεδομένα. Σε αυτήν την ενότητα θα παρουσιάσουμε το σύστημα Multimodal Human Action Database (MHAD) που προτάθηκε στο πανεπιστήμιο του Berkeley [25].

Το σύστημα αυτό αποτελείται από συγχρονισμένα και γεωμετρικά βαθμονομημένα δεδομένα από ένα οπτικό σύστημα σύλληψης κίνησης , stereo κάμερες , αισθητήρες βάθους , επιταχυνσιόμετρα και μικρόφωνα. Θα δούμε ότι η multimodal (πολυτροπική) ανάλυση της ανθρώπινης κίνησης μας δίνει καλύτερα ποσοστά αναγνώρισης δράσης από την unimodal (μονοτροπική) ανάλυση. Στην επόμενη ενότητα θα παρουσιάσουμε κάποιες άλλες βάσεις δεδομένων που χρησιμοποιούνται για την καταγραφή της ανθρώπινης κίνησης.

3.4.1 ΥΠΑΡΧΟΥΣΕΣ ΒΑΣΕΙΣ ΔΕΔΟΜΕΝΩΝ ΑΝΑΓΝΩΡΙΣΗΣ ΑΝΘΡΩΠΙΝΗΣ

ΚΙΝΗΣΗΣ.

Ένα από τα πρώτα σύνολα δεδομένων ανθρώπινης δράσης που βασίζεται στο video , ήταν το KTH , από τον Schuldt [26], και το Weizmann από τον Blank [27].

Αργότερα νέα σύνολα δεδομένων εμφανίστηκαν με πλουσιότερες ομάδες δραστηριοτήτων και πιο σύνθετα περιβάλλοντα , τα οποία εξομοιώνουν πιο πιστά τα πραγματικά σενάρια. Αυτά περιλαμβάνουν συλλογές από video clips από διάφορες τηλεοπτικές εκπομπές , ταινίες και video από χρήστες (πχ. από το Youtube):

- Hollywood2 από τον Marszalek.
- UCF50 από τον Liu.
- Olympic Sports από τον Niebles.
- HMDB51 από τον Kuehne.

Κάποια multiview σύνολα δεδομένων video που αύξησαν την διαθέσιμη ποσότητα πληροφοριών για την ανάλυση και έβαλαν την βάση για την δημιουργία αλγορίθμων εκτίμησης της ανθρώπινης κίνησης και στάσης:

- CMU Motion of Body (MoBo) από τους Gross and Shi.
- IXMAS από τον Weinland.
- ViHASi: Virtual Human Action Silhouette Data από τον Ragheb.
- i3DPost Multi-view και 3D Human Action/Interaction Database από τον Gkalelis.

Σε αντίθεση με τα παραπάνω video-based σύνολα ανθρώπινης δράσης δημιουργήθηκαν και κάποια mocap-based σύνολα δεδομένων που δεν βασίζονται πολύ στα video όπως :

- CMU Motion Capture Dataset από το CMU Graphics Lab.
- HDM05 Mocap Database από τον Muller.

Και τέλος έχουμε τα mocap-video-based σύνολα δεδομένων που μπορεί να χρησιμοποιούν και δεδομένα από μικρόφωνα , αισθητήρες , επιταχυνσιόμετρα , δεδομένα βάθους (πχ. Microsoft Kinect) κτλ όπως:

- HumanEva I and II datasets από τον Sigal.
- TUM Kitchen Dataset από τον Tenorth.
- CMU Multimodal Activity (CMU-MMAC) Dataset από τον De La Torre.
- MSR-Action3D Dataset από τον Li.
- MSRDailyActivity3D Dataset από τον Wang.

Στον παρακάτω πίνακα φαίνονται οι διαφορετικοί τρόποι ανάλυσης που χρησιμοποιούνται από τα διάφορα σύνολα δεδομένων :

Database	Video	Mocap	Depth	Audio	Others
KTH	1				
Weizmann	1				
Hollywood2	1				
UCF50	1				
Olympic Sports	1				
HMDB51	1				
CMU MoBo	6				

IXMAS	5				
ViHASi	<40				
i3DPost	8				
CMU Mocap	1	*			
HDM05		*			
HumanEva I	7	*			
HumanEva II	4	*			
TUM Kitchen	4	*			*
CMU MMAC	6	*		5	*
MSR-Action3D				1	
MSRDailyActivity3	1			1	
MHAD	12	*	2	4	*

Παρατηρούμε ότι μόνο η βάση δεδομένων MHAD χρησιμοποιεί τόσους πολλούς τρόπους ανάλυσης της κίνησης οι οποίοι αποτελούνται από multiview video , δεδομένα βάθους και χρώματος από πολλές κάμερες Kinect της Microsoft , η δυναμική της κίνησης με χρήση επιταχυνσιόμετρων και ακριβή mocap δεδομένα με τις πληροφορίες σκελετού. Επιπλέον καταγράφεται ο ήχος του περιβάλλοντος κατά την εκτέλεση της δράσης με σκοπό να αποκαλύψει διακριτά “συνθήματα” για την ανάλυση της ανθρώπινης κίνησης.

3.4.2 MULTIMODAL HUMAN ACTION DATABASE

Σε αυτήν την ενότητα θα παρουσιάσουμε τα επιμέρους στοιχεία που χρησιμοποιούνται από το σύνολο δεδομένων MHAD. Κάθε δράση συλλαμβάνεται από 5 διαφορετικά συστήματα :

- 1) Το οπτικό σύστημα mocap.
- 2) 4 multiview κάμερες στερεοσκοπικής όρασης.
- 3) 2 κάμερες Microsoft Kinect.
- 4) 6 ασύρματα επιταχυνσιόμετρα.
- 5) 4 μικρόφωνα.

Η βάση δεδομένων MHAD , περιέχει 11 δράσεις που υλοποιούνται από 12 άτομα. Όλα τα άτομα εκτέλεσαν από 5 επαναλήψεις κάθε δράσης αποδίδοντας 660 σκηνές δράσης. Σε όλα τα πειράματα οι δράσεις των 7 πρώτων ατόμων χρησιμοποιήθηκαν για εκπαίδευση και των άλλων 5 για δοκιμή. Οι δράσεις αυτές περιλαμβάνουν :

- Δράσεις με κίνηση στα 2 άνω άκρα και 2 κάτω άκρα , πχ. , ρίψεις , επιτόπιο άλμα κτλ.
- Δράσεις με υψηλή δυναμική στα άνω άκρα , πχ. , κουνώντας τα χέρια , παλαμάκια κτλ.
- Δράσεις με υψηλή δυναμική στα κάτω άκρα, πχ. , να καθίσει , να σηκώνεται κτλ.

Στην εικόνα 12 φαίνονται αυτές οι δράσεις , μαζί με το αντίστοιχο “σύννεφο” που εξάγεται από τα δεδομένα του Kinect.



ΕΙΚΟΝΑ 12 : Στιγμιότυπα όλων των δράσεων της βάσης δεδομένων MHAD , με τα αντίστοιχα σημεία από το σύννεφο που δημιουργείται από τα δεδομένα του Kinect. Πηγή : (Ofli *et al.* ,2013) at [25]

Mocap System :

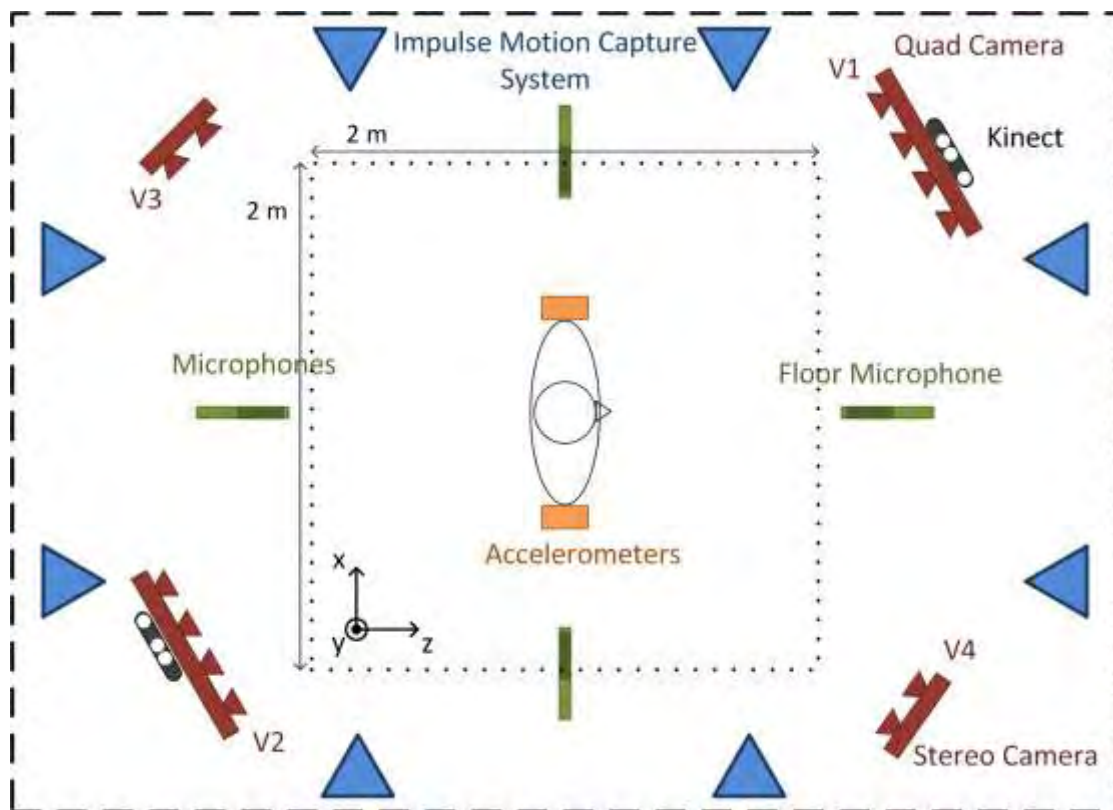
Χρησιμοποιήθηκε ένα οπτικό σύστημα καταγραφής κίνησης που συλλάμβανε την 3D θέση του ατόμου με 43 ενεργούς δείκτες LED , που προσαρμόστηκαν πάνω σε ένα εφαρμοστό κοστούμι , με συχνότητα 480 Hz.

Ο server που έπαιρνε τα δεδομένα , τα συγχρόνιζε χρονικά με τους άλλους αισθητήρες (κάμερες , Kinect , επιταχυνσιόμετρα κτλ.) Επίσης με την βοήθεια του λογισμικού MotionBuilder της AutoDesk εξάγαμε τον σκελετό κάθε αντικειμένου.

Camera System :

Χρησιμοποιήθηκαν 12 κάμερες Dragonfly2 ανάλυσης 640×480. Οι 4 κάμερες για στερεοσκοπική λήψη και οι άλλες 8 για multi-view όπως φέρεται στην εικόνα 13.

Πριν ληφθούν τα δεδομένα, οι κάμερες είχαν γεωμετρικά βαθμονομηθεί και ευθυγραμμιστεί με το σύστημα mocap.



ΕΙΚΟΝΑ 13. Διάγραμμα του συστήματος απόκτησης δεδομένων. Πηγή : (Ofli *et al.*, 2013) at [25]

Kinect System :

Για την απόκτηση δεδομένων βάθους χρησιμοποιήθηκαν 2 κάμερες Kinect με αντίθετες κατευθύνσεις όπως φαίνεται στην εικόνα 12. Κάθε κάμερα Kinect συλλάμβανε μία έγχρωμη εικόνα με ανάλυση 640×480 pixels και βάθος εικόνας 16-bit στα 30 Hz. Οι κάμερες ήταν χρονικά συγχρονισμένες με τον server και είχαν γεωμετρικά βαθμονομηθεί με το σύστημα mocap.

Επιταχυνόμετρα :

Χρησιμοποιήθηκαν 6 τριαξονικά ασύρματα επιταχυνσιόμετρα τοποθετημένα στο mocap κοστούμι για την μέτρηση της κίνησης στους καρπούς, τους αστραγάλους και τους γοφούς.

Audio system :

Χρησιμοποιήθηκαν 4 μικρόφωνα για την καταγραφή του ήχου κατά την διάρκεια των δράσεων. 3 από αυτά σε απόσταση 65cm από το έδαφος και 1 στο πάτωμα για να συλλάβει τους ήχους που δημιουργούνται στην επιφάνεια του εδάφους.

3.4.3 ΑΛΓΟΡΙΘΜΟΙ ΠΟΥ ΧΡΗΣΙΜΟΠΟΙΟΥΝΤΑΙ ΓΙΑ ΤΗΝ ΑΝΑΓΝΩΡΙΣΗ ΤΗΣ ΔΡΑΣΗΣ

Μία ακολουθία κίνησης αναπαριστάται από έναν πίνακα X διάστασης $N \times D$, όπου κάθε γραμμή αντιστοιχεί σε ένα πλαίσιο καταγραφής και αποτελεί ένα διάνυσμα χαρακτηριστικών μεγέθους D . Τα χαρακτηριστικά αυτά είναι τα δεδομένα περιστροφής κάθε άρθρωσης μιας σκελετικής ιεραρχίας. Οι ακολουθίες κίνησης συνενώνονται σε έναν ενιαίο πίνακα για να εφαρμοστεί η μέθοδος μείωσης διάστασης Locality Preserving Projections (LPP) [28]. Αυτό γίνεται επειδή σε αρκετές περιπτώσεις είναι αναγκαίο να μειωθεί η διάσταση των δεδομένων προκειμένου τα δείγματα να αναπαρασταθούν από διανύσματα χαρακτηριστικών μικρότερου μήκους. Η τεχνική μείωσης διάστασης LPP αποτελεί μία γραμμική μέθοδο μείωσης διάστασης.

Αν για παράδειγμα υπάρχουν k ακολουθίες X_1, X_2, \dots, X_k που αποτελούνται από N_1, N_2, \dots, N_k πλαίσια αντίστοιχα, τότε οι ακολουθίες συνενώνονται σε ένα πίνακα X , με $\sum_i N_i$ γραμμές και D στήλες. Μέσω την τεχνική LPP υπολογίζεται ο πίνακας μετασχηματισμού A , όπως στο [28], ο οποίος μπορεί να χρησιμοποιηθεί για τον μετασχηματισμό των ακολουθιών εκπαίδευσης και δοκιμής.

Στην συνέχεια, τα μετασχηματισμένα δεδομένα τροφοδοτούνται στον αλγόριθμο αναγνώρισης κίνησης ο οποίος περιγράφεται παρακάτω.

