



ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ

ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

**ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΜΕ ΕΦΑΡΜΟΓΕΣ ΣΤΗ
ΒΙΟΙΑΤΡΙΚΗ**

**Κατασκευή Παραμετρικού Μοντέλου Ανθρώπινου
Σώματος για Αναγνώριση Δραστηριότητας σε Video**

Κότταρη Κωνσταντίνα

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

Επιστημονικός Υπεύθυνος:

Δελήμπασης Κωνσταντίνος

Επίκουρος Καθηγητής



ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ

ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

**ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΜΕ ΕΦΑΡΜΟΓΕΣ ΣΤΗ
ΒΙΟΙΑΤΡΙΚΗ**

**Κατασκευή Παραμετρικού Μοντέλου Ανθρώπινου
Σώματος για Αναγνώριση Δραστηριότητας σε Video**

Κότταρη Κωνσταντίνα

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

Επιστημονικός Υπεύθυνος:

Δελημπασης Κωνσταντίνος

Επίκουρος Καθηγητής

Κατασκευή Παραμετρικού Μοντέλου Ανθρώπινου Σώματος για Αναγνώριση Δραστηριότητας σε Video

Κότταρη Κωνσταντίνα

Τριμελής Επιτροπή:

Δελήμπασης Κωνσταντίνος – Επίκουρος Καθηγητής

Πλαγιανάκος Βασίλης – Αναπληρωτής Καθηγητής

Σανδαλίδης Χαρίλαος – Επίκουρος Καθηγητής

Περίληψη

Οι εφαρμογές επεξεργασίας και ανάλυσης βίντεο, ιδιαίτερα αυτές που σχετίζονται με τον αυτόματο εντοπισμό της στάσης της ανθρώπινης σιλουέτας και της αναγνώρισης δραστηριότητας, είναι κομμάτι της Τεχνητής Νοημοσύνης. Στα στιγμιότυπα των βίντεο, οι σιλουέτες συχνά υστερούνται πληροφορίας βάθους. Το πρόβλημα αυτό εντοπίζεται περισσότερο στα βίντεο που προκύπτουν από τη λήψη μέσω μίας κάμερα. Στη συγκεκριμένη εργασία, χρησιμοποιείται ένα τρισδιάστατο μοντέλο ανθρώπινου σώματος, σε συνδυασμό με μία βαθμονομημένη κάμερα ευρυγώνιου φακού (fisheye), με σκοπό την απόκτηση ενδείξεων τρισδιάστατου σχήματος. Επεξηγηματικά, παράγεται ένα μοντέλο γενικευμένου τρισδιάστατου ανθρώπου μέσω ενός νεοτεριστικού μαθηματικού φορμαλισμού μίας γνωστής κλάσης επιφανειών, που ονομάζονται γενικευμένοι κύλινδροι. Ο νέος αυτός φορμαλισμός έχει σημαντικά πλεονεκτήματα σε σχέση με τους υπάρχοντες παραμετρικούς ορισμούς. Η χρήση του μοντέλου της fisheye κάμερας επιτρέπει τη δημιουργία εξομοιωμένων εικόνων των σιλουετών χρησιμοποιώντας το παραμετρικό τρισδιάστατο μοντέλο του ανθρώπου. Επιπροσθέτως, παρουσιάζεται ένας αποδοτικός αλγόριθμος βελτιστοποίησης για την ταύτιση του 3D μοντέλου με τη φιγούρα του πραγματικού ανθρώπου, ώστε να αναγνωρίζεται η στάση του παρακολουθούμενου ανθρώπου. Αρχικά, η σιλουέτα τμηματοποιείται για κάθε στιγμιότυπο και κατόπιν, γίνεται υπολογισμός της πραγματικής θέσης του ανθρώπου. Στη συνέχεια, μία διαδικασία οπτικοποίησης ρυθμίζει τις παραμέτρους του 3D μοντέλου στην προσπάθεια της ταυτοποίησής του με την πόζα (θέση και κατεύθυνση, ως προς την κάμερα) του ανθρώπου. Τα πειραματικά αποτελέσματα είναι ελπιδοφόρα, αφού η στάση, η τροχιά και ο προσανατολισμός του ανθρώπου μπορούν να εκτιμηθούν με ακρίβεια. Μεγάλο μέρος των αποτελεσμάτων της παρούσας πτυχιακής έχουν παρουσιαστεί σε διεθνές συνέδριο [41].

Abstract

The role of video processing and analysis applications increases as video acquisition and sharing becomes available. Silhouette identification, pose and activity recognition from video sequences are issues that have not yet been solved. Silhouettes in video frames captured by a single camera lack depth information. In this work, we combine a three-dimensional human body model with a calibrated fish-eye camera, to derive three-dimensional (3D) clues from the imaged scene. First, a generic 3D human model is derived from a novel mathematical formalization of a well-known class of geometric primitives, namely the generalized cylinders (GCs). This 3D model can assume various poses by means of a set of shape parameters. We used a novel parametric definition of GCs, which exhibits advantages over the existing definitions. The use of the calibrated model of the specific fish-eye camera can generate rendered silhouettes, using the aforementioned 3D human model. Moreover, we present an efficient local optimization algorithm for matching the parametric 3D model with a real human figure in order to recognize the posture of a monitored person. In the first step, the silhouette is segmented in each frame using background modelling and the real position of the human is calculated. Subsequently, an optimization process adjusts the parameters of the 3D human model in an attempt to match the position and orientation of human. The quantitative experimental results show that the pose, the trajectory and the orientation of the human can be accurately estimated. The results of this work have been presented in [41].

Λέξεις Κλειδιά - Keywords

Τεχνητή Νοημοσύνη, Fisheye Κάμερα, Επεξεργασία Βίντεο, Τρισδιάστατο Παραμετρικό Μοντέλο Ανθρώπου, Γενικευμένοι Κύλινδροι, Εντοπισμός Σιλουέτας, Αναγνώριση Στάσης, Παρακολούθηση Ανθρώπου.

Artificial Intelligence, Fisheye Camera, Video Processing, Three-dimensional Human Modelling, Generalized Cylinders, Silhouette Detection, Posture Recognition, Human Tracking.

Ευχαριστίες

Θα ήθελα να ευχαριστήσω τη μητέρα μου και τη γιαγιά μου, που σε εκείνες οφείλω τη δυνατότητά μου να σπουδάσω, άρα και να συντάξω τα παρακάτω. Επιπλέον, θέλω να ευχαριστήσω τον καθηγητή μου κ. Δελήμπαση που μου έδωσε την ευκαιρία να ασχοληθώ με αυτήν την εργασία, με καθοδήγησε και με στήριξε σε κάθε βήμα, καθώς και όλους τους καθηγητές μου που μου μεταλαμπάδευσαν πληθώρα εκ των γνώσεών τους.

1. Περιεχόμενα

Περίληψη.....	iv
Abstract	v
Λέξεις Κλειδιά - Keywords.....	vi
Ευχαριστίες.....	vii
2. Εισαγωγή	1
2.1. Γενικές προσεγγίσεις στο πρόβλημα της ανίχνευσης στάσης ανθρώπινης σιλουέτας σε ακολουθίες video	2
2.1.1. Προσέγγιση χωρίς παραμετρικό μοντέλο.....	2
2.1.2. Προσέγγιση βασισμένη σε μοντέλα.....	3
2.2. Η συνεισφορά της παρούσας εργασίας.....	4
3. Μέθοδος.....	5
3.1. Κατασκευή μοντέλου κάμερας fisheye	5
3.1.1. Βαθμονόμηση.....	8
3.1.2. Αντίστροφο μοντέλο της fisheye κάμερας – Πίνακες αναφοράς αζιμούθιου και ανύψωσης	9
3.1.3. Υπολογισμός Θέσης Τμηματοποιημένης Σιλουέτας	10
3.2. Μοντέλο Ανθρώπου	12
3.2.1. Γενικευμένος Κύλινδρος.....	12
3.2.2. Κατασκευή 3D human	15
3.2.3. Ολικοί γεωμετρικοί μετασχηματισμοί του 3D μοντέλου.....	21
3.2.4. Κίνηση του 3D μοντέλου	22
3.2.5. Προσομοίωση της απεικόνισης της ανθρώπινης Σιλουέτας με χρήση του μοντέλου της κάμερας	24
3.3. Τμηματοποίηση.....	27
3.4. Ανίχνευση Ανθρώπινης Σιλουέτας σε βίντεο.....	31
3.4.1. Αντικειμενική συνάρτηση.....	31
3.4.2. Εύρεση στάσης σιλουέτας.....	31
4. Αποτελέσματα.....	34
4.1. Αποτελέσματα στα οποία η ανθρώπινη δραστηριότητα είναι το βάδισμα	34

4.2.	Αποτελέσματα στα οποία ελέγχεται η στάση «όρθιος» - «καθιστός»	38
5.	Συμπεράσματα και Μελλοντική εργασία	44
6.	Αναφορές	46

2. Εισαγωγή

Το επιστημονικό πεδίο του αυτόματου εντοπισμού της στάσης της ανθρώπινης σιλουέτας και της αναγνώρισης της ανθρώπινης δραστηριότητας έχει κερδίσει ιδιαίτερο ενδιαφέρον πρόσφατα. Οι πιθανές εφαρμογές τέτοιων αλγορίθμων περιλαμβάνουν αυτόματη επιτήρηση εσωτερικών ή εξωτερικών χώρων, υποβοήθηση της αυτόνομης διαβίωσης ηλικιωμένων και χρόνια πασχόντων (ambient assisted living environments) κλπ. Η αύξηση της φθηνά διαθέσιμης υπολογιστικής ισχύως, η ευκολία της λήψης, μετάδοσης και αποθήκευσης μεγάλου όγκου ακολουθιών video καθώς και η αυτόματοποιημένη επικοινωνία μεταξύ των καμερών, εξυπηρετητών και λοιπών συσκευών που συμμετέχουν στην διαδικασία επιτήρησης, εντείνουν περαιτέρω την χρησιμότητα σχετικών αλγορίθμων.

Ένας αριθμός προσεγγίσεων έχει αναφερθεί στην βιβλιογραφία που βασίζεται σε 3D μοντέλα του ανθρώπινου σώματος, ή σε περιγραφείς της εικόνας. Αναλυτικά άρθρα επισκόπησης είναι τα ακόλουθα [1], [2], [3], [4]. Μεγάλος αριθμός εργασιών χρησιμοποιεί την προσέγγιση της εξαγωγής περιγραφών από τις σιλουέτες της εικόνας. Στο [3] περιγράφονται 14 εργασίες που χρησιμοποιούν σιλουέτες για να εξάγουν περιγραφείς δραστηριοτήτων. Η χρήση των σιλουετών για αναγνώριση ανθρώπινης δραστηριότητας με χρήση περιγραφών ανασκοπείται στο [2]. Οι παραπάνω αλγόριθμοι παράγουν ικανοποιητικά αποτελέσματα για σύντομα video με σχετικά απλές ανθρώπινες δραστηριότητες, χωρίς πολλά άτομα παρόντα. Σχετικά παραδείγματα αποτελούν το INRIA XMAS [5] το Weizmann σετ δεδομένων [6], το KTH [7] η βάση δεδομένων CMU MoBo [8] και η βάση Human EVA [9]. Σε αυτά τα σετ δεδομένων, συνήθως δεν υπάρχει άλλη ορατή ανθρώπινη δραστηριότητα, ενώ η τμηματοποίηση των σιλουετών είναι σχετικά εύκολη. Για λίγες βάσεις δεδομένων ακολουθιών βίντεο, οι οποίες περιέχουν σκηνές με δύσκολη τμηματοποίηση σιλουετών, πολλαπλές δραστηριότητες και άτομα που αλληλεπιδρούν μεταξύ τους, έχουν δημοσιευθεί ικανοποιητικά αποτελέσματα, όπως το [10].