Εκτελέστηκαν πειράματα αναγνώρισης της δράσης βασιζόμενα στην μέθοδο Bag-of-Words για να μοντελοποιηθεί η ακολουθία της δράσης για έναν συγκεκριμένο τρόπο σύλληψης της κίνησης. Για κάθε τρόπο σύλληψης της δράσης εξήχθησαν χαρακτηριστικά από ολόκληρη την ακολουθία και κβαντοποιήθηκαν εφαρμόζοντας ομαδοποίηση k -μέσων. Υπολογίστηκαν οι ομοιότητες 2 ακολουθιών υπολογίζοντας την x^2 απόσταση των ιστογραμμάτων που αντιπροσωπεύουν τις 2 ακολουθίες και χρησιμοποιήθηκε k -Nearest Neighbors (k -NN) κατηγοριοποιητής για να χαρακτηριστεί η δράση της κάθε ακολουθίας όπως και kernel-SVM με x^2 kernel για κατηγοριοποίηση. Στην συνέχεια χρησιμοποιήθηκε ο Multiple Kernel Learning (MKL) για τον υπολογισμό του βέλτιστου γραμμικού συνδυασμού των πολυτροπικών/multiview ομοιοτήτων.

3.4.3.1 ΜΕΘΟΔΟΣ ΒΑΣΙΣΜΕΝΟΣ ΣΕ BAG OF WORDS (BOW)

Η μέθοδος Bag-Of-Words (BOW) [29], χρησιμοποιήθηκε για την αναπαράσταση των δεδομένων κίνησης σε συνδυασμό με την μείωση διάστασης. Συγκεκριμένα,

κάθε ακολουθία αναπαριστάται από ένα ιστόγραμμα , που υπολογίζεται από τις συχνότητες εμφάνισης χαρακτηριστικών πλαισίων , τα οποία ονομάζονται dynemes.

Αρχικά εξάγεται ένα σύνολο απο dynemes πλήθους C από ένα σύνολο ακολουθιών εκπαίδευσης. Αυτό γίνεται με την εφαρμογή ομαδοποίησης k -μέσων στο σύνολο των πλαισίων όλων των ακολουθιών εκπαίδευσης. Κάθε dyneme αντιπροσωπεύει το κέντρο μιας από τις ομάδες (clusters) που προέκυψαν από την ομαδοποίηση. Αφού κάθε dyneme είναι ο μέσος όρος των πλαισίων (δειγμάτων) που ανήκουν στην ίδια ομάδα , είναι λογικό ότι τα dynemes αυτά δεν αντιστοιχούν σε κάποιο από τα αρχικά πλαίσια.

Μετά οι ακολουθίες εκπαίδευσης αντιστοιχίζονται στα dynemes με τον εξής τρόπο:

Με εφαρμογή k -μέσων , κάθε πλαίσιο μιας ακολουθίας εκπαίδευσης έχει ανατεθεί σε μία ομάδα $K= 1,2,\dots, C$, επομένως κάθε πλαίσιο μπορεί να αντιστοιχηθεί σε ένα dyneme. Άρα μία ακολουθία $X= \{X_1, X_2, \dots, X_N\}$, μπορεί να μετατραπεί σε μία ακολουθία $X_D=\{d_1, d_2, \dots, d_N\}$ που αποτελείται από N πλαίσια , όπου κάθε πλαίσιο d_i προκύπτει από την αντικατάσταση του πλαισίου x_i της X από το πλησιέστερο σε αυτό dyneme. Αν υπολογίσουμε την συχνότητα εμφάνισης καθενός από τα C dynemes σε μία ακολουθία εκπαίδευσης , τότε προκύπτει ένα ιστόγραμμα , το οποίο είναι ένα διάνυσμα $S= \{s_1, s_2, \dots, s_C\}$ με , $s_i = \frac{n_i}{N}$, όπου n_i είναι ο αριθμός εμφανίσεων του dyneme i ($i=1,2,\dots,C$) στην ακολουθία ενώ N είναι το μήκος της ακολουθίας.

Για να γίνει κατηγοριοποίηση μιας ακολουθίας δοκιμής σε μία κλάση , εξάγεται από αυτήν ένα ιστόγραμμα. Όπως και με τις ακολουθίες εκπαίδευσης , κάθε πλαίσιο της αντιστοιχίζεται σε κάποιο dyneme. Για να γίνει αυτό, για κάθε πλαίσιο της ακολουθίας δοκιμής υπολογίζονται οι αποστάσεις από όλα τα dynemes. Το κάθε πλαίσιο θα

αντιστοιχηθεί με το *dyneme* που βρίσκεται πιο κοντά σε αυτό. Μετά βρίσκουμε το ιστόγραμμα για την ακολουθία δοκιμής ακριβώς όπως και για τις ακολουθίες εκπαίδευσης.

Το τελικό βήμα της μεθόδου BOW περιλαμβάνει την κατηγοριοποίηση αυτών των ιστογραμμάτων δοκιμής με αλγορίθμους κατηγοριοποίησης εξόρυξης δεδομένων (data mining). Στην περίπτωση που μελετάμε χρησιμοποιήθηκαν οι αλγόριθμοι *k*-nearest neighbour(*k*-NN) για *k*=1 και *k*=3 και support vector machines(SVM). Παρακάτω θα πούμε λίγα λόγια για τους αλγόριθμους *k*-NN και SVM χωρίς να μπούμε σε λεπτομέρειες.

3.4.3.2 ΤΟΜΗ ΙΣΤΟΓΡΑΜΜΑΤΩΝ

Για την κατηγοριοποίηση μιας ακολουθίας δοκιμής ,το ιστόγραμμα S_{test} , συγκρίνεται με καθένα από τα ιστογράμματα S_{train} ενός συνόλου ακολουθιών εκπαίδευσης και υπολογίζεται η ομοιότητα τους. Ως μέτρο ομοιότητας χρησιμοποιούμε την τομή των ιστογραμμάτων((histogramm intersection) που δίνεται ως εξής:

Η $I(S_{train}, S_{test}) = \sum_{i=1}^C \min \{S_{train,i}, S_{test,i}\}$,όπου $S_{train,i}$ και $S_{test,i}$ είναι οι συχνότητες εμφάνισης για το bin *i* του ιστογράμματος εκπαίδευσης και δοκιμής αντίστοιχα.

Με αυτό τον τρόπο το ιστόγραμμα της ακολουθίας δοκιμής συγκρίνεται με τα ιστογράμματα όλων των ακολουθιών εκπαίδευσης και κατηγοριοποιείται στην κλάση της ακολουθίας εκπαίδευσης που εμφάνισε την μεγαλύτερη τιμή τομής ιστογραμμάτων(1-Nearest Neighbour).

3.4.3.3 SUPPORT VECTOR MACHINES – SVMs

Ένας άλλος τρόπος για να γίνει η κατηγοριοποίηση είναι με την χρήση μηχανών εδραίων διανυσμάτων(SVMs). Τα SVMs [30] πραγματοποιούν κατηγοριοποίηση για προβλήματα 2 ή περισσότερων κλάσεων, μέσω του καθορισμού ενός επιπέδου που επιτυγχάνει μέγιστο διαχωρισμό μεταξύ των δειγμάτων εκπαίδευσης που ανήκουν σε διαφορετικές κλάσεις. Στην περίπτωση που μελετάμε χρησιμοποιήθηκε kernel-SVM με πυρήνα x^2 για κατηγοριοποίηση.

3.4.3.4 MULTIPLE KERNEL LEARNING (MKL)

Ο MKL μαθαίνει τον βέλτιστο γραμμικό συνδυασμό των πυρήνων , με σκοπό την επίτευξη καλύτερων αποτελεσμάτων σε σχέση με την μάθηση ενός απλού πυρήνα. Με λίγα λόγια και χωρίς να αναλύσουμε περαιτέρω τον MKL , με την χρήση του μπορούμε να βρούμε τον βέλτιστο γραμμικό συνδυασμό πολυτροπικών μεθόδων. Στην περίπτωση που μελετάμε χρησιμοποιήθηκε η εκτέλεση του MKL από τον Vedaldi στο [31] .

3.4.4 ΠΕΙΡΑΜΑΤΑ ΑΝΑΓΝΩΡΙΣΗΣ ΤΗΣ ΔΡΑΣΗΣ

Σε αυτήν την ενότητα θα παρουσιάσουμε τα πειράματα που έγιναν και τα αποτελέσματά τους για κάθε διαφορετικό από τους 5 τρόπους που χρησιμοποιεί η βάση δεδομένων καταγραφής κίνησης MHAD.

Σε όλα τα πειράματα χρησιμοποιήθηκαν τα πρώτα 7 αντικείμενα (άτομα) για την εκπαίδευση και τα επόμενα 5 αντικείμενα για δοκιμή.

1) Multi-view Video Data :

Οι μέθοδοι αναγνώρισης της δράσης υπολογίζουν χωροχρονικά σημεία ενδιαφέροντος από τα βίντεο και εξάγουν χαρακτηριστικά από ένα χωροχρονικό παράθυρο γύρω από κάθε σημείο ενδιαφέροντος.

Ο Wang [32],πραγματοποίησε μία συγκριτική μελέτη μεταξύ αυτών των χωροχρονικών χαρακτηριστικών σε διάφορες βάσεις δεδομένων και έδειξε ότι οι μέθοδοι, Histogram of Gradients (HOG) [33] και Histogram of Flow (HOF) [34] , αποδίδουν καλύτερα. Εξήχθησαν τα HOG ,HOF χαρακτηριστικά όπως παρουσιάζει ο Laptev στο [35].

Κβαντοποιήθηκαν τα χαρακτηριστικά εκπαίδευσης HOG ,HOF σε 20 codewords και εκτελέστηκαν τα πειράματα κατηγοριοποίησης. Στον παρακάτω πίνακα φαίνονται τα αποτελέσματα για κάθε μέθοδο κατηγοριοποίησης που χρησιμοποιήθηκε :

	Video-1	Video-2	Video-3	Video-4	MKL
1-NN	91.61	91.97	93.43	84.67	
3-NN	90.88	90.51	92.43	84.31	
K-SVM	91.97	87.96	96.35	87.23	99.27

Αποτελέσματα κατηγοριοποίησης δράσης (σε ποσοστό %) από δεδομένα video από 4 διαφορετικές οπτικές γωνίες κάμερών. Πηγή : (Ofli *et al.* ,2013) at [25]

Συνολικά βλέπουμε ότι τα ποσοστά επιτυχούς κατηγοριοποίησης ήταν αρκετά υψηλά. Αυτό γίνεται επειδή τα δεδομένα video είχαν στατικό φόντο άρα και ελάχιστο θόρυβο. Επίσης ο συνδυασμός όλων των καμερών με τον αλγόριθμο MKL , μας δίνει το εξαιρετικό ποσοστό 99.27%.

2) Kinect Depth Data

Για την εξαγωγή των χαρακτηριστικών από βίντεο βάθους , για την αναγνώριση της δράσης , ακολουθήθηκε η προσέγγιση του Ni [36] , όπου επεκτείνει τα χωροχρονικά χαρακτηριστικά του Laptev στον τομέα του βάθους.

Αρχικά χωρίστηκε το video σε $m=8$ κομμάτια κανάλια video πολυεπίπεδου βάθους, Depth-Layered Multi-Channel (DLMC) videos. Αυτό γίνεται διαιρώντας το σύνολο του εύρους βάθους σε 8 ίσα στρώματα βάθους , και για κάθε στρώμα βάθους κρατήθηκαν τα εικονοστοιχεία (pixels) , που εμπίπτουν στο αντίστοιχο στρώμα βάθους.



Παράδειγμα εικόνας DLMC για ένα συγκεκριμένο κανάλι. Πηγή : (Ofli *et al.* ,2013) at [25]

Στην συνέχεια εξήχθησαν τα χαρακτηριστικά HOG/HOF από τα DLMC video και κβαντοποιήθηκαν ξεχωριστά. Τα κανάλια C-1 , C-2 και C-8 δεν έχουν σημαντικό βάθος οπότε δεν έχουν και χαρακτηριστικά.

Μετά τον υπολογισμό των codewords , υπολογίστηκαν τα m ιστογράμματα για κάθε ακολουθία βάθους και έγιναν πειράματα κατηγοριοποίησης. Στον παρακάτω πίνακα

φαίνονται τα αποτελέσματα για κάθε μέθοδο κατηγοριοποίησης που χρησιμοποιήθηκε

:

	Channel- 3	Channel- 4	Channel- 5	Channel- 6	Channel- 7	MKL
1-NN	77.37	35.40	28.47	35.77	58.03	
3-NN	76.28	30.29	19.34	29.93	55.84	
K-SVM	70.07	24.45	24.09	27.01	40.15	91.24

Απολέσματα κατηγοριοποίησης δράσης (σε ποσοστό %) από δεδομένα βάθους για τα 5 διαφορεικά κανάλια επιπέδων βάθους. Πηγή : (Ofli *et al.* ,2013) at [25]

Το στρώμα βάθους $m=3$ δίνει τα καλύτερα αποτελέσματα.

Με χρήση του αλγορίθμου MKL έχουμε αποτέλεσμα 91.24% επιτυχούς κατηγοριοποίησης .

3) Mocap Data

Με βάση τις πόζες του σκελετού που εξάγονται από τα δεδομένα καταγραφής κίνησης , υπολογίστηκαν τις γωνίες των αρθρώσεων για 21 αρθρώσεις από την ιεραρχία του σκελετού , επιτρέποντας να παρουσιαστεί η ακολουθία δράσης ως μία συλλογή από 21-διαστάσεων χαρακτηριστικών γωνιών αρθρώσεων.

Κάθε ακολουθία δράσης χωρίστηκε σε $N=60$ χρονικά παράθυρα και υπολογίστηκε η αλλαγή των γωνιών των αρθρώσεων , σε κάθε χρονικό παράθυρο. Έτσι

παρουσιάστηκε κάθε ακολουθία δράσης , σαν ένα σετ 60 χαρακτηριστικών μεγέθους 21.

Στην συνέχεια κβαντοποιήθηκαν τα χαρακτηριστικά αυτά σε 20 codewords , και εκτελέστηκαν πειράματα κατηγοριοποίησης. Τα αποτελέσματα φαίνονται στον παρακάτω πίνακα :

	Motion Capture
1-NN	74.82
3-NN	75.55
K-SVM	79.93

Απολέσματα κατηγοριοποίησης δράσης (σε ποσοστό %) από δεδομένα καταγραφής κίνησης. Πηγή : (Ofli *et al.* ,2013) at [25]

4) Accelerometer Data

Κάθε ένα από τα επιταχυνσιόμετρα παρέχει δεδομένα επιτάχυνσης, 3-διαστάσεων (στους άξονες X-Y-Z) , στον χρόνο όσο πραγματοποιείται η διαδικασία της δράσης. Όπως και πριν εξήχθησαν N=30 χρονικά παράθυρα από κάθε ακολουθία δεδομένων για κάθε επιταχυνσιόμετρο και υπολογίστηκε η διακύμανση της επιτάχυνσης προς κάθε κατεύθυνση , σε κάθε ένα από τα 30 χρονικά παράθυρα.

Ενώνοντας τις τιμές διακύμανσης και από τα 6 επιταχυνσιόμετρα , παίρνουμε για κάθε χρονικό παράθυρο ένα 3X6= 18-διαστάσεων περιγραφέα χαρακτηριστικών. Ως εκ τούτου , παρουσιάστηκε μία ακολουθία δράσης σαν ένα σετ από 30 χαρακτηριστικά μεγέθους 18.

Στην συνέχεια κβαντοποιήθηκαν τα χαρακτηριστικά αυτά σε 20 codewords , και εκτελέστηκαν πειράματα κατηγοριοποίησης. Τα αποτελέσματα φαίνονται στον παρακάτω πίνακα:

	Accelerometer
1-NN	79.20
3-NN	81.75
K-SVM	85.40

Απολέσματα κατηγοριοποίησης δράσης (σε ποσοστό %) από δεδομένα επιταχυνσιόμετρων. Πηγή : (Ofli *et al.* ,2013) at [25]

Το ποσοστά επιτυχούς κατηγοριοποίησης με χρήση επιταχυνσιόμετρων είναι αρκετά μεγάλα.

5) Audio Data

Με τον ίδιο τρόπο έγιναν και τα πειράματα με την χρήση μικροφώνων , με χρήση 60 χρονικών παραθύρων στην 18-διαστάσεων χρονοσειρά MFCC συντελεστών [37].

Τα αποτελέσματα φαίνονται στον παρακάτω πίνακα:

	Audio
1-NN	32.12
3-NN	28.10
K-SVM	35.04

Απολέσματα κατηγοριοποίησης δράσης (σε ποσοστό %) από δεδομένα ήχου. Πηγή : (Ofli *et al.* ,2013) at [25]

Τα αποτελέσματα επιτυχούς κατηγοριοποίησης με την χρήση ήχου , όπως αναμενόταν ήταν μικρά , αν και αρκετά μεγαλύτερα από το ποσοστό 9% που έχει η επιλογή στην τύχη.

3.4.5 ΜΟΝΟΤΡΟΠΙΚΑ (UNIMODAL) ΑΠΟΤΕΛΕΣΜΑΤΑ

Συγκρίνοντας όλα τα παραπάνω αποτελέσματα μεμονωμένα για κάθε τρόπο καταλήγουμε στο συμπέρασμα , ότι ο καλύτερος τρόπος αναγνώρισης της δράσης είναι με την χρήση καμερών (φτάνει έως 99.27% επιτυχούς πρόβλεψης). Μετά είναι η χρήση δεδομένων βάθους , η χρήση επιταχυνσιομέτρων , τα δεδομένα καταγραφής κίνησης και τέλος ο ήχος.

Από τους αλγορίθμους κατηγοριοποίησης ο αλγόριθμος K-SVM , έδινε συνήθως τα καλύτερα αποτελέσματα σε σχέση με τους αλγορίθμους 1-NN και 3-NN. Η χρήση MKL έδινε πάντα τα καλύτερα αποτελέσματα.

3.4.6 ΠΟΛΥΤΡΟΠΙΚΑ (MULTIMODAL) ΑΠΟΤΕΛΕΣΜΑΤΑ

Συνδυάζοντας διάφορους τρόπους καταγραφής κίνησης μεταξύ τους μέσω του Multiple Kernel Learning (MKL) , παίρνουμε τα αποτελέσματα αναγνώρισης δράσης που φαίνονται στον παρακάτω πίνακα :

Τρόποι	Ποσοστά αναγνώρισης δράσης (%)
MC + Acc	97.45
Kin + Vid	99.27
MC + Kin	93.80
Acc + Kin	97.81
Vid + Aud	99.27
MC + Acc + Kin	98.18
MC + Acc + Aud	97.45
Vid + Kin + Aud	98.91
MC + Acc + Kin + Aud	98.54
MC + Acc + Kin + Vid	100.00
MC + Acc + Kin + Vid + Aud	100.00

Ενσωμάτωση πολλαπλών μεθόδων χρησιμοποιώντας MKL. Mocap(MC) , Accelerometer(Acc) , Kinect(Kin) , Video(Vid) , Audio(Aud). Πηγή : (Ofli *et al.* ,2013) at [25]

Όπως ήταν αναμενόμενο , ο συνδυασμός περισσότερων τρόπων μεταξύ τους αύξησε τις επιδόσεις αναγνώρισης κίνησης. Για παράδειγμα συνδυασμός mocap και accelerometer , από 79.93% και 85.40% ποσοστά επιτυχίας αναγνώρισης κίνησης αντίστοιχα , ανέβηκε στο 97.45%.

Επίσης ο συνδυασμός της μεθόδου του ήχου με κάποια άλλη μέθοδο , δεν συνιστά απαραίτητα την μείωση του ποσοστού της αναγνώρισης για αυτήν την μέθοδο , καθώς η MKL μαθαίνει τον καλύτερο δυνατό γραμμικό συνδυασμό.

3.5 ΑΝΑΚΑΤΑΣΚΕΥΗ ΤΗΣ ΚΙΝΗΣΗΣ ΜΕ ΧΡΗΣΗ

ΑΡΑΙΩΝ ΔΕΔΟΜΕΝΩΝ ΕΠΙΤΑΧΥΝΣΙΟΜΕΤΡΩΝ

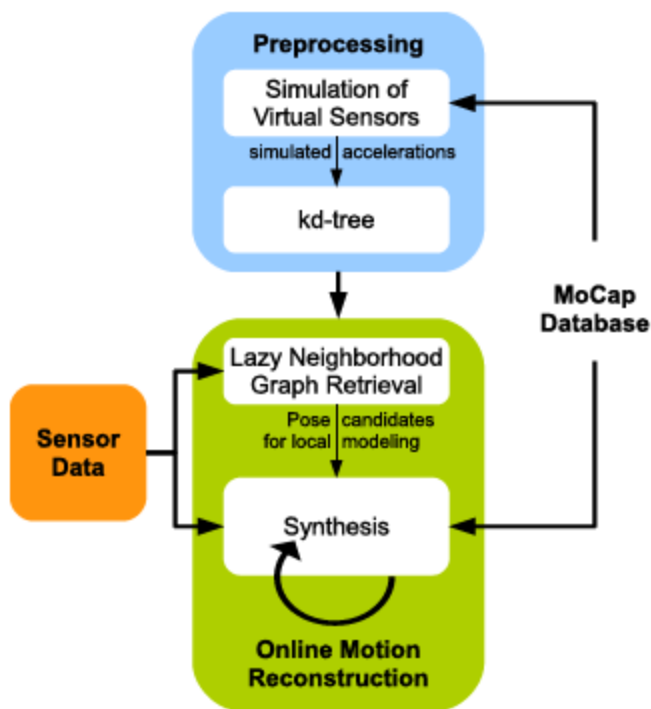
Σε αυτήν την ενότητα θα παρουσιάσουμε ένα πλαίσιο από τον JOCHEN TAUTGES [38] για την παραγωγή animation για όλο το ανθρώπινο σώμα , με χρήση μόνο τεσσάρων 3D επιταχυνσιόμετρων που είναι συνδεδεμένα στα άκρα ενός ανθρώπου. Η προσέγγιση αυτή βασίζεται σε μία γνωστή βάση δεδομένων καταγραφής κίνησης την HDM05.

3.5.1 ΕΙΣΑΓΩΓΗ

Η αύξηση των διαθέσιμων και ποιοτικών δεδομένων καταγραφής κίνησης (mocap data) έχει γίνει κινητήρια δύναμη για χρήση data-driven μεθόδων για την δημιουργία animation. Θα παρουσιάσουμε ένα data-driven σύστημα για την παραγωγή animations για ολόκληρο το ανθρώπινο σώμα. Ως είσοδο ελέγχου χρησιμοποιήθηκαν 4 3D επιταχυνσιόμετρα. Οι κινήσεις δόθηκαν σε μία μορφή από μία βάση γνώσης που αποτελείται από έναν μεγάλο αριθμό από ακολουθίες κινήσεων, που έχουν καταγραφεί από marker-based mocap συστήματα. Στην προσέγγιση που μελετάμε η βάση γνώσης (HDM05) , περιέχει κινήσεις διαφορετικών τύπων από διαφορετικά άτομα. Στην φάση της προεπεξεργασίας , αντλήθηκαν

κατάλληλα προσομοιωμένες μετρήσεις επιταχυνσιόμετρων από αποθηκευμένες ακολουθίες κίνησης στην βάση γνώσης(HDM05), που τις καθιστά συγκρίσιμες με την είσοδο από τους αισθητήρες. Επιπλέον για μεταγενέστερη χρήση , στην βάση γνώσης χρησιμοποιήθηκε η δομή k-dimensional tree (kd-tree). Ένα kd-tree [39] είναι μία δομή δεδομένων που χρησιμοποιείται για την γρήγορη επίλυση προβλημάτων κοντινότερου γείτονα. Ένα kd-tree είναι μία γενίκευση ενός δυαδικού δέντρου αναζήτησης , που αποθηκεύει τα σημεία σε χώρο κ-διαστάσεων. Μπορούμε να χρησιμοποιήσουμε μία kd-tree δομή για να αποθηκεύσουμε σημεία στο καρτεσιανό επίπεδο , σε τρισδιάστατο χώρο. Κατά τον χρόνο εκτέλεσης , η είσοδος του αισθητήρα προκαλεί μία αναζήτηση nearest-neighbor(NN) για κάθε frame(πλαίσιο). Για κάθε τρέχον frame , οι ανακτώμενες πόζες χρησιμοποιούνται για να κάνουν ενημέρωση μιας δομής δεδομένων, που παραπέμπει σε υποακολουθίες της κινήσεις στην βάση γνώσης, όπου εξηγούν καλύτερα την είσοδο με βάση τα προηγούμενα frames. Αυτή η δομή δεδομένων που ορίζεται ως online lazy neighborhood graph(OLNG) , χρησιμοποιήθηκε στην συνέχεια στο βήμα της ανακατασκευής για να υπολογίσει το animation του τρέχοντος frame.

Στην εικόνα 14.1 βλέπουμε την γενική επισκόπηση του συστήματος που θα αναλύσουμε.



ΕΙΚΟΝΑ 14.1 . Γενική επισκόπηση συστήματος δημιουργίας animation. Πηγή : (Tautges *et al.* , 2011) at [38]

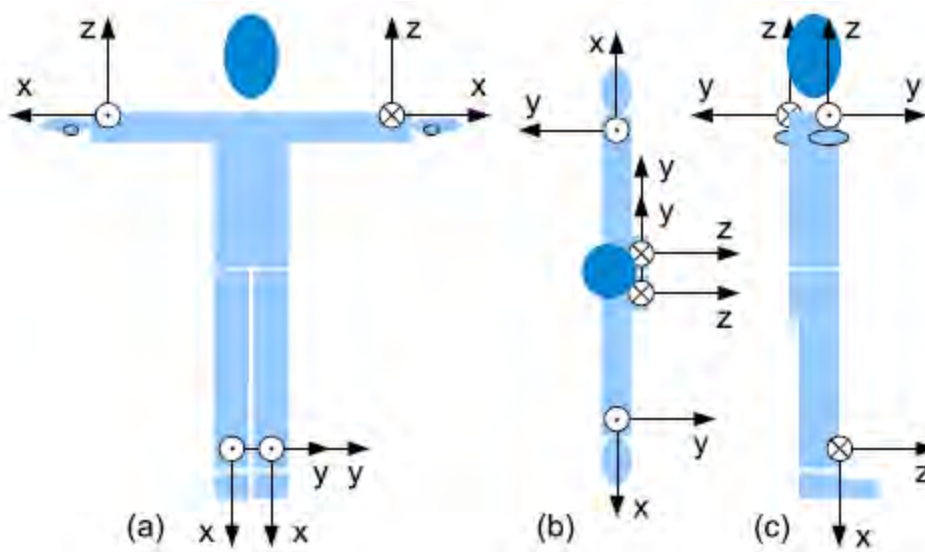
3.5.2 ΕΛΕΓΧΟΣ ΕΙΣΟΔΟΥ ΚΑΙ ΒΑΣΗ ΓΝΩΣΗΣ

Σε αυτήν την προσέγγιση χρησιμοποιήθηκαν 4 συσκευές Xsens MTx (Xsens 2010) , που παρέχουν την είσοδο ελέγχου στο πλαίσιο ανακατασκευής μας. Οι αισθητήρες είναι τοποθετημένες στα κάτω μέρη των χεριών και ποδιών , δίπλα απο τους καρπούς και τους αστραγάλους αντίστοιχα. Αυτές οι συσκευές δίνουν πολλές πληροφορίες για τον ρυθμό της στροφής , το μαγνητικό πεδίο και τον προσανατολισμό. Χρησιμοποιήθηκαν μόνο οι βαθμονομημένες μετρήσεις από τις συσκευές επιταχυνσιόμετρων , κάτι που σημαίνει ότι μπορούν να χρησιμοποιηθούν πολύ πιο μικρές και φθηνές συσκευές (αισθητήρες) , χρησιμοποιώντας μόνο επιταχυνσιόμετρα.

Για να γίνουν αυτά τα δεδομένα εισόδου συγκρίσιμα με τα δεδομένα από την βάση γνώσης , θα πρέπει οι 4 συσκευές να τοποθετηθούν κατάλληλα πάνω στα άκρα του ανθρώπου έτσι ώστε :

1. Οι άξονες X , των αισθητήρων, στα χέρια και στα πόδια να δείχνουν μακριά από το κέντρο του σώματος. Δηλαδή όταν το χέρι είναι τεντωμένο οριζόντια ο άξονας X δείχνει οριζόντια με το έδαφος και μακριά από το σώμα. Ο άξονας X στον αισθητήρα στο πόδι δείχνει προς το πάτωμα όταν ο άνθρωπος είναι όρθιος.
2. Ο άξονας Z , στην συσκευή στα χέρια, δείχνει κάθετα στο έδαφος προς τον ουρανό όταν το χέρι είναι τεντωμένο οριζόντια με το έδαφος και την παλάμη να κοιτάει προς το έδαφος.
3. Οι άξονες Z στα πόδια δείχνουν προς τα εμπρός.
4. Οι άξονες Y είναι κάθετες στα X , Z και δείχνουν μακριά από το σώμα του ανθρώπου σε κάθε άκρο.

Στην εικόνα 14.2 φαίνεται η αναπαράσταση της ιδανικής ρύθμισης των αισθητήρων.