2.1. Γενικές προσεγγίσεις στο πρόβλημα της ανίχνευσης στάσης ανθρώπινης σιλουέτας σε ακολουθίες video

Μια δημοφιλής προσέγγιση καθορισμού στάσης ανθρώπου σε ακολουθίες video βασίζεται σε 3D μοντέλο του ανθρώπινου σώματος. Στην προσέγγιση αυτή, παράγεται μία προβολή του μοντέλου σε μία πιθανή στάση και υπολογίζεται η συμφωνία της προβολής αυτής με την τρέχουσα εικόνα του video. Τα 3D μοντέλα του ανθρώπινου σώματος είναι συνήθως πολυπαραμετρικά, ώστε να μπορούν να προσομοιώνουν ικανό αριθμό στάσεων του σώματος. Έτσι η διαδικασία του προσδιορισμού των τιμών των παραμέτρων του μοντέλου δύσκολα μπορεί να επιτευχθεί σε πραγματικό χρόνο, όταν δεν λαμβάνονται υποψη πιθανές συσχετίσεις μεταξύ των τιμών τους σε διαδοχικά frames. Ένα σχετικό παράδειγμα αποτελεί το βάδισμα.

Η αντίστροφη προσέγγιση του προβλήματος περιλαμβάνει τον εντοπισμό χαρακτηριστικών μελών του ανθρώπινου σώματος σε frame του video και στη συνέχεια ο συνδυασμός τους δίνει την στάση του σώματος. Παράδειγμα αυτής της προσέγγισης αποτελεί το [11]. Χαρακτηριστικό πρόβλημα της προσέγγισης αυτής είναι ο δύσκολος εντοπισμός των επιμέρους μελών του ανθρώπινου σώματος στο frame, λόγω της ύπαρξης πολλών ψευδών θετικών ανιχνεύσεων [12].

2.1.1. Προσέγγιση χωρίς παραμετρικό μοντέλο

Στην προσέγγιση αυτή γίνεται απευθείας αντιστοίχιση τιμών της εικόνας, ή περιγραφικών αυτής με στάση της σιλουέτας. Οι Rosales and Sclaroff [13] υπολόγισαν τις ροπές H_u από εικόνες και χρησιμοποίησαν ένα νευρωνικό δίκτυο (ΝΔ) για να αντιστοιχίσουν τις τιμές των ροπών σε στάσεις. Η εκπαίδευση του ΝΔ γίνεται με βάση ένα 3D μοντέλο του ανθρώπινου σώματος που χρησιμοποιεί κυλίνδρους. Παρόμοια είναι και η προσέγγιση [14], η οποία χρησιμοποιεί HMM αντί του ΝΔ.

Σιλουέτες και περιγράμματα: Οι σιλουέτες μπορούν να χρησιμοποιηθούν για αναγνώριση στάσης, καθώς δεν εξαρτώνται από χρωματικά χαρακτηριστικά, με την προϋπόθεση ότι η τμηματοποίηση είναι επιτυχής. Οι κεντρικές ροπές [15], ή οι ροπές H_u [16] μπορούν να χρησιμοποιηθούν για την αναγνώριση της στάσης. Τα περιγράμματα των σιλουετών μπορούν επίσης να χρησιμοποιηθούν για τον ίδιο σκοπό, με χρήση διαφόρων μετρικών απόστασης μεταξύ διαφορετικών στάσεων, όπως μετασχηματισμοί

απόστασης (Chamfer distance) [17] ή κόστος ελαστικής παραμόρφωσης.

2.1.2. Προσέγγιση βασισμένη σε μοντέλα

Τα 3D μοντέλα αποτελούνται από πρωτογενή γεωμετρικά αντικείμενα, είτε από τριγωνοποιημένες επιφάνειες. Συχνά χρησιμοποιούμενα πρωτογενή γεωμετρικά αντικείμενα είναι σφαίρες [18], κύλινδροι ή μετασχηματισμένα υπερελιψοειδή [19]. Εκτενής επισκόπηση του θέματος γίνεται στο [3].

Οι κύλινδροι αποτελούν εύχρηστα πρωτογενή γεωμετρικά αντικείμενα (geometric primitives) τα οποία μπορούν να συνδυαστούν εύκολα για να παραστήσουν ένα μοντέλο του ανθρώπινου σώματος [19].

Σε μία πιο πρόσφατη εργασία [20] ένα 3D μοντέλο αποτελούμενο από κυλίνδρους χρησιμοποιήθηκε σε συνδυασμό με μία νέα τεχνική βελτιστοποίησης για να ανακτήσει την στάση της σιλουέττας από 3D ακολουθίες video. Στη συνέχεια η ανθρώπινη δραστηριότητα αναγνωρίστηκε με χρήση Hidden Markov Model (HMM).

Μοντέλο αποτελούμενο από κυλίνδρους έχει χρησιμοποιηθεί και για ακολουθίες video σε συνδυασμό με κάμερα βάθους [21].

2.2. Η συνεισφορά της παρούσας εργασίας

Η συνεισφορά της παρούσας εργασίας μπορεί να συνοψισθεί ως εξής:

- Χρησιμοποίηση ενός νέου ορισμού της επιφάνειας των γενικευμένων κυλίνδρων (Generalized Cylinders –GCs) για την κατασκευή ενός παραμετρικού 3D μοντέλου του ανθρώπινου σώματος, συγκριτικά με μοντέλα αποτελούμενα από πλούστερα γεωμετρικά αντικείμενα
- Ενσωμάτωση επιπλέον βαθμών ελευθερίας στο συγκεκριμένο προτεινόμενο 3D μοντέλο, ώστε να προσομοιώνονται οι φάσεις του βαδίσματος και οι στάσεις «όρθιος», «καθιστός»
- Εφαρμογή αλγορίθμων ανάλυσης βίντεο σε ακολουθίες που συλλέγονται από ευρυγώνιο (fisheye) φακό κάμερας τοποθετημένης σε οροφή κτιρίου καθημερινής δραστηριότητας. Ας τονιστεί ότι η αυτόματη παρακολούθηση της τροχιάς μίας σιλουέτας και η αναγνώριση της στάσης της, αποτελούν ιδιαίτερα δυσεπίλυτα προβλήματα για τέτοιου είδους ακολουθίες βίντεο, (συγκριτικά με ακολουθίες που λαμβάνονται από απλή κάμερα τοποθετημένη σε κατακόρυφο τοίχο)
- Αξιοποίηση ενός νέου μοντέλου βαθμονόμησης της fisheye κάμερας για την παραγωγή εξομοιωμένων στιγμιότυπων ανθρώπινων στάσεων
- Ορισμός κατάλληλης αντικειμενικής συνάρτησης και αποτελεσματική μέθοδος τοπικής βελτιστοποίησης της για την εύρεση των τιμών των παραμέτρων του μοντέλου και τον καθορισμό θέσης, τροχιάς, προσανατολισμού και στάσης της σιλουέτας

Μεγάλο μέρος των αποτελεσμάτων της παρούσας πτυχιακής έχουν παρουσιαστεί σε διεθνές συνέδριο [41].

3. Μέθοδος

Η μεθοδολογία μας αποτελείται από τέσσερα επιμέρους στάδια. Στο πρώτο, κατασκευάζουμε και βαθμονομούμε το μοντέλο της κάμερας. Στο δεύτερο χρησιμοποιούμε το μοντέλο του γενικευμένου κυλίνδρου για να προσομοιώσουμε το ανθρώπινο σώμα και με μαθηματικές μετατροπές σε αυτόν, να του προσδίδουμε κίνηση. Σε επόμενο στάδιο, βιντεοσκοπούμε φυσιολογική ανθρώπινη κίνηση και τμηματοποιούμε τα βίντεο. Τέλος, χρησιμοποιούμε το ανθρώπινο μοντέλο μας για να ανιχνεύσουμε την πραγματική κίνηση στο τμηματοποιημένο βίντεο.

3.1. Κατασκευή μοντέλου κάμερας fisheye

Έχουμε χρησιμοποιήσει τη μοντελοποίηση παραγωγής εικόνας με χρήση ευρυγώνιας κάμερας όπως περιγράφεται στο [22]. Το μοντέλο μας αποτελείται από έναν ημισφαιρικό φακό, και από τον αισθητήρα (CCD) , το οποίο καθορίζεται από ένα επίπεδο πίσω από τον ημισφαιρικό φακό, όπως φαίνεται στο σχήμα 1. Στο paper αυτό υιοθετείται η γενική περίπτωση κατά την οποία ο αισθητήρας μπορεί να μην είναι κάθετος.

Η γενική μορφή του μοντέλου fisheye M γράφεται ως εξής :

$$(j,i) = M(x,y,z) \quad (1)$$

όπου (j,i) είναι οι άξονες των pixel στο στιγμιότυπο του βίντεο, και (x,y,z) είναι οι άξονες του πραγματικού χώρου που αντιστοιχούν στο απεικονιζόμενο σημείο. Δοθέντος ενός pixel (j,i) ενός στιγμιότυπου του βίντεο, βρίσκουμε την οπτική κατεύθυνση ορισμένη σε σφαιρικές συντεταγμένες μέσω δύο γωνιών: το αζιμούθιο θ και η γωνία ανύψωσης φ (elevation) (σχήμα 1):

$$(\theta, \varphi) = M_1(i, j). \quad (2)$$

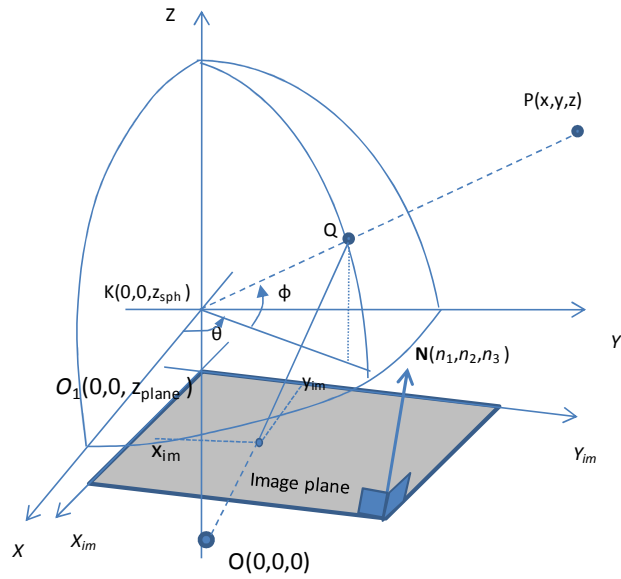
Η περιγραφή του μοντέλου για τη fisheye κάμερα βασίζεται στη φυσική του σχηματισμού εικόνας, όπως περιγράφεται στα [23],[24] και παρουσιάζεται στο [25]. Το μοντέλο αποτελείται από :

- ένα σφαιρικό οπτικό στοιχείο (φακός) με οποιαδήποτε ακτίνα R_0 , με κέντρο το $K(0, 0, z_{sph})$ (σχήμα 2) και

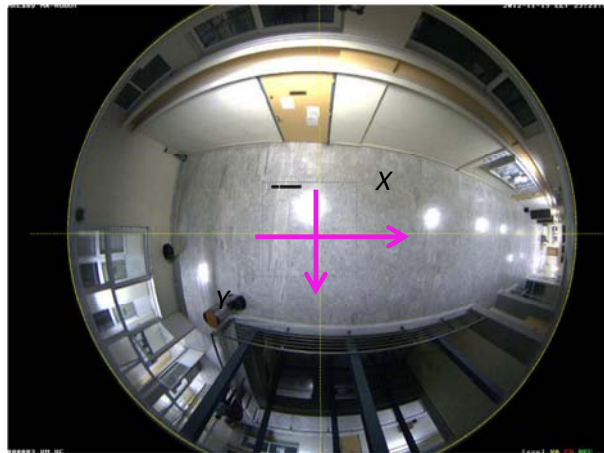
- το επίπεδο του ανιχνευτή CMOS το οποίο διέρχεται από το σημείο $(0, 0, z_{plane})$ και καθορίζεται από το μοναδιαίο κάθετο διάνυσμα \mathbf{n} .