ΕΙΚΟΝΑ 14.2 . Σχηματική αναπαράσταση των αισθητήρων που προσαρμόστηκαν στο άτομο. Πηγή : (Tautges , 2012)

Έγινε η υπόθεση ότι η βάση γνώσης αποτελείται από μία ακολουθία απο πόζες, από ένα σύνολο $[1 : N] := \{1, \dots, N\}$, με το N να δηλώνει τον συνολικό αριθμό των frames. Επίσης θεωρήθηκε ότι κάθε πόζα δίνεται από την γωνία των αρθρώσεων και συμβολίζεται με q_n , $n \in [1 : N]$. Για να βρεθούν οι θέσεις των αρθρώσεων μιας πόζας , πρέπει να εφαρμοστεί εμπρός κινηματική(forward kinematics) [40] ,με βάση ένα συγκεκριμένο μοντέλο σκελετού. Ο σκελετός περιέχει πληροφορίες σχετικά με την τοπολογία και τα μήκη των οστών του actor , όπως και τους βαθμούς ελευθερίας κάθε άρθρωσης. Στα επόμενα που θα μελετήσουμε έγινε η υπόθεση ότι όλοι οι σκελετοί από την βάση γνώσης έχουν την ίδια τοπολογία.

Ένας βασικός μηχανισμός της προσέγγισης που μελετάμε είναι η αναγνώριση μεγάλων σε διαστάσεις δεδομένων γωνιών των αρθρώσεων , χρησιμοποιώντας χαμηλών σε διαστάσεις δεδομένων επιταχυνσιομέτρων. Σε αυτό το σενάριο , πρέπει να γίνει σύγκριση 2 διαφορετικών δεδομένων κίνησης που παριστάνονται από

διαφορετικές διαστάσεις. Για να γεφυρωθεί αυτό το χάσμα , προσομοιώθηκαν οι μετρήσεις επιτάχυνσης για όλες τις κινήσεις από την βάση γνώσης , υπολογίζοντας όλες τις επιταχύνσεις των εικονικών αισθητήρων που τοποθετούνται στα άκρα των εικονικών actors , με τον ίδιο τρόπο όπως και οι πραγματικοί αισθητήρες στους πραγματικούς actors. Μετά τον υπολογισμό των θέσεων αυτών των εικονικών αισθητήρων με χρήση forward kinematics , υπολογίστηκαν οι δεύτερες παράγωγοι στον χρόνο και να πάρουμε τις σχετικές επιταχύνσεις στο συνολικό frame. Μετά προστέθηκε η συνιστώσα της επιτάχυνσης που αντιστοιχεί στην βαρύτητα , η οποία μετράται από κάθε επιταχυνσιόμετρο και μετατράπηκε το προκύπτων αποτέλεσμα , στο τοπικό σύστημα συντεταγμένων των εικονικών αισθητήρων.

Για να γίνει η προσομοίωση των δεδομένων των αισθητήρων , υπολογίστηκαν από πριν κάποιες ποσότητες που θα χρησιμοποιηθούν στην φάση της σύνθεσης της κίνησης. Αυτές οι ποσότητες περιλαμβάνουν τις θέσεις x_n , τις ταχύτητες u_n και τις επιταχύνσεις a_n των αρθρώσεων. Οι τύποι των παραγώγων είναι οι εξής:

$$u(t) = \frac{-x(t+2h) + 8x(t+h) - 8x(t-h) + x(t-2h)}{12h} , \text{ για την πρώτη παράγωγο.}$$

$$a(t) = \frac{-u(t+2h) + 16u(t+h) - 30u(t) + 16u(t-h) - u(t-2h)}{12h^2} , \text{ για την δεύτερη παράγωγο.}$$

Για όλες τις δοκιμές χρησιμοποιήθηκε ένας σκελετός όπου τα μήκη των οστών του είναι ο μέσος όρος του μήκους των όλων σκελετών που υπάρχουν στην βάση γνώσης HDM05. Επίσης παραμελήθηκαν οι αρθρώσεις στα πόδια και χέρια του σκελετού , με αποτέλεσμα να έχουμε συνολικά 21 αρθρώσεις και συνολικά 43 βαθμούς ελευθερίας περιστροφών.

Οι προσομοιωμένες επιταχύνσεις των αισθητήρων , που συμβολίζονται με a_n μετατρέπονται με χρήση kd-tree διάστασης $4*3=12$. Σε αυτές τις χαμηλές διαστάσεις , οι Andoni and Indyk [41] , αναφέρουν ότι τα kd-trees είναι κατάλληλα για γρήγορες αναζητήσεις πλησιέστερου γείτονα. Στην περίπτωση που μελετάμε έγιναν αναζητήσεις πλησιέστερου γείτονα για να προσδιοριστούν οι πόζες στην βάση γνώσης που είναι παρόμοιες με τις μετρήσεις των αισθητήρων.

3.5.3 ONLINE LAZY NEIGHBORHOOD GRAPH (OLNG)

Στο σενάριο που μελετάμε , η είσοδος ελέγχου συγκρίνεται με ακολουθίες mocap δεδομένων από έτοιμες βάσεις δεδομένων , με βάση τις 3D επιταχύνσεις. Οι Slyper και Hodgins[2008] , χρησιμοποίησαν την χρονική αλληλουχία των κινήσεων, με την αναζήτηση σταθερού μήκους ακολουθιών από αναγνώσεις επιταχύνσεων. Με τον τρόπο αυτό όμως δεν μπορούσαν στην προσέγγιση τους να αντιμετωπίσουν χρονικές μεταβολές , όπως πχ. διάφορες κινήσεις που εκτελούνται σε διαφορετικές ταχύτητες.

Στην προσέγγιση του Tautges που μελετάμε έγινε η ενσωμάτωση της χρονικής αλληλουχίας , χρησιμοποιώντας μία δομή δεδομένων που αναφέρεται ως online lazy neighborhood graph (OLNG). Η δομή OLNG που χρησιμοποιήθηκε στην προσέγγιση αυτή είναι παρόμοια με την δομή lazy neighborhood graph(LNG) που παρουσίασε ο [Kruger το 2010] αλλά με ορισμένες βελτιώσεις σε σχέση με την αρχική μέθοδο του Kruger. Θα δούμε πώς μπορεί η LNG να δημιουργηθεί σταδιακά κάνοντας την αποτελεσματική και ικανή να λειτουργήσει online.

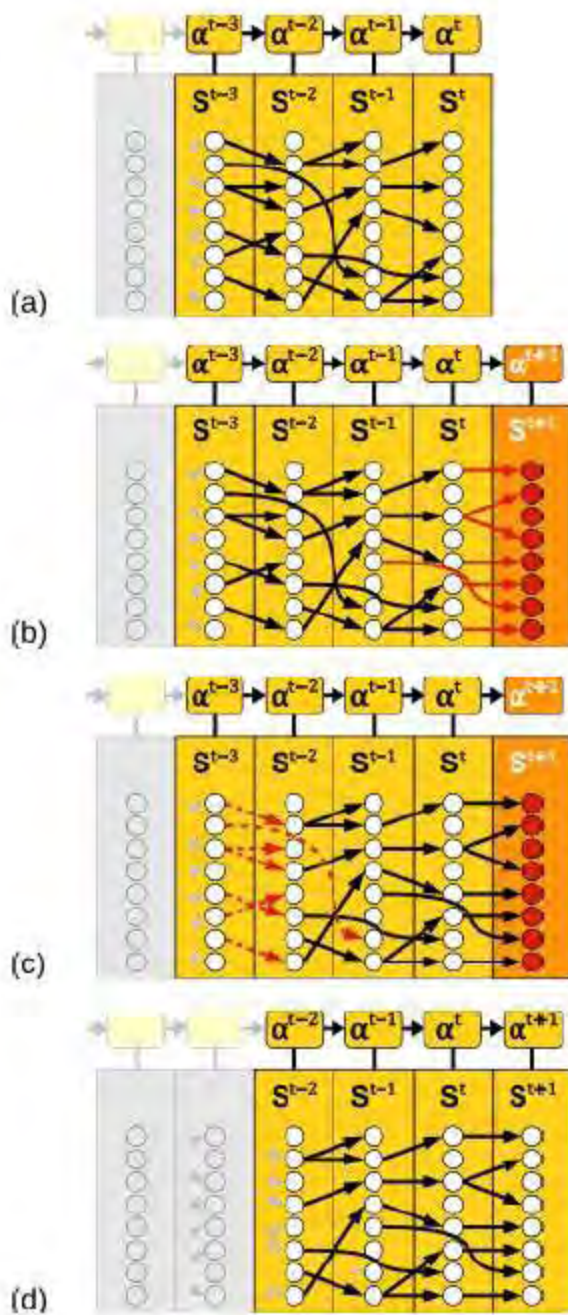
Η βασική ιδέα του OLNG συνοψίζεται ως εξής : δεδομένου ότι κάθε πλαίσιο(frame) στην βάση γνώσης έχει έναν μοναδικό δείκτη (με διαδοχικά frame να έχουν διαδοχικούς δείκτες) , αντί να ψάχνουμε μία ακολουθία κίνησης που καθολικά ταιριάζει με το σήμα εισόδου , μπορούμε να ψάξουμε μία ακολουθία από διαδοχικά αύξουσες δείκτες των οποίων τα αντίστοιχα frames τοπικά ταιριάζουν με το σήμα εισόδου.

Θεωρήθηκε ότι η είσοδος ελέγχου αποτελείται από ένα (μετά από δειγματοληψία) συνεχές ρεύμα από επιταχύνσεις αισθητήρων $(\dots, a^{t-2}, a^{t-1}, a^t, \dots)$ όπου το a^t δηλώνει τις επιταχύνσεις την χρονική στιγμή $t \in Z$, και θέλουμε να βρούμε πόζες που ταιριάζουν με την πραγματική απόδοση στον χρόνο t . Για λόγους που είπαμε παραπάνω, θα γίνει εκμετάλλευση της χρονικής αλληλουχίας των κινήσεων , με αναζήτηση ακολουθιών επιταχύνσεων αντί με απλά δείγματα του χρόνου. Για να βρεθούν αυτές οι ακολουθίες θεωρήθηκαν οι τελευταίες M μετρήσεις αισθητήρων $(a^{t-M+1}, \dots, a^{t-1}, a^t)$. Με άλλα λόγια , ψάχνουμε για ακολουθίες επιτάχυνσης στην βάση γνώσης που ταιριάζουν με τις μετρήσεις των αισθητήρων κατά την διάρκεια των τελευταίων M χρονικών πλαισίων.

Θέτουμε με K , τον αριθμό των (τοπικά) πλησιέστερων γειτόνων , και με S^t το σύνολο μεγέθους K , των δεικτών που αντιπροσωπεύουν τους K πλησιέστερους γείτονες του a^t , όπου μπορεί να υπολογιστεί αποδοτικά χρησιμοποιώντας kd-tree. Στόχος τώρα είναι να προσδιοριστεί ένα υποσύνολο του συνόλου S^t , που είναι αξιόπιστο για την δημιουργία ενός τοπικού στατιστικού μοντέλου για τις πόζες την χρονική στιγμή t . Θεωρήθηκε ότι μία πόζα είναι αξιόπιστη , αν μία ακολουθία κίνησης κατάλληλου μήκους που περιέχει αυτήν την πόζα, ταιριάζει με το σήμα ελέγχου. Το

ταίριασμα των ακολουθιών μπορεί να βρεθεί με τον προσδιορισμό των μοναπατιών από διαδοχικούς δείκτες , σε μία δομή δεδομένων όπως η LNG που παρουσίασε ο Kruger. Προκειμένου να έχουμε μία αριθμητική τιμή αξιοπιστίας της υπό εξέτασης πόζας , πρέπει να βρεθεί για κάθε κόμβο στο σύνολο S^t , το μονοπάτι με το μικρότερο κόστος το οποίο οδηγεί μέχρι τον συγκεκριμένο κόμβο.

Οι κόμβοι (πιθανές πόζες) του ONLG ορίζονται από τα πλαίσια που δίνει η k -nearest-neighbors αναζήτηση και από κατευθυνόμενες ακμές που κωδικοποιούν την χρονική αλληλουχία των πλαισίων , μεταξύ των κόμβων που δημιουργούν έγκυρες συνέχειες. Ο γράφος του ONLG κατασκευάζεται από τους κόμβους που ορίζονται ως n_k^t με $1 < t \leq M$ και $1 \leq k \leq K$ που αναγνωρίζονται με τον μοναδικό δείκτη της βάσης δεδομένων $i(n_k^t)$, και εισάγοντας ακμές μεταξύ των κόμβων που αποτελούν μία έγκυρη συνέχεια. Στην παρακάτω εικόνα 16(a) βλέπουμε μία παρουσίαση του γράφου:



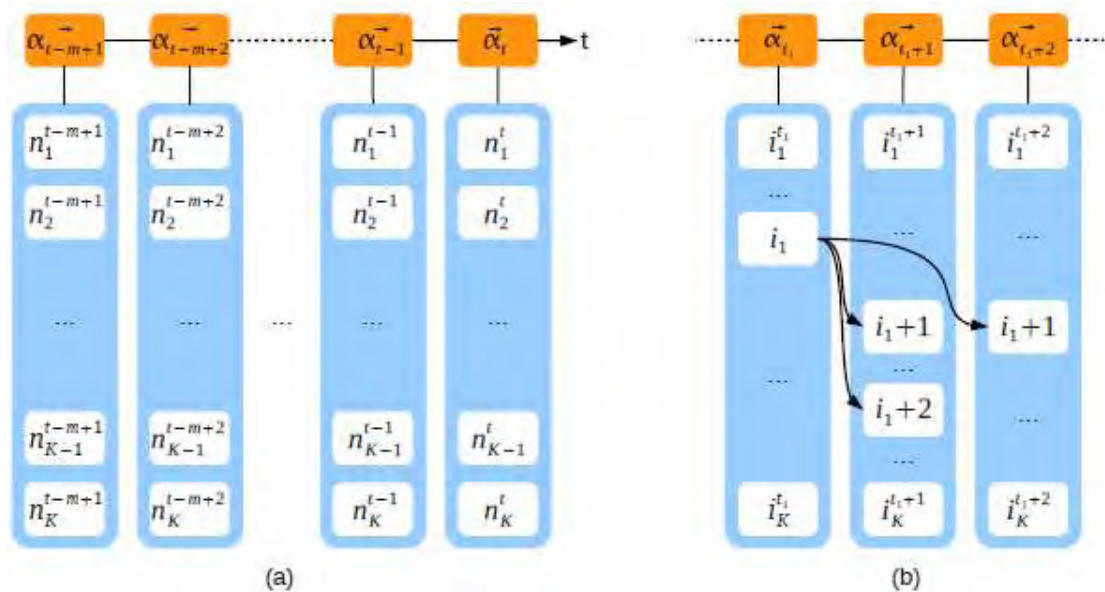
Εικόνα 16. Online Lazy Neighborhood Graph (OLNG) για $M=4$ και $K=8$. Πηγή : (Tautges *et al.* , 2011) at [38]

Αυτές οι έγκυρες συνέχειες ορίστηκαν ως εξής : το $i_2 = i(n_{k_2}^{t_2})$ είναι μία έγκυρη συνέχεια του $i_1 = i(n_{k_1}^{t_1})$, εάν μία από τις παρακάτω 3 συνθήκες ισχύουν :

- $i_1 + 1 = i_2$ και $t_1 + 1 = t_2$.

- $i_1 + 2 = i_2$ και $t_1 + 1 = t_2$
- $i_1 + 1 = i_2$ και $t_1 + 2 = t_2$

Ο γράφος του OLNG με τους κόμβους φαίνεται στην εικόνα 15(a) , ενώ οι τρεις προϋποθέσεις δημιουργίας ακμής στην εικόνα 15(b).



ΕΙΚΟΝΑ 15 . (a) Οι $M \times K$ κόμβοι του OLNG γράφου. (b) Οι ακμές που δημιουργούνται από τις προϋποθέσεις έγκυρης συνέχειας. Πηγή : (Tautges , 2012)

Κάθε μονοπάτι παράγει μία ακολουθία από δείκτες , όπου αυτή η ακολουθία αντιστοιχεί σε μία υπακολουθία κίνησης στην βάση γνώσης. Τα κόστη των μονοπατιών υπολογίζονται όπως και στην περίπτωση του LNG από τον Kruger αλλά με μερικές παραλλαγές και κάποια πρόσθετα κόστη.

Στον LNG το κόστος του κάθε κόμβου συνδέεται με την ανάλογη απόσταση από τα αποτελέσματα της k-nearest-neighbor αναζήτησης. Το κόστος κάθε μονοπατιού δίνεται από το άθροισμα όλων των σχετικών κόμβων. Τώρα θέλουμε να βρούμε τα μονοπάτια, που διασχίζουν τον γράφο από αριστερά προς δεξιά, με το ελάχιστο

κόστος. Με την προσθήκη ενός κόμβου πηγής S στο γράφημα και με την σύνδεση του με όλους αριστερούς κόμβους n_k^1 με $1 \leq k \leq K$, το πρόβλημα εύρεσης του μονοματιού με το μικρότερο κόστος μετατρέπεται σε πρόβλημα *single-source shortest-paths*, όπου μπορεί να λυθεί σε γραμμικό χρόνο.

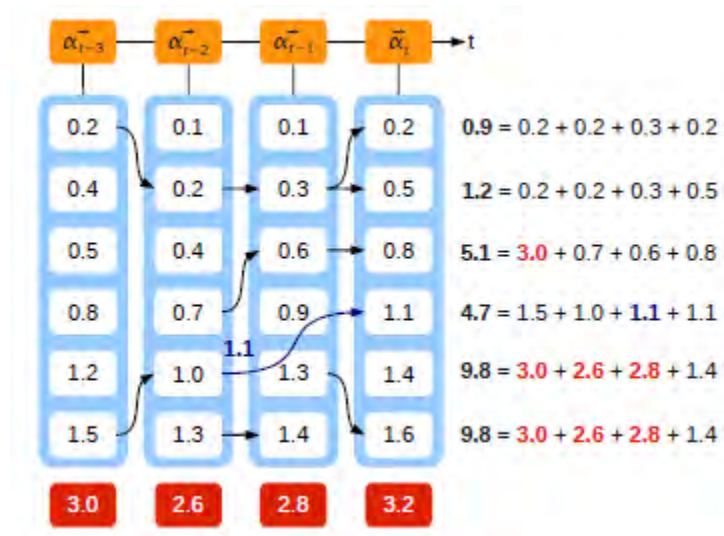
Στην OLNG προσέγγιση που μελετάμε, τα επιπλέον κόστη που αναφέραμε προηγουμένως είναι απαραίτητα, αφού τα μήκη των μονοπατιών, εν αντιθέση με την LNG προσέγγιση, δεν έχουν το ίδιο μήκος. Αν και τα μεγαλύτερα μονοπάτια προτιμώνται, περιέχουν περισσότερους κόμβους που προσθέτουν επιπλέον κόστη στο συνολικό κόστος των μονοπατιών, έτσι τα μικρότερα σε μήκος μονοπάτια πρέπει να “τιμωρούνται” κατάλληλα. Αυτό γίνεται με την χρησιμοποίηση ενός επιπλέον κόστους, το οποίο για κάθε χρονική στιγμή t δεν υπάρχει στο μονοπάτι. Στην περίπτωση που μελετάμε αυτό το κόστος υπολογίστηκε ως το διπλάσιο κόστος του K μακρινότερου γείτονα στην συγκεκριμένη χρονική στιγμή t . Για τον ίδιο ακριβώς λόγο πρέπει να “τιμωρηθούνε” και τα μικρότερα μονοπάτια. Εισήχθηκε ένα κόστος για τις ακμές που παραλείπουν μία στήλη (που προκύπτουν από την προϋπόθεση 3 της έγκυρης συνέχειας), όπου το κόστος των ακμών αυτών συνδέεται με ένα από τα κόστη των κόμβων που συνδέει. Οι ακμές που δεν παραλείπουν μία στήλη έχουν κόστος 0.

Έτσι το τελικό κόστος των μονοπατιών δίνεται από το άθροισμα των παρακάτω:

- *Το κόστος όλων των κόμβων που σχετίζονται με το μονοπάτι, μετρώντας τις διαφορές από τα δείγματα επιτάχυνσης.*
- *Το κόστος όλων των ακμών που σχετίζονται με το μονοπάτι.*

- Το επιπλέον κόστος που εισήχθηκε προηγουμένως για τα μικρότερα σε μήκος μονοπάτια.

Με βάση αυτά τα κόστη είναι πλέον εύκολο να συγκριθούν οι πόζες στο σύνολο S^t και να κρατηθεί μόνο ένα υποσύνολο για την κατασκευή του τοπικού στατιστικού μοντέλου για τις πόζες. Στην εικόνα 17 βλέπουμε ένα παράδειγμα υπολογισμού του κόστους των μονοπατιών, με τους αριθμούς στους κόμβους να ορίζουν τα κόστη κάθε κόμβου.



ΕΙΚΟΝΑ 17. Παράδειγμα υπολογισμού κόστους μονοπατιών στον OLNG με $M=4$ και $K=6$. Πηγή : (Tautges, 2012)

Τώρα θα περιγράψουμε την διαδικασία για την αποτελεσματική ενημέρωση του OLNG. Έγινε η υπόθεση ότι ο OLNG έχει ήδη κατασκευαστεί για τις μετρήσεις (a^{t-M+1}, \dots, a^t) , και πώς για κάθε κόμβο στο σύνολο S^t έχει βρεθεί το μονοπάτι με το ελάχιστο κόστος που οδηγεί σε κάθε κόμβο. Τώρα μία νέα μέτρηση a^{t+1} φτάνει. Αρχικά για την a^{t+1} μέτρηση, βρίσκονται οι K πλησιέστεροι γείτονες με την χρήση του kd-tree και αποθηκεύονται στο σύνολο S^{t+1} . Ο OLNG επεκτείνεται με την εισαγωγή

μίας νέας στήλης από K κόμβους. Εισάγονται οι νέες ακμές με τέτοιο τρόπο ώστε να πληρούνται οι προϋποθέσεις έγκυρης συνέχειας (εικόνα 16(b)) , με συνέπεια να επεκταθούν τα προηγούμενα μονοπάτια ελάχιστου κόστους. Τέλος οι κόμβοι που αντιστοιχούν στο a^{t-M+1} , μαζί με τις εμπλεκόμενες ακμές αφαιρούνται για να πάρουμε τον ανανεωμένο πλέον OLNG(εικόνα 16(d)). Αυτή η σταδιακή κατασκευή του OLNG, μας δίνει την δυνατότητα χρήσης του σε online εφαρμογές.

3.5.4 ΑΝΑΚΑΤΑΣΚΕΥΗ ΤΗΣ ΚΙΝΗΣΗΣ

Ο στόχος της ανακατασκευής της κίνησης είναι να προσεγγίσει μία κίνηση που έχει εκτελεστεί. Το σύστημα οδηγείται από χαμηλού επιπέδου είσοδο ελέγχου. Είναι αδύνατον να ανακατασκευαστεί άμεσα μία κίνηση πολλών διαστάσεων. Έτσι για να γίνει η εκτίμηση μιας κίνησης ολόκληρου του σώματος, οι βαθμοί ελευθερίας που λείπουν από τα δεδομένα επιταχυνσιόμετρων , θα πρέπει να συντεθούν χρησιμοποιώντας την γνώση που υπάρχει στις υπάρχουσες βάσεις δεδομένων κίνησης. Μία μέθοδος για να γίνει η σύνθεση της κίνησης είναι να δημιουργηθεί μία νέα κίνηση που βασίζεται σε παρόμοιες ήδη υπάρχουσες πόζες ή κινήσεις.

Χρησιμοποιήθηκε η δομή OLNG , που παρέχει για κάθε χρονικό τμήμα t , ένα σετ από K μονοπάτια με το αντίστοιχο κόστος τους. Στην πράξη τα περισσότερα μονοπάτια έχουν μικρό μήκος και μεγάλο κόστος. Εξετάστηκαν μόνο τα I μονοπάτια που είναι υποσύνολο των K , $I \ll K$, τα οποία έχουν το μικρότερο κόστος.

Ορίστηκε ως $C^t = \{c_1^t, \dots, c_I^t\}$ το κόστος αυτών των I μονοπατιών σε κάθε χρονικό βήμα t .

Ορίστηκε ως $Q^t = \{q_1^t, \dots, q_l^t\}$ το σύνολο των διαμορφωμένων γωνιών που δίνεται από τα τελευταία πλαίσια αυτών των μονοπατιών.

Ορίστηκε ως $X^t = \{x_1^t, \dots, x_l^t\}$ τις θέσεις.

Ορίστηκε ως $V^t = \{u_1^t, \dots, u_l^t\}$ τις ταχύτητες.

Ορίστηκε ως $A^t = \{a_1^t, \dots, a_l^t\}$ τις επιταχύνσεις.

Όλες αυτές οι ποσότητες υπολογίζονται στο στάδιο της προ-επεξεργασίας και μπορούν να ληφθούν από την βάση γνώσης σε πραγματικό χρόνο.

Ορίστηκαν ως $W^t = \{w_1^t, \dots, w_l^t\}$ τα κανονικοποιημένα βάρη των μονοπατιών βασιζόμενοι στα κόστη τους, όπου κάθε τιμή w_i^t δίνεται από τον τύπο :

$$w_i^t = \frac{\max(C^t) - c_i^t}{\sum_{j=1}^l (\max(C^t) - c_j^t)} \quad (1).$$

Με αυτόν τον τρόπο έχει διατυπωθεί πλέον το πρόβλημα ανακατασκευής της κίνησης ως πρόβλημα ελαχιστοποίησης της ενέργειας.

Για κάθε χρονικό βήμα προσπαθούμε να βρούμε μία πόζα q_{best} που βέλτιστα ικανοποιεί τους περιορισμούς που επιβάλλονται (από τα δεδομένα επιτάχυνσης), ενώ είναι συγχρόνως και συνεπής με παρόμοια frames κίνησης που έχουν ανακτηθεί από την βάση δεδομένων. Ακριβέστερα, το πρόβλημα της ελαχιστοποίησης της συνάρτησης ενέργειας βασίζεται σε 2 συνιστώσες :

- Τον προγενέστερο όρο δεδομένων (prior term) που επιβάλλει εύλογα αποτελέσματα ανακατασκευής της κίνησης.
- Τον όρο ελέγχου (control term) που οδηγείται από τα μετρούμενα δεδομένα επιτάχυνσης.

Έχουμε την εξής σχέση:

$$q_{best} = \operatorname{argmin}(w_{prior} * E_{prior}(q) + w_{contr} * E_{contr}(q)). \text{ ως προς τον όρο } q. \text{ (2)}$$

Αρκεί να βρεθεί η πόζα q , που ελαχιστοποιεί την παραπάνω σχέση. Οι όροι w_{prior} και w_{contr} είναι σταθερές που καθορίζονται από τον χρήστη.

Στην συνέχεια υπολογίστηκαν οι όροι $E_{prior}(q)$ και $E_{contr}(q)$.

Θεωρήθηκε ότι ήδη ανακατασκευάστηκε η κίνηση μέχρι την χρονική στιγμή t . Τώρα στον χρόνο $t+1$ μία νέα είσοδος ελέγχου a^{t+1} έρχεται από τους αισθητήρες, που χρησιμοποιείται για να κάνει update τον OLNG, ο οποίος μας δίνει τις πρόσθετες πληροφορίες Q^{t+1} και W^{t+1} .

3.5.4.1 THE PRIOR TERM

Το μοντέλο prior που χρησιμοποιήθηκε αποτελείται από τρεις διαφορετικές συνιστώσες :

- I. Το *pose prior*, χαρακτηρίζει την πιθανότητα μιας πόζας να καθορίζεται από τα δείγματα των βάσεων δεδομένων.
- II. Το *motion prior*, μετρά την πιθανότητα μιας πόζας να αφορά την χρονική εξέλιξη μιας κίνησης.
- III. Το *smoothness prior*, που μειώνει την αντίληψη μεμονομένων εικόνων στην κίνηση (Jerkiness).

Βασιζόμενοι σε αυτό το μοντέλο η συνάρτηση ενέργειας $E_{prior}(q)$, με βάρη w_{pose} , w_{motion} και w_{smooth} ισούται με:

$$E_{prior}(q) = w_{pose} * E_{pose}(q) + w_{motion} * E_{motion}(q) + w_{smooth} * E_{smooth}(q) \quad (3).$$

- **Pose prior :**

Τα σύνολα από τις πόζες Q^{t+1} και τα αντίστοιχα βάρη W^{t+1} που παρέχονται από το OLNG , χρησιμοποιούνται για να χαρακτηρίσουν την συνάρτηση πυκνότητας πιθανότητας στον χώρο της πόζας. Χρησιμοποιήθηκε μία προσέγγιση βασισμένη σε πυρήνα για να γίνει η προσέγγιση της πιθανότητας p_{pose} να συνθεθεί μία υποψήφια πόζα q .

$$p_{pose}(q) \propto \sum_{i=1}^I w_i^{t+1} * K(|q_i^{t+1} - q|) \quad (4)$$

Το K είναι μία συμμετρική συνάρτηση πυρήνα.