Ορίζουμε την τομή Q της γραμμής KP με το σφαιρικό οπτικό στοιχείο του fisheye φακού για οποιοδήποτε σημείο P με συντεταγμένες πραγματικού χώρου (x,y,z) . Το σημείο P οπτικοποιείται στην κεντρική προβολή (x_{im}, y_{im}) του σημείου Q στο επίπεδο της εικόνας, χρησιμοποιώντας το $O(0,0,0)$ ως κέντρο της προβολής. Η προσέγγιση αυτή λειτουργεί ακόμα και όταν το επίπεδο δεν είναι οριζόντιο και ο άξονας του σφαιρικού φακού δεν είναι ευθυγραμμισμένος με τον κάθετο άξονα του ανιχνευτή της εικόνας. Γίνεται πλέον προφανές ότι όλα τα σημεία του πραγματικού χώρου, τα οποία ανήκουν στην ημιευθεία KP απεικονίζονται στο ίδιο σημείο (x_{im}, y_{im}) στο επίπεδο της εικόνας. Η ημιευθεία KP είναι μοναδικά ορισμένο από το αζιμούθιο και τη γωνία ανύψωσης, ϑ και φ , αντίστοιχα. Η σχεδίαση της γεωμετρίας του fisheye μοντέλου φαίνεται στο σχήμα 1.

Η κάμερα fisheye δεν έχει κινητά μέρη. Συνεπώς, η σχέση μεταξύ z_{sph} και z_{plane} καθορίζει το σχηματισμό της εικόνας. Θέτουμε το z_{plane} σε μία οποιαδήποτε τιμή, μικρότερη του R_0 και ορίζουμε $z_{sph} = pz_{plane}$, όπου p είναι η κύρια παράμετρος του fisheye μοντέλου. Για να μπορούμε να χειριστούμε πιθανή μη ευθυγράμμιση του αισθητήρα της κάμερας σε σχέση με το φακό, η οποία θα μπορούσε να έχει ως αποτέλεσμα παραμορφώσεις της εικόνας [26], εισάγουμε δύο επιπλέον παραμέτρους στο μοντέλο της κάμερας: την X και την Y θέση του κέντρου του σφαιρικού οπτικού φακού $K(x_{sph}, y_{sph}, z_{sph})$ σχετικά με τον οπτικό άξονα της κάμερας. Έτσι, πλέον οι παράμετροι της κάμερας είναι οι εξής: \mathbf{n} , p, x_{sph} , and y_{sph} . Το σχήμα 1 δείχνει τη γεωμετρία του μοντέλου της fisheye κάμερας για $x_{sph} = 0$ και $y_{sph} = 0$.



Σχήμα 1. Η γεωμετρία του προτεινόμενου fisheye μοντέλου κάμερας.



Σχήμα 2. Ένα τυπικό στιγμιότυπο της fisheye κάμερας. Η ακτίνα και το κέντρο του οπτικού πεδίου (center of distortion – CoD) εντοπίζονται μέσω των κίτρινων γραμμών.

3.1.1. Βαθμονόμηση

Με στόχο την εφαρμογή του προτεινόμενου μοντέλου της fisheye κάμερας στον αλγόριθμό μας, θα πρέπει να ορίσουμε τις τιμές των άγνωστων παραμέτρων $(p, x_{sph}, y_{sph}, \mathbf{n})$. Αυτό πραγματοποιείται μέσω της βαθμονόμησης. Αρχικά, παρέχουμε τη θέση των N_p οδηγών σημείων $\{(X_{im}^k, Y_{im}^k)\}, k=1,2,\dots,N_p$ σε ένα στιγμιότυπο του βίντεο, καθώς και τις πραγματικές του συντεταγμένες $\{(x_{real}^k, y_{real}^k, z_{real}^k)\}$, (οι εκθέτες δε δηλώνουν δυνάμεις). Σύμφωνα με την εξίσωση (1), και δεδομένων των τιμών των παραμέτρων του μοντέλου, οι αναμενόμενες θέσεις των οδηγών σημείων σε ένα στιγμιότυπο του βίντεο δίνονται από την παρακάτω σχέση.

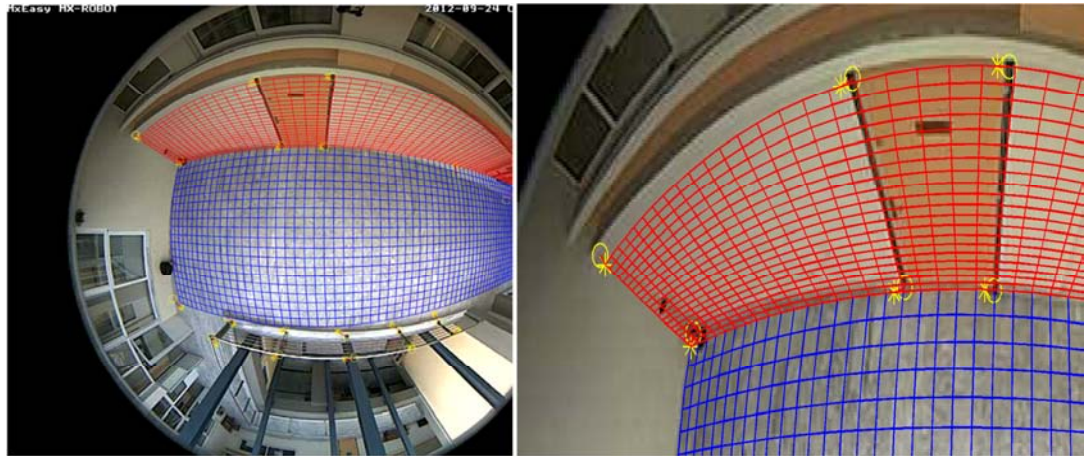
$$(x_{im}^k, y_{im}^k) = M(x_{real}^k, y_{real}^k, z_{real}^k; p, x_{sph}, y_{sph}, \mathbf{n}). \quad (3)$$

Οι τιμές των παραμέτρων αποκτώνται μέσω της ελαχιστοποίησης του λάθους μεταξύ των αναμενόμενων και των παρατηρηθέντων συντεταγμένων των εικονοστοιχείων στα οποία απεικονίζονται τα οδηγά σημεία:

$$(p, x_{sph}, y_{sph}, \mathbf{n}) = \arg \min_{p, x_{sph}, y_{sph}, \mathbf{n}} \left(\sum_{k=1}^{N_p} \left((X_{im}^k - x_{im}^k)^2 + (Y_{im}^k - y_{im}^k)^2 \right) \right). \quad (4)$$

Λόγω του μεγάλου αριθμού των παραμέτρων και της πολυπλοκότητας της αντικειμενικής συνάρτησης, αξιοποιήσαμε την μέθοδο βελτιστοποίησης που βασίζεται σε διαφορική εξέλιξη, όπως περιγράφεται στο [27]. Οι επιτρεπόμενες ακτίνες τιμών των παραμέτρων του μοντέλου καθορίστηκαν ως εξής: η παράμετρος p στο διάστημα $[0.5, 1.5]$, οι x_{sph}, y_{sph} στο διάστημα $[-R_0/4, R_0/4]$ και κάθε συνιστώσα του \mathbf{n} στο διάστημα $[0, 1]$.

Το αποτέλεσμα της παραπάνω διαδικασίας παρουσιάζεται στο παρακάτω σχήμα (σχήμα 3).



(a)

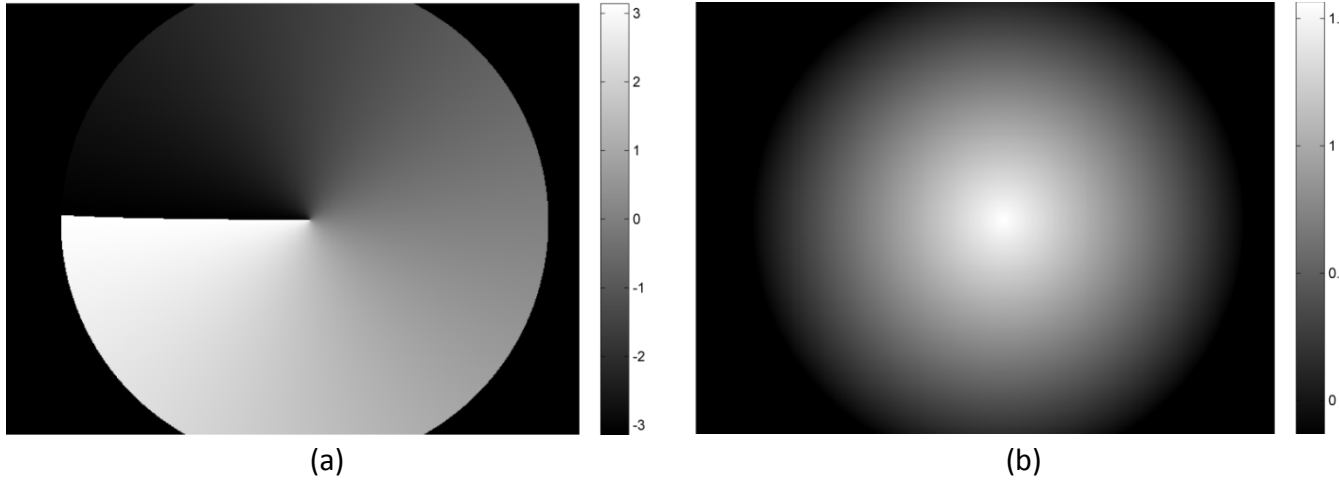
(b)

Σχήμα 3. Οπτικοποίηση του αποτελέσματος της βαθμονόμησης του fisheye μοντέλου (a), με μεγαλύτερη λεπτομέρεια (b). Τα οδηγία σημεία που καθορίζονται από τον χρήστη φαίνονται με κίτρινους κύκλους, ενώ οι υπολογισμένες θέσεις βάσει του μοντέλου απεικονίζονται με κίτρινους αστερίσκους.