Υπολογίστηκε ο όρος E_{pose} , με χρήση της τετραγωνικής ρίζας του kernel , για να έχουμε ελαχιστοποίηση της ενέργειας.

$$E_{pose} = \sum_{i=1}^I w_i^{t+1} * \sqrt{(|q_i^{t+1} - q|)} \quad (5)$$

- **Motion prior :**

Η κίνηση των αρθρώσεων από πόζα σε πόζα , θα πρέπει να γίνεται με αληθοφανή τρόπο.

Βρέθηκαν οι θέσεις των αρθρώσεων την χρονική στιγμή $t+1$, υπολογίζοντας την επέκταση δεύτερης τάξης Taylor , στις θέσεις x^t (που συνδέονται με τα q^t) , χρησιμοποιώντας τα V^{t+1} και A^{t+1} .

Για το δείγμα i όπου $i \in \{1, \dots, n\}$ η εκτιμώμενη θέση x_i^{t+1} , δίνεται από τον τύπο :

$$x_i^{t+1} = x^t + u_i^{t+1} * \Delta t + \frac{1}{2} a_i^{t+1} * \Delta t^2 \quad (6)$$

Όμοια με πριν έχουμε :

$$E_{motion}(q) = \sum_{i=1}^I w_i^{t+1} * \sqrt{(|x_i^{t+1} - x|} \quad (7)$$

- **Smoothness prior :**

Μία πόζα q (με θέσεις x), θα είναι λογική εάν οι επιταχύνσεις από τους αισθητήρες είναι συνεπής με τις επιταχύνσεις από τα δεδομένα βάσεων δεδομένων. Όπως και πριν η πιθανότητα μιας υποψήφιας πόζας μετράται με kernel βασισμένο στην εκτίμηση της πυκνότητας (Kernel density estimation) :

$$E_{smooth}(q) = \sum_{i=1}^I w_i^{t+1} * \Delta t * \sqrt{(|a_i^{t+1} - a|} \quad (8)$$

$$\mu = \Delta t^{-2} * (x - 2x^2 + x^{t-1}) \quad (9)$$

3.5.4.2 THE CONTROL TERM

Οι επιταχύνσεις έχουν ήδη χρησιμοποιηθεί για την ανάκτηση υποακολουθιών κίνησης (άρα και πόζες), που είναι πολύ πιθανόν να είναι παρόμοιες με τις κινήσεις που όντως πραγματοποιούνται. Καθώς η επακόλουθη σύνθεση της κίνησης βασίζεται σε αυτές τις πόζες, αυτό το βήμα παρέχει ήδη έναν βαθμό ελέγχου και περιορίζει τα πιθανά αποτελέσματα. Παρόλα αυτά η χρήση αυτών των επιταχύνσεων για σήμα

ελέγχου δεν εγγυάται ομοιότητα στον χώρο της πόζας , που είναι απαραίτητη για την ανακατασκευή μιας σταθερής κίνησης. Ακριβώς για αυτόν τον λόγο το control term υπολογίζεται με βάση τις ελάχιστες ή μέγιστες(extremal) τιμές των θέσεων που ταιριάζουν με τις πραγματικές θέσεις του αισθητήρα.

Υποθέτοντας τις σωστές θέσεις x^t την χρονική στιγμή t , τότε η πυκνότητα πιθανότητας για τις θέσεις των αρθρώσεων την χρονική στιγμή $t+1$, υπολογίζεται με την αριθμητική ολοκλήρωση της εξίσωσης της κίνησης με χρήση V^t .

$$x_i^{t+1} = \langle x^t + u^t * \Delta t \rangle + \frac{1}{2} a^t * \Delta t^2 \quad (10).$$

Υποθέτοντας ότι η βάση δεδομένων περιλαμβάνει παρόμοιες κινήσεις με αυτές που πραγματοποιούνται , χρησιμοποιήθηκε το x_i^{t+1} για να πάρουμε τον εξής όρο που πρέπει να ελαχιστοποιηθεί:

$$E_{contr}(x) = \sum_{i=1}^l w_i^{t+1} * \sqrt{(x_i^{t+1} - \langle x \rangle)} \quad (11).$$

3.5.4.3 ENERGY MINIMIZATION

Υιοθετήθηκε μία gradient-descent-based μέθοδο (την συνάρτηση Isqnonlin του Matlab) [42] , για την ελαχιστοποίηση της συνάρτησης (2) :

$$q_{best} = \text{argmin}(w_{prior} * E_{prior}(q) + w_{contr} * E_{contr}(q)).$$

Η μέθοδος συγκλίνει γρήγορα μετά από μερικές επαναλήψεις.

Τα βάρη που επιλέγει ο χρήστης στην εξίσωση 2 , κρατήθηκαν σταθερά με τις παρακάτω τιμές:

- $w_{contr} = 1$
- $w_{prior} = 6$

- $w_{pose}=0.6$
- $w_{motion}=0.2$
- $w_{smooth}=0.2$

Η μικρή αλλαγή αυτών των τιμών δεν επηρεάζουν την συνολική ποιότητα των αποτελεσμάτων μας.

3.5.5 ΑΠΟΤΕΛΕΣΜΑΤΑ

Η αποτελεσματικότητα του συστήματος δοκιμάστηκε με προσομοίωση καθώς και με πραγματικά δεδομένα αισθητήρων. Για όλες τις αριθμητικές συγκρίσεις χρησιμοποιήθηκε το μέσο σφάλμα RMS (Root mean square) των θέσεων.

$$RMS_{error} = \sqrt{\frac{\sum_{i=1}^n (y'_i - y_i)^2}{n}} . \text{ όπου } y_i \text{ είναι η παρατηρούμενη τιμή και } y'_i \text{ είναι η}$$

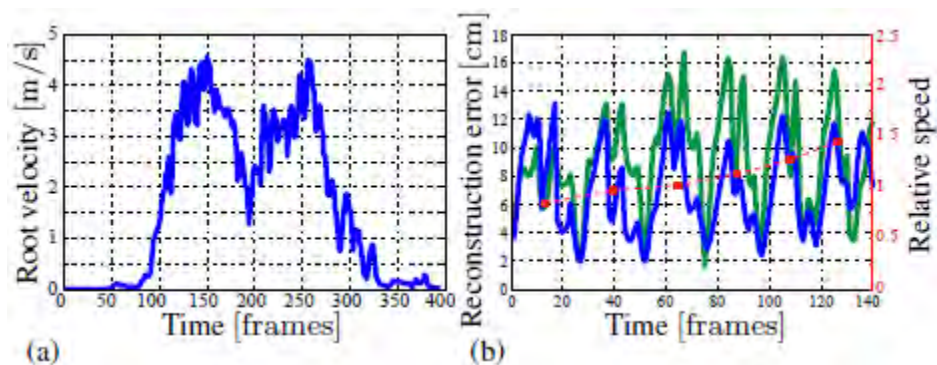
προβλεπόμενη τιμή.

➤ ΑΠΟΤΕΛΕΣΜΑΤΑ ΒΑΣΙΣΜΕΝΑ ΣΤΑ ΔΕΔΟΜΕΝΑ ΤΩΝ ΑΙΣΘΗΤΗΡΩΝ

Η ανακατασκευή της κίνησης σε αυτήν την περίπτωση έγινε με την βοήθεια 2 ατόμων , σε εξωτερικό περιβάλλον. Σε αντίθεση με τα οπτικά συστήματα καταγραφής κίνησης , οι αδρανειακές συσκευές που χρησιμοποιήθηκαν στην περίπτωση που μελετάμε , επιβάλλουν πολύ μικρούς περιορισμούς στα άτομα που εκτελούν τις δράσεις , το περιβάλλον εκτέλεσης της δράσης , τις συνθήκες φωτισμού και το γενικό setup.

Λόγω της δυνατότητας χρονικής στρέβλωσης της δομής OLNG , η προσέγγιση αυτή δεν είναι ευαίσθητη στις χρονικές διακυμάνσεις. Η ανάλυση περιορίστηκε σε μία ακολουθία από κινήσεις jumping jack (επιτόπιο άλμα όπου τα πόδια ανοίγουν

διάπλατα, τα χέρια ακουμπούν μεταξύ τους πάνω από το κεφάλι και στην συνέχεια επιστρέφουν στην φυσική τους θέση τόσο τα πόδια όσο και τα χέρια), που εκτελέστηκαν σε διαφορετικές ταχύτητες. Τα αποτελέσματα του πειράματος (το σφάλμα ανακατασκευής της κίνησης), δίνονται στην εικόνα 18, μαζί με μία παραλλαγή της μεθόδου όπου δεν λαμβάνει υπόψη την χρονική στρέβλωση.

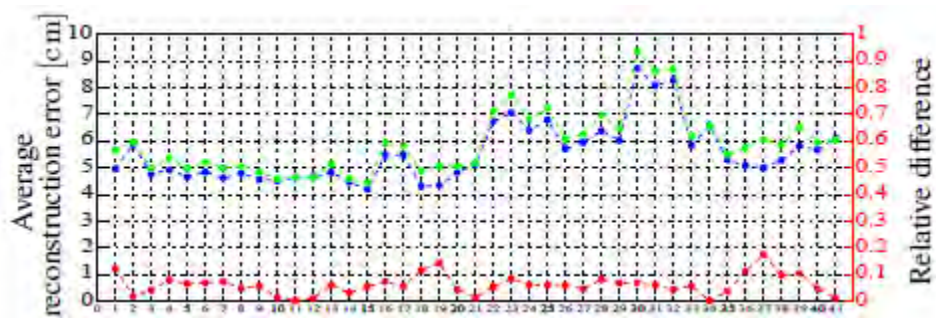


Εικόνα 18. (a) Εκτιμώμενη ταχύτητα για την δράση τρέξιμο-περπάτημα-τρέξιμο.(b) Μεταβολή της σχετικής ταχύτητας για μία κίνηση jumping jack στον χρόνο.Η διακεκομμένη κόκκινη γραμμή μας δείχνει την σχετική ταχύτητα και οι κόκκινες κουκίδες τα παλαμάκια μετά από 6 κινήσεις jumping jack. **Μπλε:** Το σφάλμα ανακατασκευής της κίνησης με την μέθοδο που μελετάμε. **Πράσινο:** Το σφάλμα ανακατασκευής της κίνησης με μία μέθοδο που δεν λαμβάνει υπόψη την χρονική στρέβλωση της κίνησης. Πηγή : (Tautges *et al.*, 2011) at [38]

Παρατηρούμε ότι τα σφάλματα ανακατασκευής της μεθόδου που μελετάμε είναι αρκετά μικρότερα από την άλλη μέθοδο. Αυτό δείχνει το πλεονέκτημα της δομής OLNG, σε σχέση με έναν συμβατικό γραμμικό αλγόριθμο αναζήτησης που δεν λαμβάνει υπόψη τις χρονικά στρεβλωμένες κινήσεις.

Για να γίνει μία ποσοτική σύγκριση μεταξύ των πραγματικών δεδομένων με δεδομένα προσομοίωσης (MoCap data), καταγράφηκαν 41 διαφορετικές κινήσεις

χρησιμοποιώντας τους αισθητήρες Xsens' MTx sensors συγχρόνως με ένα οπτικό σύστημα καταγραφής κίνησης με 12 κάμερες. Μετά την ανακατασκευή αυτών των κινήσεων και με τα 2 ήδη δεδομένων, τα σφάλματα ανακατασκευής φαίνονται στην εικόνα 19.



ΕΙΚΟΝΑ 19. **Μπλε:** Σφάλματα ανακατασκευής με χρήση πραγματικών αισθητήρων , για κάθε κίνηση.
Πράσινο: Σφάλματα ανακατασκευής από δεδομένα προσομοίωσης(επιταχύνσεις) για κάθε κίνηση.
Κοκκίνο: Η διαφορά των σφαλμάτων μεταξύ τους. Πηγή : (Tautges *et al.* , 2011) at [38]

Από την εικόνα 15 παρατηρούμε ότι τα σφάλματα ανακατασκευής με χρήση δεδομένων προσομοίωσης είναι ελαφρώς μεγαλύτερα από τα πραγματικά δεδομένα. Παρόλα αυτά , και οι 2 καμπύλες σφαλμάτων είναι πολύ όμοιες μεταξύ τους , με πολύ μικρές ποσοτικές διαφορές. Αυτές οι παρατηρήσεις μας δίνουν το δικαίωμα να χρησιμοποιηθούν ήδη καταγεγραμμένα δεδομένα προσομοίωσης. Χρησιμοποιήθηκαν τα δεδομένα από την βάση δεδομένων HDM05 [43].

➤ ΑΠΟΤΕΛΕΣΜΑΤΑ ΒΑΣΙΣΜΕΝΑ ΣΕ ΠΡΟΣΟΜΟΙΩΜΕΝΑ ΔΕΔΟΜΕΝΑ ΑΙΣΘΗΤΗΡΩΝ (HDM05)

Σε αυτήν την ενότητα θα μελετήσουμε πώς η ανακατασκευή της κίνησης επηρεάζεται από το μέγεθος και την ποικιλομορφία της χρησιμοποιούμενης βάσης δεδομένων καταγραφής κίνησης . Επιπλέον θα μελετήσουμε την επίδραση του μήκους του παραθύρου M που χρησιμοποιείται στον OLNG. Θα δούμε πώς το μέγεθος του ατόμου που εκτελεί τις δράσεις επηρεάζει τα αποτελέσματα , ιδικά στην

περίπτωση που είναι πολύ πιο μεγάλος ή μικρόσωμος από τα άτομα που χρησιμοποιήθηκαν για την δημιουργία της βάσης δεδομένων HDM05. Στο τέλος αξιολογούμε πώς ο αριθμός και η τοποθέτηση των αισθητήρων επηρεάζει την ποιότητα των αποτελεσμάτων μας.

1) Τα σενάρια που χρησιμοποιήθηκαν για τις δοκιμές:

Όπως αναφέρθηκε από πριν , η βάση γνώσης που χρησιμοποιήθηκε ήταν η βάση δεδομένων HDM05. Οι κινήσεις που περιέχει η HDM05 πραγματοποιήθηκαν από 5 διαφορετικά άτομα , που αναφέρονται με τα αρχικά τους (bd, bk, dg, mm, tr). Στην συνέχεια ορίστηκαν διαφορετικές βάσεις γνώσεις , με το ίδιο πρότυπο ονομασίας που χρησιμοποιήθηκε στην τεκμηρίωση του HDM05 ([Müller et al. 2007]) , για να περιγράψει τα αρχεία κίνησης :

HDM_{actor}_{part}-{scene}_{take}_{framerate}.