3.1.2. Αντίστροφο μοντέλο της fisheye κάμερας – Πίνακες αναφοράς αζιμούθιου και ανύψωσης

Όπως με κάθε προβολικό μετασχηματισμό, για οποιοδήποτε εικονοστοιχείο ενός στιγμιότυπου, το μοντέλο της κάμερας μπορεί να επιστρέψει την ευθεία του χώρου, η οποία καθορίζεται από το αζιμούθιο ϑ και ανύψωσης φ και προβάλλεται στο εικονοστοιχείο αυτό, και όχι ένα συγκεκριμένο σημείο του χώρου. Το αντίστροφο μοντέλο αντιστοιχίζει σε κάθε εικονοστοιχείο (j,i) τις γωνίες ϑ, φ (σχήμα 1), της ευθείας που απεικονίζεται σε αυτό. Εφαρμόζοντας τη μέθοδο, η οποία περιγράφεται στο [22, παράγραφος 2.3], αποθηκεύουμε τις γωνίες ϑ και φ στους πίνακες αναφοράς Θ, Φ . Οι πίνακες αυτοί φαίνονται στο σχήμα 4.

Η παραπάνω διαδικασία εκτελείται μόνο μία φορά μετά τη βαθμονόμηση του μοντέλου της fisheye κάμερας. Όπως αναμένεται, η ακτίνα τιμών του Θ είναι $[-\pi, \pi]$, και του Φ είναι $[0, \pi/2]$, με τη μέγιστη τιμή του δεύτερου να εμφανίζεται στο εικονοστοιχείο του κέντρου του οπτικού πεδίου (center of distortion – CoD).



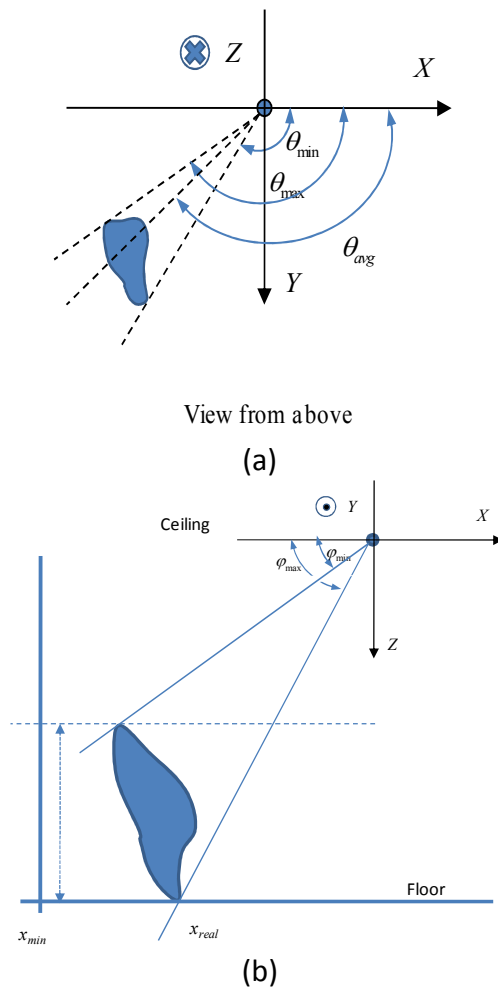
Σχήμα 4. Γραφική αναπαράσταση των πινάκων αναφοράς Θ (a), Φ (b) του αζιμούθιου και της ανύψωσης, αντίστοιχα.

3.1.3. Υπολογισμός Θέσης Τμηματοποιημένης Σιλουέτας

Ας υποθέσουμε ότι ένα αντικείμενο του πραγματικού κόσμου που παρατηρείται από άνω οπτική γωνία, όπως φαίνεται στο σχήμα 5(a), τμηματοποιείται σε ένα στιγμιότυπο του βίντεο. Χρησιμοποιώντας τους πίνακες αναφοράς Θ και Φ , υπολογίζεται η ελάχιστη, μέγιστη και μέση αζιμουθιακή γωνία και γωνία ανύψωσης, τα οποία φαίνονται στο σχήμα 5(b).

Η πραγματική θέση (x_{real}, y_{real}) του τμηματοποιημένου αντικειμένου, υποθέτοντας πως πρόκειται για όρθια ανθρώπινη σιλουέτα η οποία πατάει στο πάτωμα, υπολογίζεται μέσω του παρακάτω τύπου, σύμφωνα με [22].

$$\begin{aligned}
 x_{real} &= \frac{z_{max}}{\sin \varphi_{max}} \cos \varphi_{max} \cos \theta_{avg}, \\
 y_{real} &= \frac{z_{max}}{\sin \varphi_{max}} \cos \varphi_{max} \sin \theta_{avg}.
 \end{aligned}
 \tag{5}$$



Σχήμα 5. Απεικόνιση ενός αντικείμενου του πραγματικού κόσμου μέσω της fisheye κάμερας. Στην εικόνα (a) φαίνεται η κάτοψη του αντικείμενου και οι αζιμουθιακές γωνίες (η μέγιστη και η ελάχιστη). Στην εικόνα (b) φαίνονται οι αντίστοιχες γωνίες ανύψωσης.

3.2. Μοντέλο Ανθρώπου

3.2.1. Γενικευμένος Κύλινδρος

Γενικευμένοι Κύλινδροι (ΓΚ) έχουν χρησιμοποιηθεί συχνά σε εφαρμογές γραφικών και ανάλυσης εικόνας. Στην παρούσα εργασία παράγεται η παραμετρική εξίσωση του ΓΚ χρησιμοποιώντας έναν νέο ορισμό του τοπικού συστήματος συντεταγμένων [28]. Ορίζουμε C_1 ως μία τμηματικά λεία καμπύλη στον R^3 , η οποία περιγράφεται ως εξής:

$$r_1(t) = (x(t), y(t), z(t)), t \in [a, b] \subset R \quad (6)$$

σε ένα Καρτεσιανό Επίπεδο Συντεταγμένων OXYZ και C_2 να είναι μία επίπεδη καμπύλη με μία παραμετρική εξίσωση:

$$r_2(u) = (X, Y) = (m_1(u), m_2(u)), \quad u \in J \subseteq R \quad (7)$$

σε ένα ορθογώνιο τοπικό Καρτεσιανό Σύστημα OXY.

Ας θεωρήσουμε τώρα την επιφάνεια S , η οποία δημιουργείται μετακινώντας την καμπύλη C_2 κατά μήκος της C_1 , έτσι ώστε το επίπεδό της C_2 να είναι κάθετο προς το εφαπτόμενο διάνυσμα του C_1 και η αρχή του OXY να ανήκει στην C_1 . Τότε η εξίσωση του διανύσματος της θέσης της επιφάνειας $r(t, u) = (x(t, u), y(t, u), z(t, u))$ για ένα τυχαίο σημείο, δίνεται από:

$$\begin{aligned} x(t, u) &= x(t) + \frac{m_1(u)y'(t)}{P_2(t)} + \frac{m_2(u)x'(t)z'(t)}{P_1(t)P_2(t)} \\ y(t, u) &= y(t) - \frac{m_1(u)x'(t)}{P_2(t)} + \frac{m_2(u)d(t)y'(t)z'(t)}{P_1(t)P_2(t)} \\ z(t, u) &= z(t) - \frac{m_2(u)((x'(t))^2 + (y'(t))^2)}{P_1(t)P_2(t)} \end{aligned} \quad (8)$$

όπου, $P_1(t) = \sqrt{(x'(t))^2 + (y'(t))^2 + (z'(t))^2}$ και $P_2(t) = \sqrt{x'(t)^2 + y'(t)^2}$.

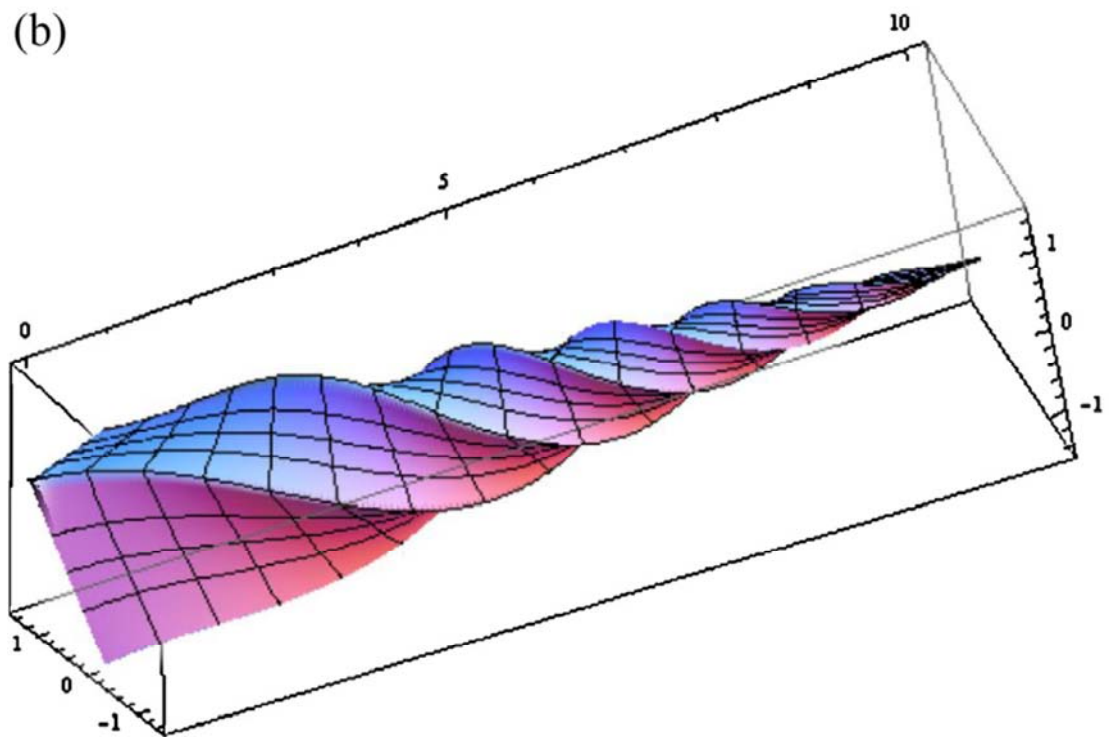
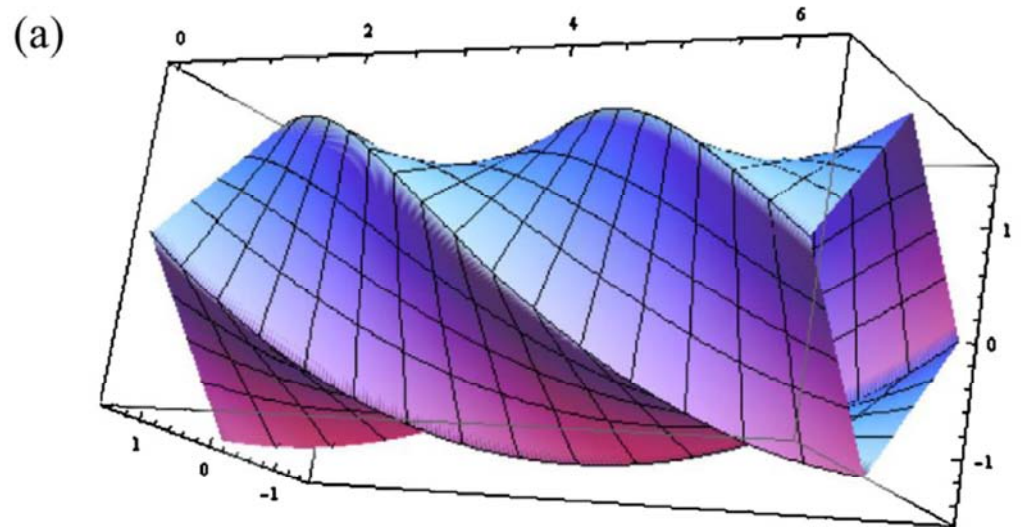
Η πλήρης απόδειξη της (8) βρίσκεται στη δημοσίευση [28].