Όταν χρησιμοποιούμε αστερίσκο (*) σε ένα πεδίο, ο αστερίσκος αυτός αντιπροσωπεύει όλες τις δυνατές τιμές του πεδίου. Με χρήση του "M" , προστίθενται αντίγραφα αντικατοπτρισμού, με φυσική συμμετρία του άξονα του σκελετού(αντιστρέφοντας "αριστερά" και "δεξιά"). Με χρήση του "R" , προστίθενται και time-reversed αντίγραφα στην βάση γνώσης. Για παράδειγμα η βάση γνώσης HDM_bd_01-**-**_25M , αντιπροσωπεύει όλα τα motion clips του μέρους 1 , που τα εκτέλεσε το άτομο bd , μαζί με τα αντίγραφα αντικατοπτρισμού.

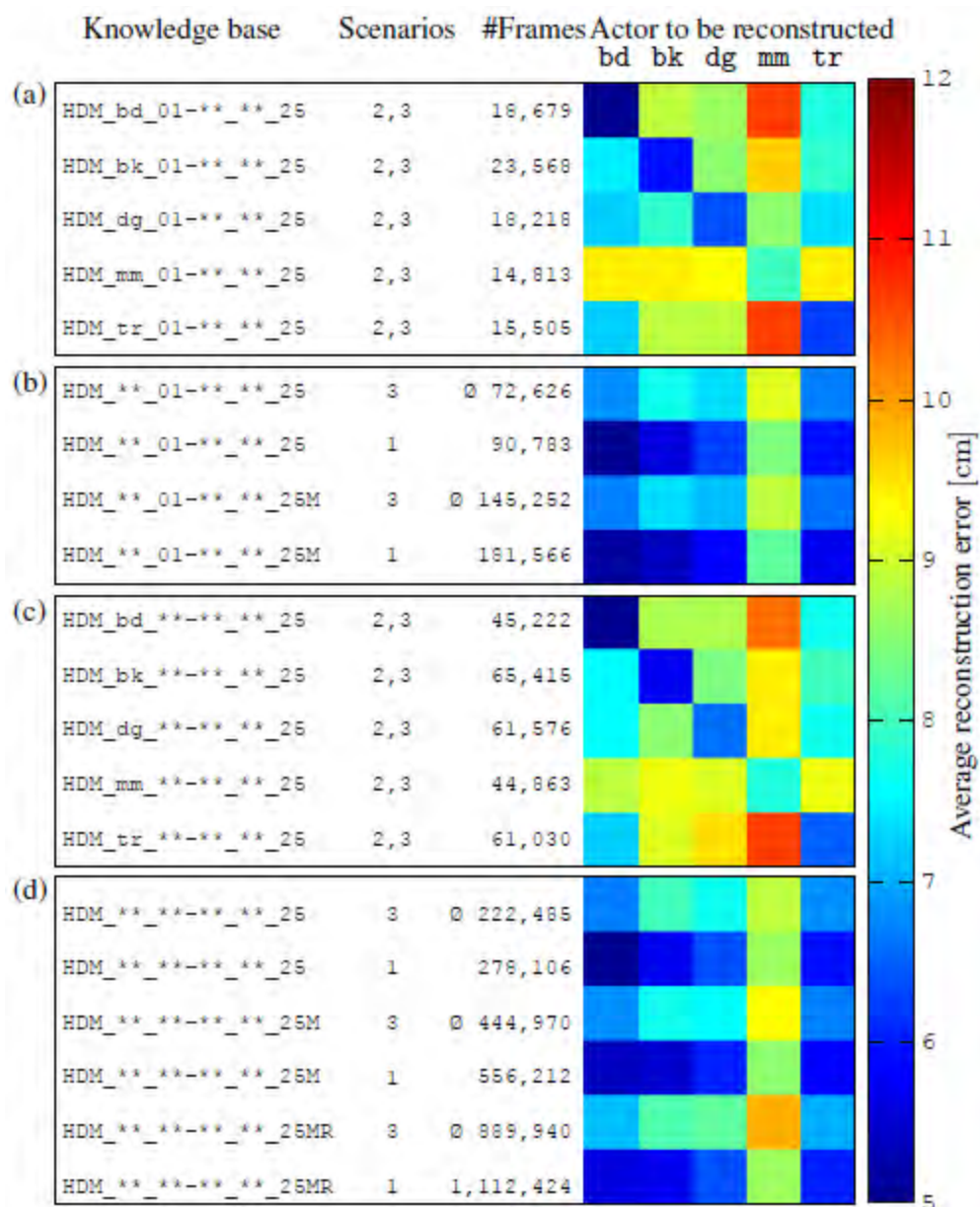
Σε όλα τα πειράματα , τα χρησιμοποιούμενα δεδομένα ελέγχου ελήφθησαν από εικονικούς αισθητήρες, όπως είχαμε πεί σε προηγούμενη παράγραφο , χρησιμοποιώντας ένα σύνολο από κινήσεις δοκιμής από την HDM05. Η πραγματική κίνηση δοκιμής ποτέ δεν συμπεριλήφθει στην βάση γνώσης. Ορίστηκαν 3

διαφορετικά σενάρια ανακατασκευής της κίνησης , όπου το είδος κάθε σεναρίου καθορίζεται από αν το άτομο που πραγματοποιεί την δοκιμή της κίνησης , περιελήφθει στην βάση γνώσης ή όχι. Στα πειράματα μας κάνουμε χρήση 3 διαφορετικών σεναρίων (βάση γνώσης):

- I. Περιέχει επίσης και τις κινήσεις του ατόμου, που θα ανακατασκευαστούν εκτός της συγκεκριμένης κίνησης(DB1) (Κινήσεις και από τα 5 άτομα).
- II. Περιέχει μόνο τις κινήσεις του ατόμου, που θα ανακατασκευαστούν(DB2).
- III. Δεν περιέχει κινήσεις του ατόμου, που θα ανακατασκευαστούν(DB3).

2) Το μέγεθος και η ποικιλομορφία της βάσης γνώσης:

Δημιουργήθηκαν 5 σύνολα κινήσεων δοκιμής , ένα για κάθε άτομο, όπου το καθένα περιέχει 6 motion clips από το part 1(περπάτημα-τρέξιμο-άλματα) της βάσης HDM05. Επιπλέον έγινε σύνθεση ενός συνόλου από 20 διαφορετικές βάσεις γνώσης , που διέφεραν μεταξύ τους σε μέγεθος και ποικιλομορφία , και χρησιμοποιήθηκαν για την ανακατασκευή των κινήσεων δοκιμής όλων των ατόμων. Στην συνέχεια αυτές οι κινήσεις συγκρίθηκαν με τις πρωτότυπες κινήσεις δοκιμής , με χρήση του RMS error των σημείων των αρθρώσεων. Τα αποτελέσματα φαίνονται στην εικόνα 20.



ΕΙΚΟΝΑ 20. Μέσο σφάλμα ανακατασκευής για μία βάση γνώσης, για κάθε actor. Πηγή : (Tautges *et al.*, 2011) at [38]

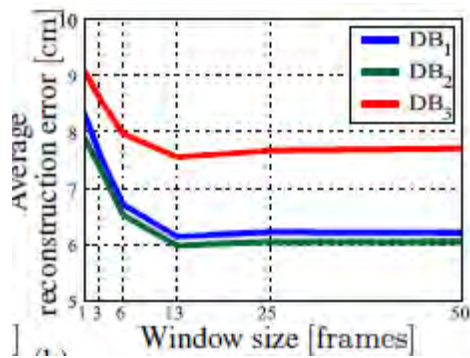
Στην εικόνα 20 φαίνονται οι 20 βάσεις γνώσης , τα σενάρια που προκύπτουν από τον συνδυασμό της βάσης γνώσης με τις αντίστοιχες κινήσεις δοκιμής, το μέγεθος της βάσης γνώσης (σε frames) , και για κάθε άτομο το μέσο RMS error.

Παρατηρούμε ότι τα αποτελέσματα είναι πολύ καλύτερα στα σενάρια 1 και 2 , όπου συμπεριλαμβάνονται οι κινήσεις του ατόμου στην χρησιμοποιούμενη βάση γνώσης. Αυτό φαίνεται από την εικόνα 20(b) και 20(d) αν παρατηρήσουμε τις διαφορές γραμμή-γραμμή ,όπως επίσης και απο την εικόνα 20(a) και 20(c) αν παρατηρήσουμε τις μπλέ διαγωνίους που αντιπροσωπεύουν το σενάριο 2. Επίσης παρατηρούμε ότι τα αποτελέσματα για τον mm δεν είναι τόσο καλά και αυτό μπορεί να οφείλεται στο στυλ δράσης του συγκεκριμένου ατόμου.

Μία άλλη σημαντική παρατήρηση είναι ότι τα αποτελέσματα γίνονται ελαφρώς χειρότερα όταν η βάση γνώσης μεγαλώνει και είναι λιγότερο ομοιογενής. Αυτό φαίνεται με σύγκριση των σειρών , που αντιστοιχούν στα ίδια σενάρια , στις εικόνες 20(b) και 20(d). Γενικά η επίδραση της ποικιλομορφίας (actor mm) , παίζει μεγαλύτερο ρόλο από την απλή αύξηση του μεγέθους της βάσης γνώσης.

3) Μέγεθος παραθύρου M :

Η σύνθεση της κίνησης σε μεγάλο βαθμό εξαρτάται από τον OLNG. Η τιμή της παραμέτρου M, που χρησιμοποιείται για την ανάκτηση των μονοπατιών , ίσως παίζει σημαντικό ρόλο. Ωστόσο η παράμετρος M , στην προσέγγιση που μελετάμε , ορίζει μόνο ένα άνω όριο στο μήκος των μονοπατιών. Άρα η τιμή του δεν είναι τόσο κρίσιμη όπως φαίνεται και στην παρακάτω εικόνα.



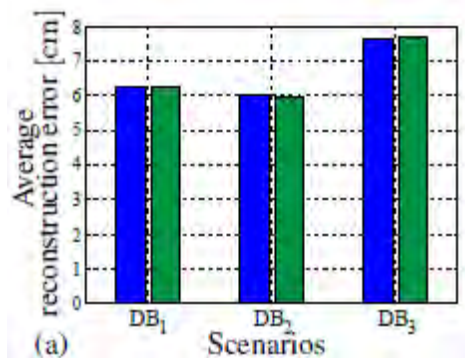
Πηγή : (Tautges *et al.* , 2011) at [38]

Παρατηρούμε ότι για τιμές $M \geq 6$, το μέσο σφάλμα παραμένει περίπου το ίδιο.

4) Μέγεθος και αναλογίες των ατόμων που εκτελούν τις δράσεις:

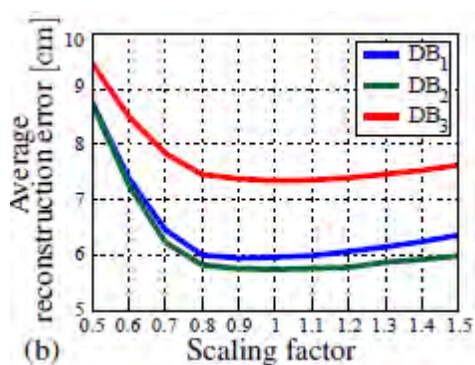
Εκτελέστηκαν 2 πειράματα για να δούμε την επίδραση του μεγέθους του ατόμου στην ανακατασκευή της κίνησης.

Στο πρώτο πείραμα χτίστηκε μία βάση γνώσης , χρησιμοποιώντας πληροφορίες του αρχικού σκελετού , για πέντε διαφορετικά άτομα με ύψη από 170cm έως 200cm , που περιλαμβάνονται στην βάση HDM05. Όπως βλέπουμε από την παρακάτω εικόνα , το σφάλμα ανακατασκευής είναι σχεδόν ανεπηρέαστο από αλλαγές του μεγέθους του actor.



Εξάρτηση του μέσου σφάλματος ανακατασκευής με το μέγεθος και τις αναλογίες των ατόμων. **Μπλε:** Αρχικές αναλογίες στην βάση γνώσης. **Πράσινο:** Χρήση σκελετών με μέσο μήκος οστών. Πηγή : (Tautges *et al.* , 2011) at [38]

Στο δεύτερο πείραμα χρησιμοποιήθηκε πάλι βάση γνώσεων πάνω σε σκελετούς με βάση τον μέσο όρο των μηκών των οστών από όλα τα άτομα. Τώρα όμως έγινε κλιμάκωση στα μήκη των οστών , του ατόμου όπου θα ανακατασκευαστεί η κίνηση , με εύρος [0.5-1.5]. Μετά την ανακατασκευή της κίνησης τα μήκη των οστών μετατρέπονται πάλι στα κανονικά μεγέθη για να γίνει η σύγκριση. Στην επόμενη εικόνα βλέπουμε τα αποτελέσματα για κάθενα από τα 3 ήδη σεναρίων:

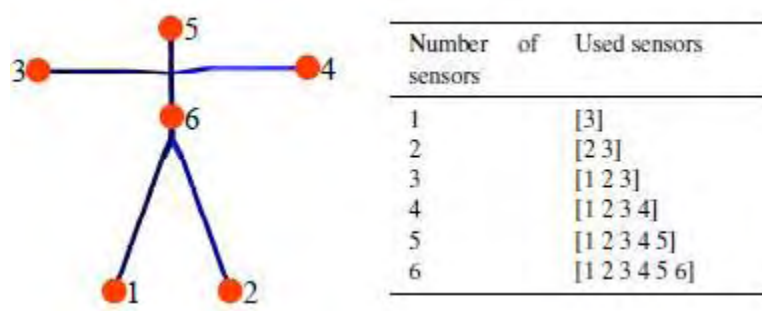


Πηγή : (Tautges *et al.* , 2011) at [38]

Παρατηρούμε ότι με τιμή scaling factor από 0.8 και πάνω τα αποτελέσματα παραμένουν περίπου ίδια και για τα 3 ήδη σεναρίων. Άρα ούτε τώρα το μέγεθος και οι αναλογίες του ατόμου παίζουν σημαντικό ρόλο στην ανακατασκευή της κίνησης.