Αν εκφράσουμε την επίπεδη επιφάνεια C_2 σε πολικές συντεταγμένες $r_2 = r_2(u), u \in [0, 2\pi]$ και εισάγουμε έναν παράγοντα αλλαγής κλίμακας $s(t)$ και μία γωνία στροφής $\phi(t)$ γύρω από το εφαπτόμενο διάνυσμα στη C_1 , συναρτήσει της θέσης κατά μήκος της C_1 , τότε η εξίσωση επιφάνειας (8) του ΓΚ γίνεται:

$$\begin{aligned}
 x(t,u) &= x(t) + \frac{s(t)y'(t)r(u+\phi(t))\cos(u+\phi(t))}{P_2(t)} + \frac{s(t)x'(t)z'(t)r(u+\phi(t))\sin(u+\phi(t))}{P_1(t)P_2(t)} \\
 y(t,u) &= y(t) - \frac{s(t)x'(t)r(u+\phi(t))\cos(u+\phi(t))}{P_2(t)} + \frac{s(t)y'(t)z'(t)r(u+\phi(t))\sin(u+\phi(t))}{P_1(t)P_2(t)} \\
 z(t,u) &= z(t) - \frac{s(t)(x'(t))^2 + (y'(t))^2 r(u+\phi(t))\sin(u+\phi(t))}{P_1(t)P_2(t)}
 \end{aligned} \tag{9}$$

όπου, $(t,u) \in [a,b] \times [0,2\pi]$ και $P_1(t)$ και $P_2(t)$ όπως ορίζονται παραπάνω.

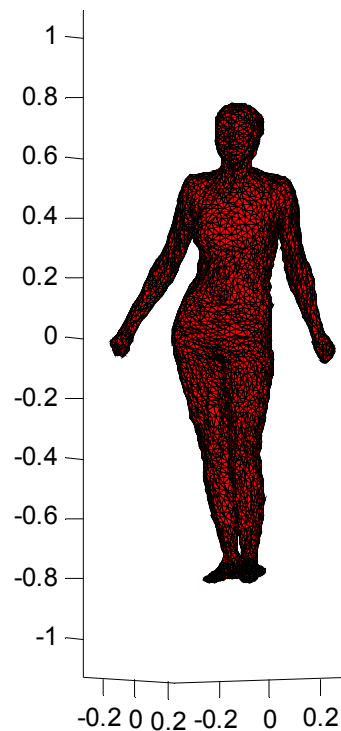
Ο ΓΚ της εξίσωσης (8) είναι μία παραμετρική αναπαράσταση γενικευμένων κυλίνδρων με αυθαίρετα σύνθετη κεντρική καμπύλη C_1 , όπως και κλειστή επίπεδη καμπύλη C_2 . Η μεγάλη ευελιξία του να περιγράφει ποικίλες επιφάνειες φαίνεται στο σχήμα 6.



Σχήμα 6. Δύο παραδείγματα επιφανειών από την εξίσωση (9).

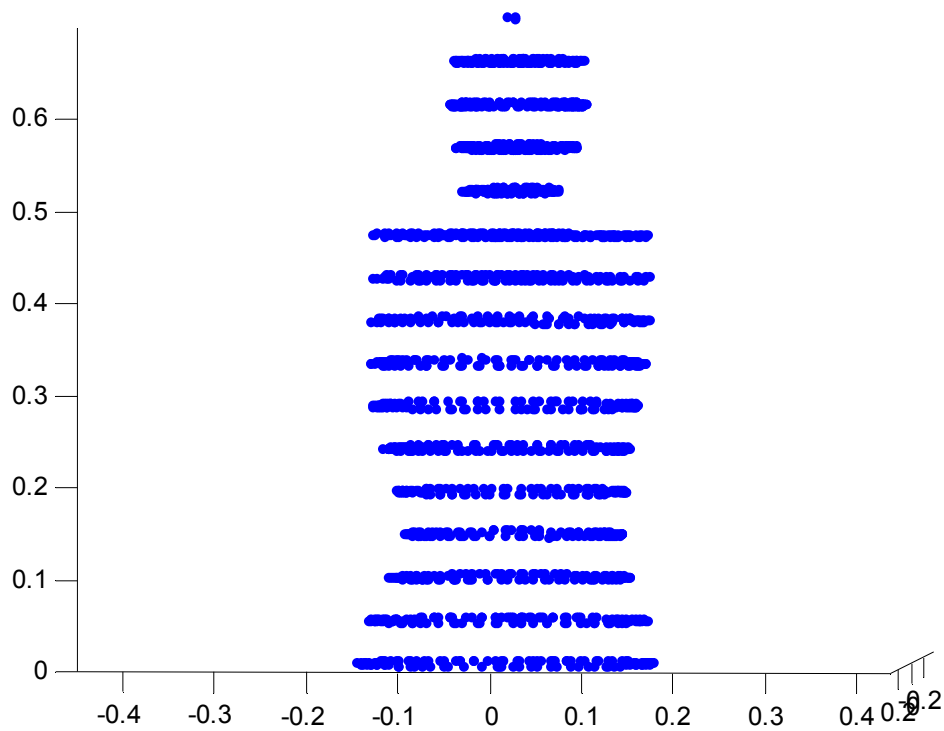
3.2.2. Κατασκευή 3D human

Σε αυτή την εργασία, χρησιμοποιούμε ένα δωρεάν διαθέσιμο τρισδιάστατο μοντέλο (3D) τριγωνοποιημένης επιφάνειας ενός όρθιου ανθρώπου (σχήμα 7). Το μοντέλο αυτό είναι ορισμένο στο Καρτεσιανό Σύστημα Συντεταγμένων και έχει περίπου 27.000 κορυφές [29]. Σκοπός μας είναι να προσομοιώσουμε το σχηματισμό εικόνας της ανθρώπινης Σιλουέτας μέσω της fisheye κάμερας σε πραγματικό χρόνο, για αυτόν τον λόγο παραβλέπουμε την πληροφορία της τριγωνοποιημένης επιφάνειάς του και το χειριζόμαστε ως νέφος σημείων.

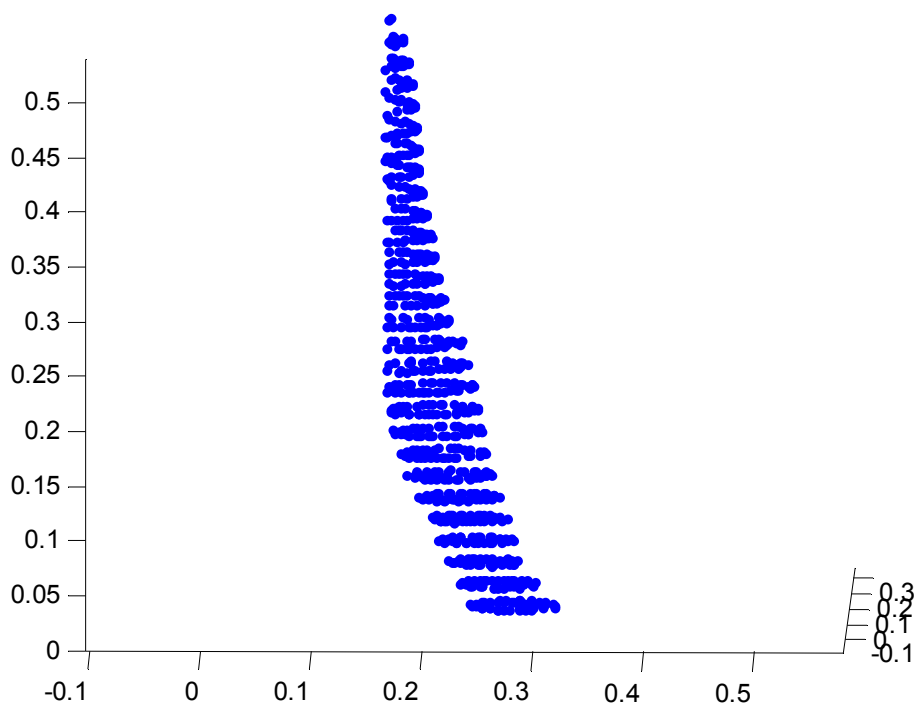


Σχήμα 7. Μοντέλο ανθρώπου.

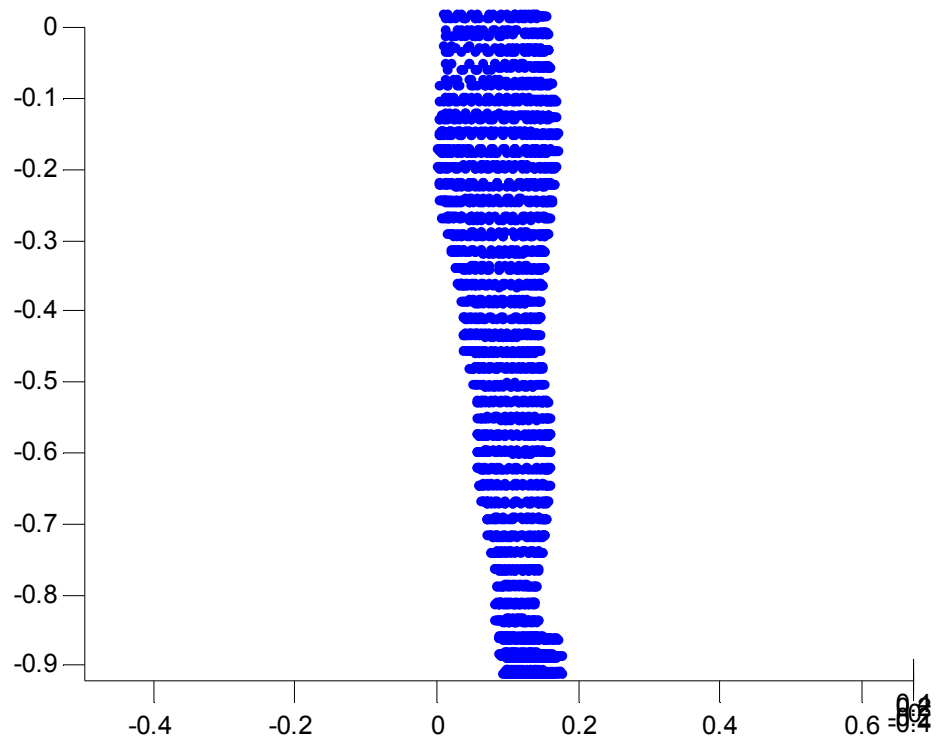
Στη συνέχεια υπολογίζουμε τις τομές του μοντέλου με μία σειρά από οριζόντια επίπεδα, ανά δύο εκατοστά στον άξονα Z κατά μήκος του κορμού (σχήμα 8), των χεριών (σχήμα 9) και των ποδιών (σχήμα 10).



Σχήμα 8. Ελλειπτικές τομές κορμού.

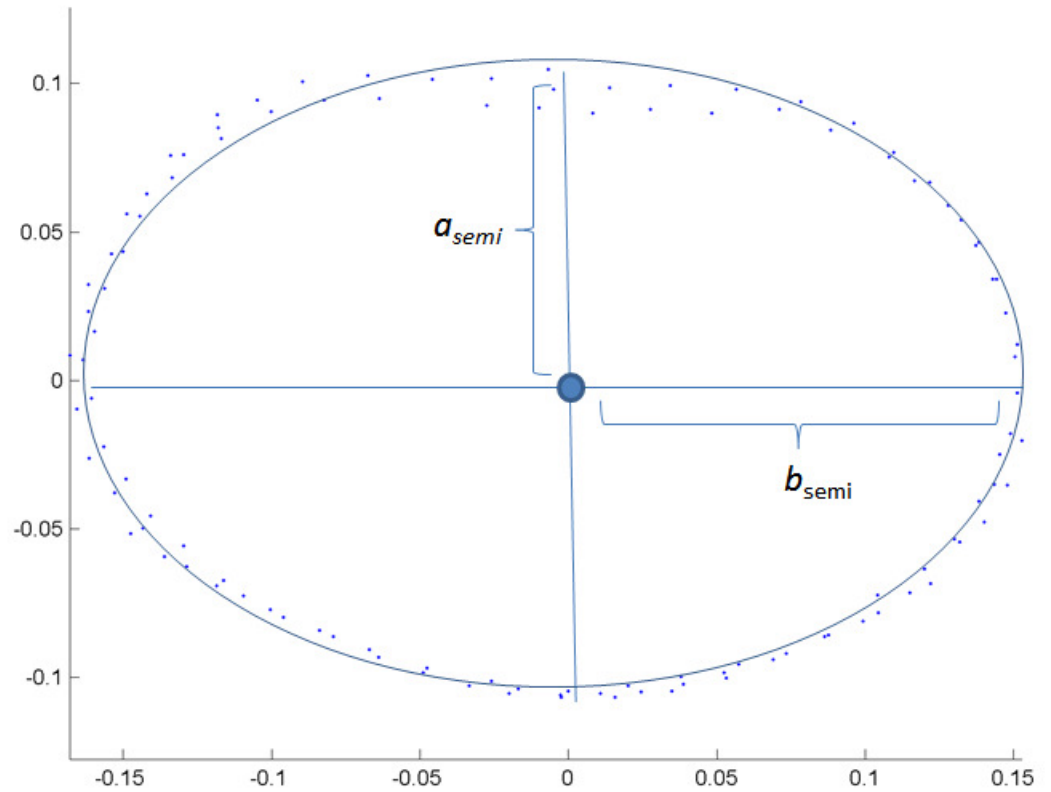


Σχήμα 9. Ελλειπτικές τομές χεριού.



Σχήμα 10. Ελλειπτικές τομές ποδιού.

Κάθε μία από τις τομές αυτές θεωρούμε ότι προσεγγίζεται από μία έλλειψη, οι ημιάξονες a_{semi} , b_{semi} της οποίας είναι παράλληλοι στους άξονες X και Y του συστήματος συντεταγμένων όπως φαίνονται στο σχήμα 11.



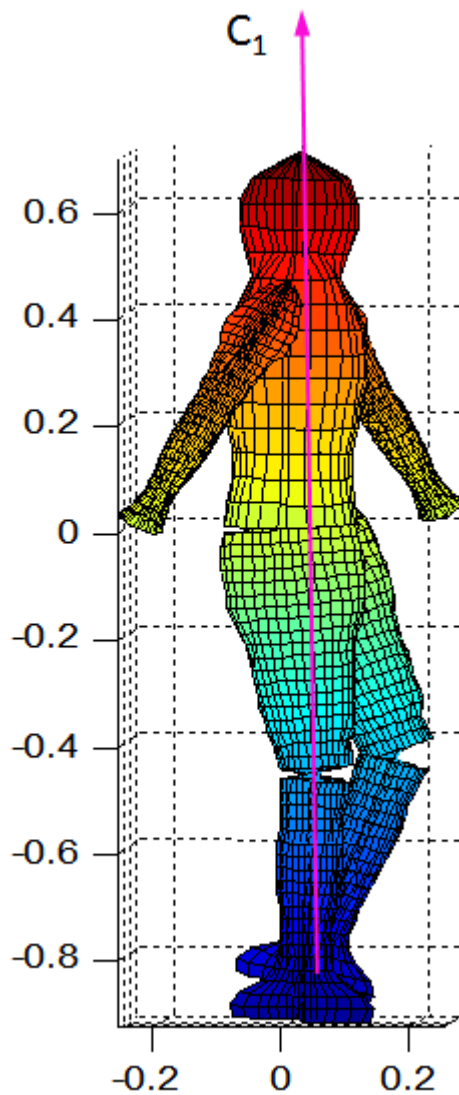
Σχήμα 11. Ελλειπτική τομή κορμού προσδιορισμένη από άξονες.

Η παραμετρική εξίσωση του Γενικευμένου Κυλίνδρου (9), απλοποιείται περαιτέρω αν θέσουμε τις ελλείψεις αυτές ως τις κλειστές καμπύλες C_2 ως εξής :

$$\begin{aligned}
 x &= a_0 t + \frac{b_0 r \cos(u)}{P_2} + \frac{a_0 c_0 r \sin(u)}{P_2} \\
 y &= b_0 t - \frac{a_0 r \cos(u)}{P_2} + \frac{b_0 c_0 r \sin(u)}{P_2} \\
 z &= c_0 t - \frac{(a_0^2 + b_0^2) r \sin(u)}{P_2}
 \end{aligned}
 \tag{10}$$

όπου $r = \frac{a_{semi} b_{semi}}{\sqrt{(b_{semi} \cos(u))^2 + (a_{semi} \sin(u))^2}}$ και τα a_0 , b_0 , c_0 καθορίζουν την κατεύθυνση του κύριου άξονα C_1 .

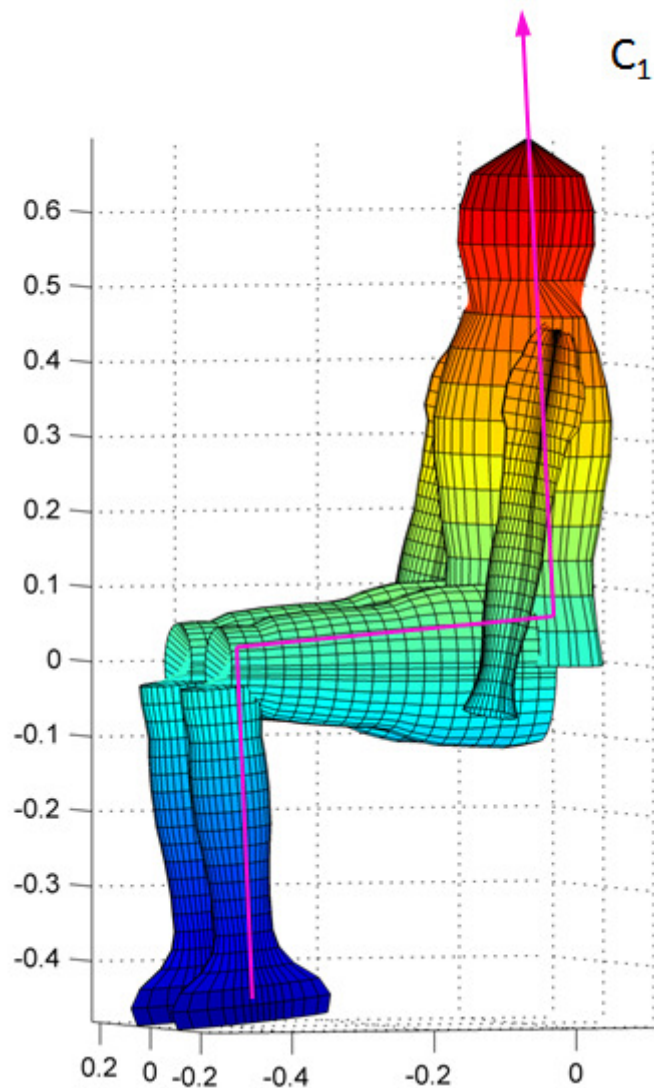
Στο σχήμα 12 βλέπουμε το αποτέλεσμα της μοντελοποίησης των τομών του μοντέλου με ελλείψεις και την εισαγωγή των ελλείψεων στη συνάρτηση του Γενικευμένου Κυλίνδρου (10). Ο κορμός και το κεφάλι έχουν σαν καμπύλη C_1 (σχήμα 12) μία ευθεία παράλληλη με τον άξονα Z στο επιλεγθέν σύστημα συντεταγμένων. Παρατηρούμε, λοιπόν, την ομοιότητα του ΓΚ που παράχθηκε με την ανθρώπινη φιγούρα.



Σχήμα 12. 3D μοντέλο ανθρώπου.

Στο μοντέλο εντοπίζεται μία στροφική αστάθεια στην περιοχή των δύο γονάτων. Το φαινόμενο αυτό έχει εξηγηθεί στο [27, section 4] και δεν επηρεάζει τη λειτουργία του μοντέλου.

Αν αλλάξουμε τις παραμέτρους (βαθμούς ελευθερίας) a_0, b_0, c_0 , της καμπύλης C_1 , τότε γίνεται γραμμική κατά τμήματα. Επιλέγοντας κατάλληλα τους συντελεστές του κάθε τμήματος προσομοιώνουμε την καθιστή στάση του ανθρώπου. Ενδεικτικά, βλ. σχήμα 13.



Σχήμα 13. Καθιστό μοντέλο 3D ανθρώπου.

3.2.3. Ολικοί γεωμετρικοί μετασχηματισμοί του 3D μοντέλου

Το γεωμετρικό μοντέλο που παράγεται, όπως περιγράφηκε στην προηγούμενη παράγραφο, διαθέτει τοπικό σύστημα συντεταγμένων $O_1(x,y,z)$, στο οποίο, το κέντρο βάρους του μοντέλου βρίσκεται στην αρχή των αξόνων $(0,0,0)$, η κατεύθυνση (πόδια – κεφάλι) συμπίπτει με τον άξονα Z, ενώ ο άξονας X ορίζεται από την κατεύθυνση αριστερό – δεξί χέρι.