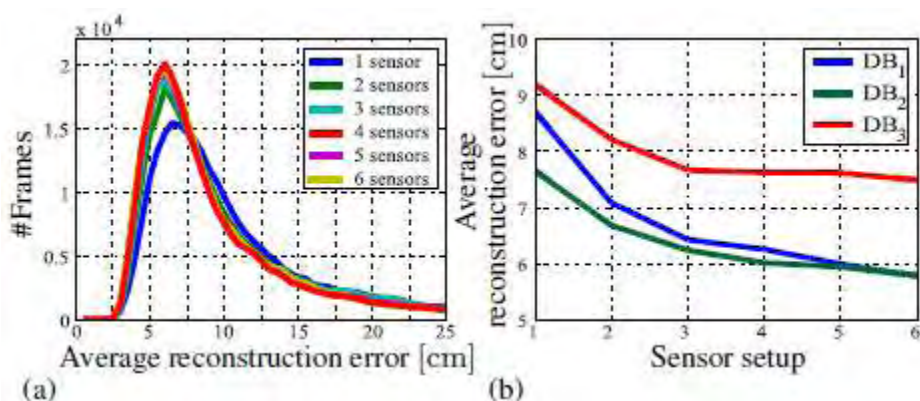
5) Δοκιμές με διαφορετικές ρυθμίσεις των αισθητήρων:

Έγιναν πειράματα με διαφορετικό αριθμό αισθητήρων σε διάφορα σημεία του σώματος, όπως φαίνεται στην παρακάτω εικόνα:



Πηγή : (Tautges *et al.*, 2011) at [38]

Τα αποτελέσματα φαίνονται στην παρακάτω εικόνα :



Πηγή : (Tautges *et al.*, 2011) at [38]

Γενικά , επιπλέον αισθητήρες τείνουν να βελτιώσουν την ποιότητα της ανακατασκευής. Παρόλα αυτά ένα πολύ μεγάλο μέρος κινήσεων μπορεί να ανακατασκευαστεί , με πολύ καλά αποτελέσματα , και με χρήση λιγότερων αισθητήρων. Αυτό συμβαίνει κυρίως σε κινήσεις που εκτελούνται σχεδόν με τον ίδιο τρόπο από διαφορετικά άτομα(περπάτημα , τρέξιμο κτλ).

4

Επίλογος

Στην παρούσα εργασία μελετήθηκαν κάποιοι αλγόριθμοι , για την αναγνώριση της ανθρώπινης κίνησης , την ανακατασκευή της κίνησης αλλά και την ανίχνευση ή μη μιας πτώσης.

Στο δεύτερο κεφάλαιο ασχοληθήκαμε με την προσπάθεια ανίχνευσης των πτώσεων των ηλικιωμένων , με χρήση επιταχυνσιόμετρων και γυροσκοπίων κυρίως. Προτάθηκε ένα γενικό σύστημα αξιολόγησης και τέλεσης πειραμάτων.Μετά μελετήσαμε 3 διαφορετικούς αλγορίθμους και παραλλαγές αυτών. Αν και κατά περιπτώσεις είχαμε καλά αποτελέσματα στην προσπάθεια ανίχνευσης πτώσης , ακόμα οι ανιχνευτές πτώσης δεν χρησιμοποιούνται ευρέως. Ο κύριος λόγος όπως είδαμε, είναι ότι δεν υπάρχει ακόμα αυτό το σύστημα που θα έχει 100% επιτυχία και θα είναι ευκολοφόρετο, εύκολο στην χρήση και άνετο στην εφαρμογή πάνω στον χρήστη. Μελλοντικές βελτιώσεις στις συσκευές αυτές θα μπορούσαν να βελτιώσουν σε μεγάλο βαθμό το βιοτικό επίπεδο των ηλικιωμένων.

Στο τρίτο κεφάλαιο ασχοληθήκαμε με την καταγραφή της κίνησης (motion capture). Αναφέραμε διάφορες βάσεις δεδομένων καταγραφής της κίνησης με τα αρχεία που χρησιμοποιούν και τα πεδία εφαρμογής του motion capture. Μετά μελετήσαμε την

πολυτροπική βάση δεδομένων MHAD και αναφέραμε κάποιους αλγορίθμους αναγνώρισης της ανθρώπινης κίνησης. Στο τέλος μελετήσαμε έναν real-time αλγόριθμο ανακατασκευής της κίνησης με χρήση αραιών δεδομένων από επιταχυνσιόμετρα με την βοήθεια της βάσης δεδομένων HDM05.

Βιβλιογραφία

- [1] Salva A, Bolibar I, Pera G, Arias C. Incidence and consequences of falls among elderly people living in the community. Med Clin (Barc) 2004.
- [2] Lord SR, Sherrington C, Menz HB. Falls in older people: risk factors and strategies for prevention. Cambridge University Press, 2001.
- [3] Lord, C.J. and DP Colvin “Falls in the elderly: Detection and Assessment”, in Ann. Int. Conf. IEEE Eng. Med. Biol. Society, 1991.
- [4] Williams G. et al. “A smart fall and activity monitor for telecare applications”, in 20th int. Conf. IEEE-EMBS 1998, Hong Kong, 29 Oct-1 Nov, pp 1151-4, 1998.
- [5] Hwang, J.Y. et al. “Development of Novel Algorithm and Real-time monitoring Ambulatory system using Bluetooth module for fall detection in the elderly” in Proceedings of the 26th Annual International Conference of the IEEE-EMBS. San Francisco, CA, USA. September 1-5, 2004.
- [6] Mihailidis A. “An intelligent emergency response system: preliminary development and testing of automated fall detection” , J. Telemed. Telecare, 2005.
- [7] N. Noury, Senior Member, IEEE , A. Fleury, Student Member, IEEE, P. Rumeau, A.K. Bourke, G. Ó Laighin, V. Rialle, J.E. Lundy, “Fall detection – Principles and

Methods” , proceedings of the 29th Annual International Conference of the IEEE EMBS Cité Internationale, Lyon, France August 23-26, 2007.

[8] J. Shea, “An Investigation of Falls in the Elderly”, <http://www.signalquest.com/> , 2007.

[9] Garrett Brown Computer Science University of Michigan , “An Accelerometer Based Fall Detector: Development, Experimentation, and Analysis” , Summer Undergraduate Program in Engineering at Berkeley (SUPERB), 2005

[10] E. Weisstein, Dot Product, Available from Mathworld-A Wolfram Web Resource URL: <http://mathworld.wolfram.com/DotProduct.html>.

[11] E. Weisstein, Butterworth filter design, Available from Mathworld-A Wolfram Web Resource URL: <http://www.mathworks.com/help/signal/ref/butter.html>.

[12] Maarit Kangas , Antti Konttila, Per Lindgren , Ilkka Winblad , Timo Jamsa , “Comparison of low-complexity fall detection algorithms for body attached accelerometers” , Available online at www.sciencedirect.com ,2008.

[13] Yoshida T, Mizuno F, Hayasaka T, Tsubota K, Wada S, Yamaguchi T. , “Awearable computer system for a detection and prevention of elderly users from falling” , Proc ICBM , 2005.

[14] Kangas M, Konttila A, Winblad I, Jamsa T. , “Determination of simple threshold for accelerometer-based parameters for fall detection” , Proc IEEE EMBS, 2007.

[15] Najafi B, Aminian K, Loew F, Blanc Y, Robert PA., “Measurement of stand-sit and sit-stand transitions using a miniature gyroscope and its application in fall risk evaluation in the elderly”, IEEE Trans Biomed Eng ,2002.

- [16] Nyan MN, Tay FE, Tan AW, Seah KH. , “Distinguishing fall activities from normal activities by angular rate characteristics and high-speed camera characterization” , Med Eng Phys ,2006.
- [17] A.K. Bourke, G.M. Lyons , “A threshold-based fall-detection algorithm using a bi-axial gyroscope sensor” , Available online at www.sciencedirect.com , 2006.
- [18] Hsiao ET, Robinovitch SN, “Common protective movements govern unexpected falls from standing height” , Biomech ,1998.
- [19] Ashish Sharma, MukeshAgarwal, Anima Sharma,PankhuriDhuria , “MOTION CAPTURE PROCESS, TECHNIQUES AND APPLICATIONS ” , IJRITCC, 2013.
- [20] Weijun Tao , Tao Liu , Rencheng Zheng and Hutian Feng , “Gait Analysis Using Wearable Sensors”, Available online at www.mdpi.com/journal/sensors , February 2012 .
- [21] Biovision file format. Available online at <http://research.cs.wisc.edu/graphics/Courses/cs-838-1999/Jeff/BVH.html> .
- [22]ASF/AMC File Format. Available online at <http://research.cs.wisc.edu/graphics/Courses/cs-838-1999/Jeff/ASF-AMC.html> .
- [23] C3D File Format , Available online at <http://www.c3d.org> .
- [24] Vicon Motion Systems , Available online at <http://www.vicon.com> .
- [25] Ferda Ofli, Rizwan Chaudhry, Gregorij Kurillo, Ren´e Vidal and Ruzena Bajcsy, “Berkeley MHAD: A Comprehensive Multimodal Human Action Database” , 2013.
- [26]] C. Schuldt, I. Laptev, and B. Caputo, “Recognizing human actions: a local SVM approach”, In Pattern Recognition (ICPR), 2004.

- [27] M. Blank, L. Gorelick, E. Shechtman, M. Irani, and R. Basri , “Actions as space-time shapes”. In Computer Vision (ICCV), 2005.
- [28] Xiaofei He, Partha Niyogi, “Locality Preserving Projections (LPP)” , 2002.
- [29] H. Wang, M. M. Ullah, A. Klaser, I. Laptev kai C. Schmid , “. Evaluation of local spatio-temporal features for action recognition” , BMVC 2009 – British Machine Vision Conference, London, United Kingdom, 2009.
- [30] C. J. C. Burges , “A Tutorial on Support Vector Machines for Pattern Recognition”, Data Min. Knowl. Discov. , 1998.
- [31] A. Vedaldi, V. Gulshan, M. Varma, and A. Zisserman , “Multiple kernels for object detection” , In Computer Vision (ICCV), 2009.
- [32] H. Wang, M. M. Ullah, A. Klaser, I. Laptev, and C. Schmid , “Evaluation of local spatio-temporal features for action recognition” , In British Machine Vision Conf., 2009.
- [33] Navneet Dalal and Bill Triggs , “Histograms of Oriented Gradients for Human Detection” , CVPR ,2005.
- [34] Rizwan Chaudhry, Avinash Ravichandran, Gregory Hager and Rene Vidal , “βιβλιογραφία Histograms of Oriented Optical Flow and Binet-Cauchy Kernels on Nonlinear Dynamical Systems for the Recognition of Human Actions” , CVPR , 2009.
- [35] I. Laptev, M. Marszalek, C. Schmid, and B. Rozenfeld , “Learning realistic human actions from movies. In Computer Vision and Pattern Recognition” , CPVR , 2008.
- [36] B. Ni, G. Wang, and P. Moulin , “RGBD-HuDaAct: A colordepth video database for human daily activity recognition” , In Computer Vision Workshops (ICCVW), 2011.

- [37] Beth Logan , “Mel-Frequency Cepstrum Coefficients for Music Modeling” , 2000.
- [38] Jochen Tautges, Arno Zinke , Bjorn Kruger,Jan Baumann,Andreas Weber, Thomas Helten, Meinard Muller, Hans-Peter Seidel, Bernd Eberhardt , “Motion Reconstruction Using Sparse Accelerometer Data” , 2011.
- [39] KDTree , Available online at <http://web.stanford.edu/class/cs106l/handouts/assignment-3-kdtree.pdf> .
- [40] Serdar Kucuk , Zafer Bingul, “Robot Kinematics: Forward and Inverse Kinematics” , Available online at <http://cdn.intechopen.com/pdfs-wm/379.pdf> , 2006.
- [41] Andoni, Alexandr , Indyk, Piotr , “Near-optimal hashing algorithms for approximate nearest neighbor in high dimensions”, 2008.
- [42] Isqnonlin , Available online at <http://www.mathworks.com/help/optim/ug/Isqnonlin.html> .
- [43] Muller, M., Roder, T., Clausen, M., Eberhardt, B., Kruger, B., and Weber, A. , “Documentation: Mocap Database HDM05” , Computer Graphics Technical Report , 2007.

ΠΑΡΑΡΤΗΜΑ Α

ΕΠΙΤΑΧΥΝΣΙΟΜΕΤΡΟ

Το επιταχυνσιόμετρο είναι μία ηλεκτρομηχανική συσκευή που έχει την ικανότητα να μετρά δυνάμεις επιτάχυνσης. Αυτές οι επιταχύνσεις χωρίζονται σε 2 κατηγορίες :

1. Στατικές : επιτάχυνση της βαρύτητας
2. Δυναμικές : επιταχύνσεις , επιβραδύνσεις ή στροφές από αλλαγή της ταχύτητας ή της διεύθυνσης της κίνησης.

Συνδιάζοντας την χρήση επιταχυνσιομέτρων , γυροσκοπίων και ηλεκτρονικών κυκλωμάτων μπορούμε να υπολογίσουμε την θέση στο χώρο ενός αντικειμένου , την στάση και την ταχύτητα αυτού με αρκετά μεγάλη ακρίβεια.

Έχει εφαρμογές στην ενεργοποίηση του αερόσακου , στον αυτόματο πιλότο και γενικά σε οποιαδήποτε εφαρμογή απαιτείται μέτρηση επιτάχυνσης.

ΓΥΡΟΣΚΟΠΙΟ

Τα γυροσκόπια είναι συσκευές οι οποίες χρησιμοποιούν τις αρχές της αδράνειας και της διατήρησης της στροφορμής για να διατηρήσουν σταθερό τον προσανατολισμό τους σε σχέση με μια αρχικώς καθορισμένη διεύθυνση.

Το παραπάνω χαρακτηριστικό τους κάνει τα γυροσκόπια ιδανικές συσκευές για την μέτρηση της :

- Γωνιακής επιτάχυνσης.
- Γωνιακής ταχύτητας.
- Γωνίας κλίσης.

ΓΩΝΙΑΚΗ ΤΑΧΥΤΗΤΑ

Γωνιακή ταχύτητα είναι το διανυσματικό μέγεθος που εκφράζει την ταχύτητα του σώματος που εκτελεί κυκλική κίνηση. Ισούται με τον ρυθμό μεταβολής του τόξου που διαγράφει το σώμα και μετράται σε ακτίνια ανά δευτερόλεπτα (rad/sec).

Ακτίνιο είναι η μονάδα μέτρησης της γωνίας και ισούται με $180/\pi$ μοίρες.

ΓΩΝΙΑΚΗ ΕΠΙΤΑΧΥΝΣΗ

Γωνιακή επιτάχυνση ονομάζουμε τον ρυθμό μεταβολής της γωνιακής ταχύτητας ενός σώματος. Μετράται σε ακτίνια ανά δευτερόλεπτα στο τετράγωνο. (rad/sec^2).

Συμβολίζεται διεθνώς με το γράμμα a .

BLUETOOTH

Το Bluetooth είναι ένα πρωτόκολλο το οποίο παρέχει προτυποποιημένη, ασύρματη επικοινωνία ανάμεσα σε PDA ,κινητά τηλέφωνα,φορητούς υπολογιστές,προσωπικούς υπολογιστές καθώς και ψηφιακές φωτογραφικές μηχανές ή ψηφιακές κάμερες.