Το σύστημα συντεταγμένων του πραγματικού κόσμου, όπως αυτό έχει οριστεί στο σχήμα 14 έχει τον άξονα Z να συμπίπτει με την κατακόρυφη κατεύθυνση κάμερα – δάπεδο. Ο X και Y άξονας του συστήματος συντεταγμένων του χώρου φαίνονται στο σχήμα 5 (a και b). Για να συμπίπτει το τοπικό σύστημα συντεταγμένων (του μοντέλου) με το σύστημα συντεταγμένων του πραγματικού κόσμου, απαιτείται ένας απλός μετασχηματισμός του άξονα Z του τοπικού συστήματος του μοντέλου:

$$z' = -z + z_{\max} + \max(z)$$

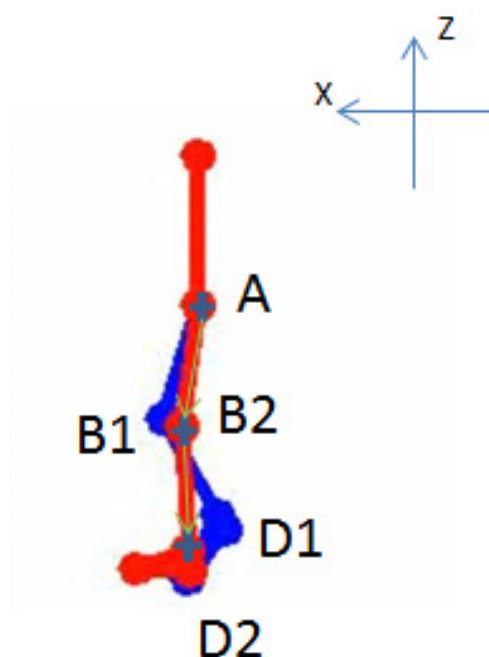
όπου, z_{\max} το ύψος της οροφής (ισοδύναμα $z=z_{\max}$ η εξίσωση του επιπέδου του δαπέδου) και $\max(z)$ η μέγιστη τιμή της Z συντεταγμένης των σημείων του μοντέλου.

Στην εργασία αυτή, θεωρούμε τον ολικό μετασχηματισμό της περιστροφής γύρω από τον άξονα Z, καθώς και μετατοπίσεις κατά τον X και Y άξονα (συνολικά 3 βαθμοί ελευθερίας του ολικού γεωμετρικού μετασχηματισμού). Έτσι, αν $[x, y, z]^T$ είναι σημείο του μοντέλου, όπως παράγεται από την εξίσωση (10), τότε το μετασχηματισμένο σημείο $[x', y', z']^T$ υπολογίζεται ως εξής:

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{pmatrix} 1 & 0 & 0 & dx \\ 0 & 1 & 0 & dy \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos(\theta_{rot}) & -\sin(\theta_{rot}) & 0 & 0 \\ \sin(\theta_{rot}) & \cos(\theta_{rot}) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (11)$$

3.2.4. Κίνηση του 3D μοντέλου

Θέλοντας να προσδώσουμε κίνηση στο μοντέλο, μελετήθηκε το βίντεο ανάλυσης βηματισμού (gait analysis) της διεύθυνσης [30] , του [31]. Συγκεκριμένα, για κάθε μία από τις δέκα φάσεις του περπατήματος λήφθηκαν τα αντίστοιχα στιγμιότυπα από το video. Στα στιγμιότυπα αυτά ορίσαμε το σύστημα συντεταγμένων όπως ακριβώς στο ανθρώπινο μοντέλο μας και επιλέξαμε τρία σημεία σε κάθε πόδι, όπως φαίνεται στο σχήμα 14.



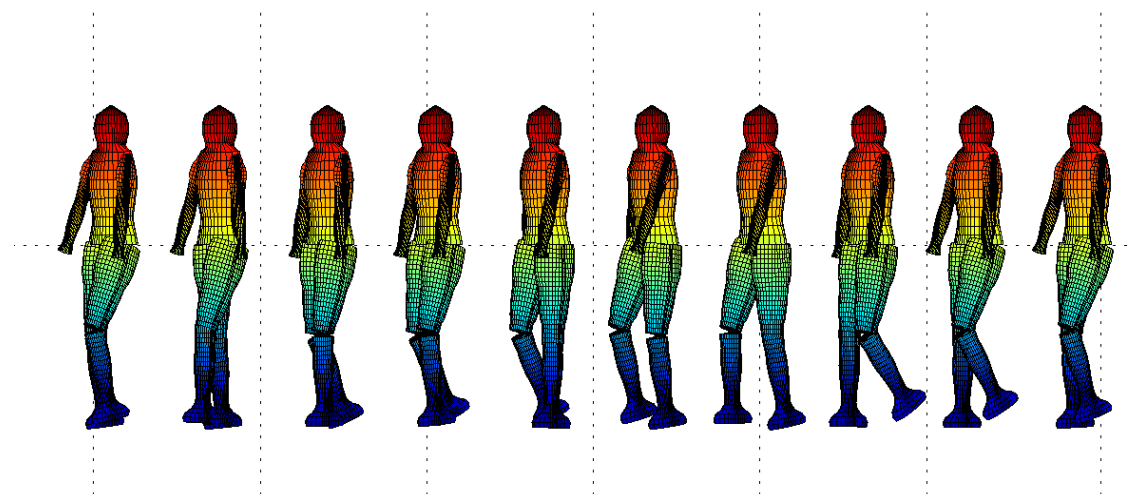
Σχήμα 14. Στιγμιότυπο βίντεο [30] και τα σημεία που επιλέχθηκαν σε κάθε κάτω άκρο.

Τα τρία αυτά σημεία καθορίζουν τα μοναδιαία διανύσματα του μηρού και του γονάτου, τα οποία φαίνονται στον πίνακα 1. Οι συντεταγμένες (a_0 , b_0 , c_0) των διανυσμάτων αυτών χρησιμοποιήθηκαν στη σχέση (10). Έτσι, βρήκαμε την τεθλασμένη που περιγράφει την κίνηση κάθε ποδιού στον βηματισμό. Η συνεχής αυτή τεθλασμένη χρησιμοποιήθηκε ως καμπύλη C_1 .

Πίνακας 1. Τα τρία σημεία των μοναδιαίων διανυσμάτων του μηρού και του γονάτου.

Φάση	Δεξιός Μηρός	Δεξί Γόνατο	Αριστερός Μηρός	Αριστερό Γόνατο
1	(-0.3693, 0, 0.9293)	(0.2102, 0, 0.9776)	(-0.1240, 0, 0.9923)	(0.1215, 0, 0.9926)
2	(-0.3267, 0, 0.9451)	(-0.0471, 0, 0.9989)	(0.0404, 0, 0.9992)	(0.0915, 0, 0.9958)
3	(-0.2066, 0, 0.9784)	(0.0200, 0, 0.9998)	(-0.0624, 0, 0.9981)	(0.3770, 0, 0.9262)
4	(-0.2450, 0, 0.9695)	(0.1083, 0, 0.9941)	(-0.1609, 0, 0.9870)	(0.4559, 0, 0.8900)
5	(-0.0220, 0, 0.9998)	(-0.0225, 0, 0.9997)	(-0.2296, 0, 0.9733)	(0.4084, 0, 0.9128)
6	(-0.0866, 0, 0.9962)	(0.1458, 0, 0.9893)	(-0.3613, 0, 0.9324)	(0.2720, 0, 0.9623)
7	(0.0167, 0, 0.9999)	(0.1871, 0, 0.9823)	(-0.3113, 0, 0.9503)	(0.0303, 0, 0.9995)
8	(-0.1833, 0, 0.9831)	(0.5760, 0, 0.8175)	(-0.1620, 0, 0.9868)	(-0.0161, 0, 0.9999)
9	(-0.3375, 0, 0.9413)	(0.5093, 0, 0.8606)	(-0.0879, 0, 0.9961)	(-0.0154, 0, 0.9999)
10	(-0.3624, 0, 0.9320)	(0.3210, 0, 0.9471)	(-0.0905, 0, 0.9959)	(0.0476, 0, 0.9989)

Οι επιφάνειες που παρήγαγε η εξίσωση 10 για κάθε μία από τις φάσεις του βηματισμού φαίνονται στο σχήμα 15.

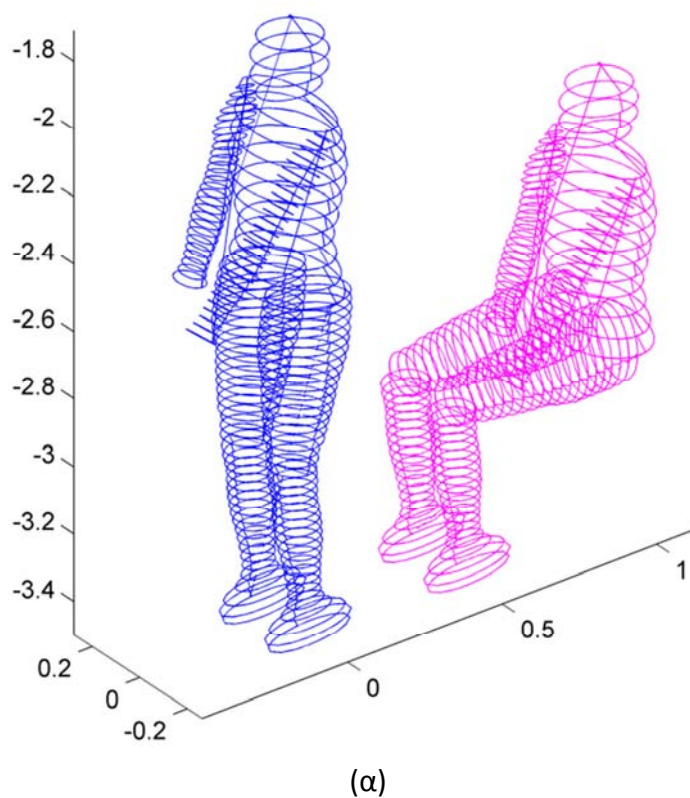


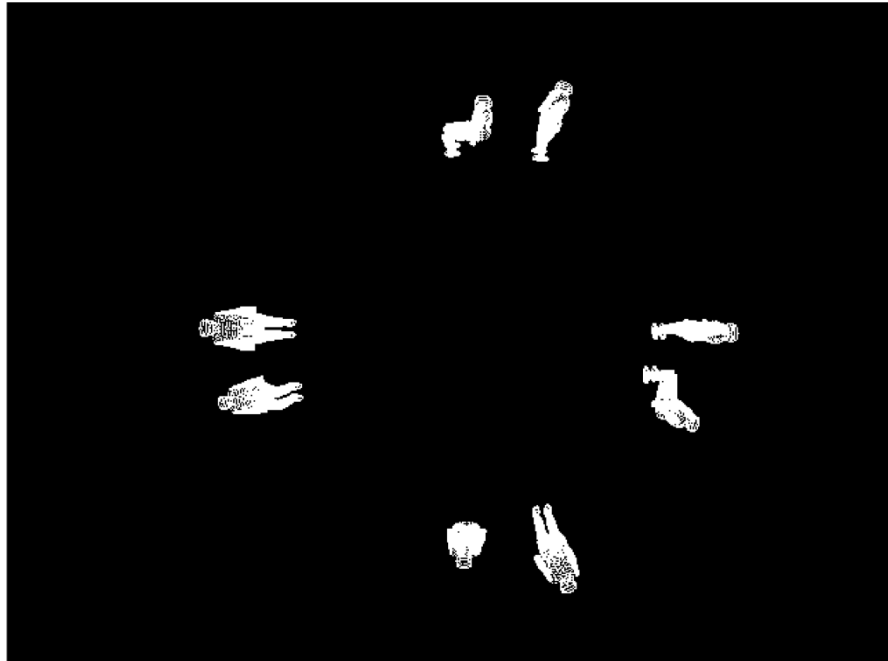
Σχήμα 15. Στιγμιότυπα βηματισμού 3D μοντέλου.

3.2.5. Προσομοίωση της απεικόνισης της ανθρώπινης Σιλουέτας με χρήση του μοντέλου της κάμερας

Για κάθε σημείο του μοντέλου (x_k, y_k, z_k) υπολογίζεται το pixel $(i_k, j_k) = M(x_k, y_k, z_k)$ στο οποίο γίνεται η απεικόνιση, βάσει της εξίσωσης (1). Στη συνέχεια δημιουργείται μία εικόνα I_M με αριθμό γραμμών και στηλών ίσο με αυτά του πραγματικού βίντεο, αρχικοποιημένη με μηδενική τιμή. Τα pixel στα οποία απεικονίζονται τα σημεία του μοντέλου αποκτούν τιμή 1: $I_M(i_k, j_k) = 1$

Στο σχήμα 16 (α) φαίνεται το 3D μοντέλο του όρθιου και καθιστού ανθρώπου. Η απεικόνιση του μέσω του βαθμονομημένου μοντέλου της fisheye κάμερας (παράγραφος 4.1.1) παρουσιάζεται στο (β).

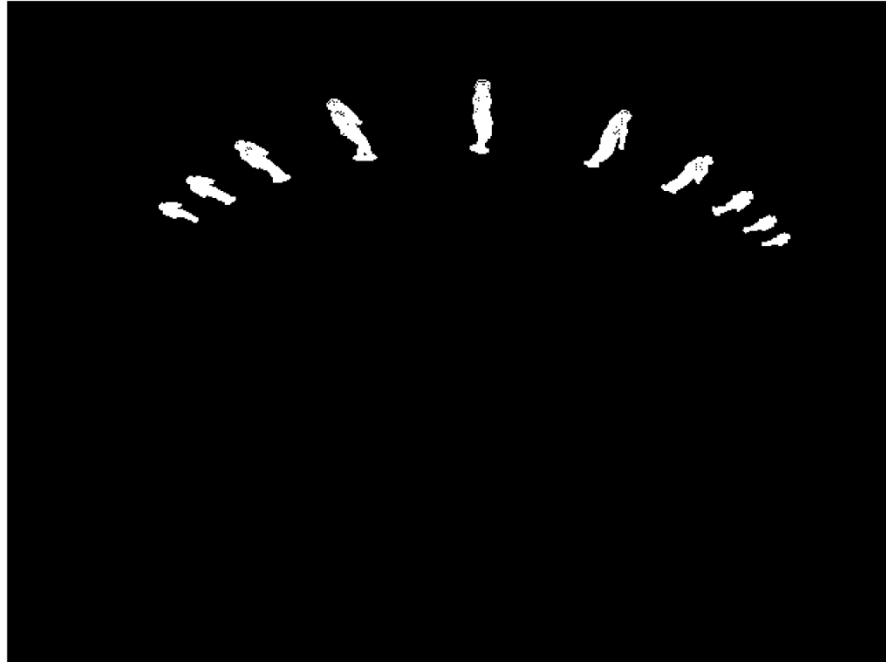




(β)

Σχήμα 16. 3D μοντέλο του όρθιου και καθιστού ανθρώπου (α) και η απεικόνιση του μέσω του βαθμονομημένου μοντέλου της fisheye κάμερας (β).

Ομοίως, στο σχήμα 15 φαίνεται το 3D μοντέλο για κάθε στιγμιότυπο του βηματισμού, ενώ στο σχήμα 17 παρουσιάζεται η απεικόνιση κάθε ενός στιγμιότυπου από το βαθμονομημένο μοντέλο της fish-eye κάμερας.



Σχήμα 17. Η απεικόνιση κάθε ενός στιγμιότυπου του βαδίσματος από το βαθμονομημένο μοντέλο της fish-eye καμερας.

3.3. Τμηματοποίηση

Στη βιβλιογραφία υπάρχουν πολλών ειδών μεθοδολογίες για μοντελοποίηση φόντου. Για παράδειγμα, στο [32] το μοντέλο του φόντου προσδιορίζεται απλά από τα προηγούμενα στιγμιότυπα και εφαρμόζεται καθολικό κατώφλι για την εξαγωγή του προσκηνίου. Αυτή η μέθοδος είναι απλή στη χρήση, αλλά και επιρρεπής σε λάθη. Το φόντο μπορεί να μοντελοποιηθεί από την εφαρμογή του φίλτρο ενδιάμεσης τιμής [33] σε έναν προκαθορισμένο αριθμό στιγμιότυπων, τα οποία κρατιούνται σε κάποια μνήμη. Αυτή η προσέγγιση απαιτεί ένα σημαντικό μέγεθος μνήμης και αριθμό υπολογισμών και δε μπορεί να εκτελεστεί σε πραγματικό χρόνο. Στη βιβλιογραφία, προτίνεται μια κλάση αλγορίθμων μοντελοποίησης αναδρομικών φόντων, με στόχο τη μείωση των υπολογιστικών απαιτήσεων. Οι παραπάνω αλγόριθμοι χρησιμοποιούν μία σταδιακή επικαιροποίηση του φόντου, όπως το φίλτρο ενδιάμεσης τιμής [34] και κινητός σταθμισμένος μέσος όρος με Γκαουσιανά βάρη [35]. Μία κατανοητή ανασκόπηση των αλγορίθμων της μοντελοποίησης του φόντου για ανίχνευση του προσκηνίου παρουσιάζεται στο [36]. Μία δημοφιλής μεθοδολογία είναι των Γκαουσιανών μιγμάτων (mixture of Gaussians – MoG), η οποία εισάγεται στο [37], όπου το χρονικό ιστόγραμμα κάθε εικονοστοιχείο μοντελοποιείται ως γραμμικός συνδιασμός Γκαουσιανών συναρτήσεων πυκνότητας πιθανότητας. Ωστόσο, η μέθοδος αυτή είναι υπολογιστικά ακριβή. Σε αυτή την εργασία εφαρμόζουμε τμηματοποίηση βίντεο στην οποία χρησιμοποιούμε τον αλγόριθμο τμηματοποίησης με διπλό υπόβαθρο (illumination-sensitive – IS), όπως είχε αρχικά περιγραφεί στο [38] και τροποποιηθεί στο [39]. Θα μπορούσε να χρησιμοποιηθεί οποιοσδήποτε αλγόριθμος τμηματοποίησης, όμως ο επιλεγθής ήταν αυτός λόγω της απλότητας του και τον αποτελεσματικό του χειρισμό στις αλλαγές του φωτισμού. Παράδειγμα της χρήσης αυτής σε βίντεο που λήφθηκε με τη fisheye κάμερα, φαίνεται στο σχήμα 18 .

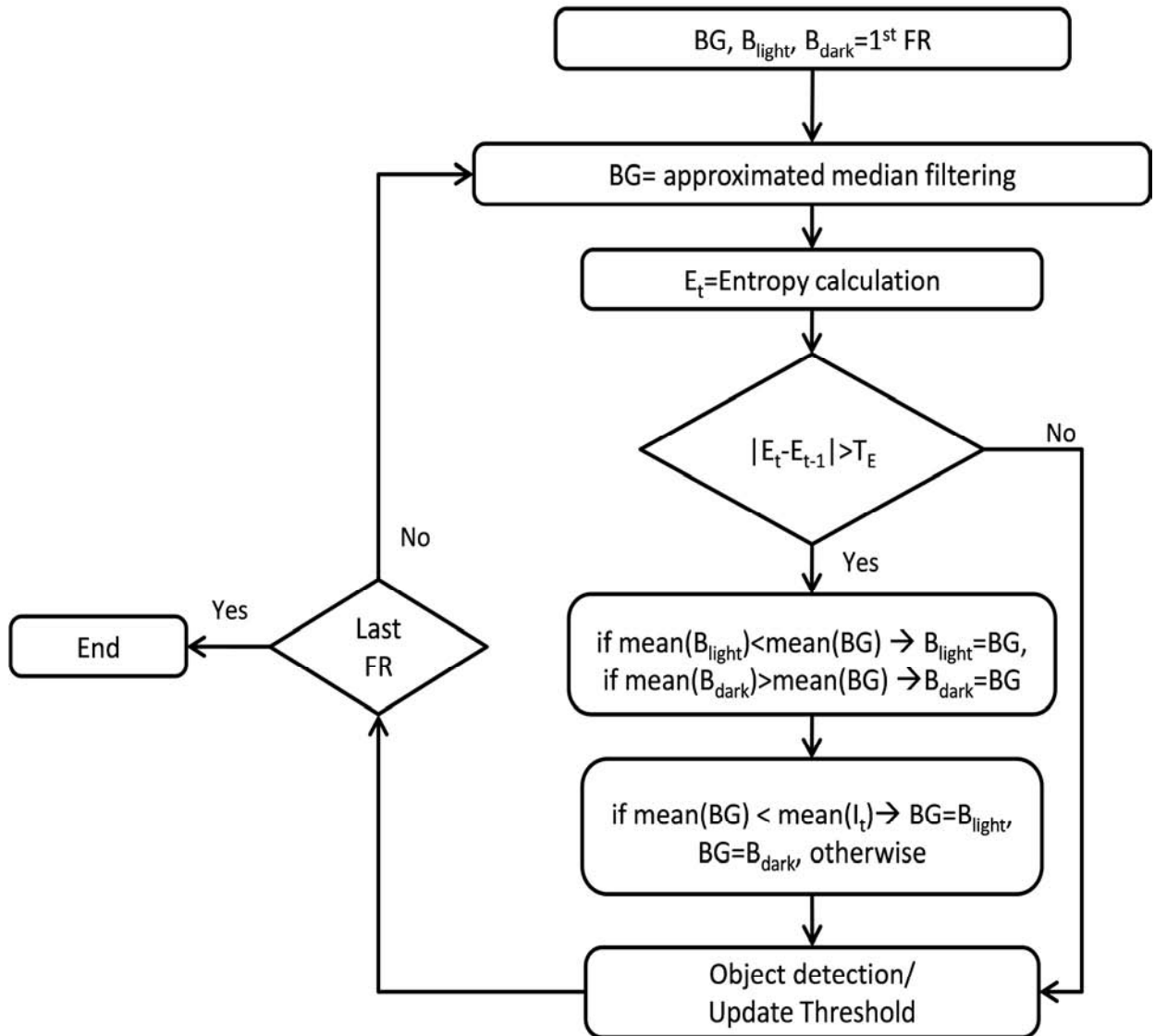


Σχήμα 18. Παράδειγμα εφαρμογής του αλγορίθμου της τμηματοποίησης σε ένα στιγμιότυπο του βίντεο.

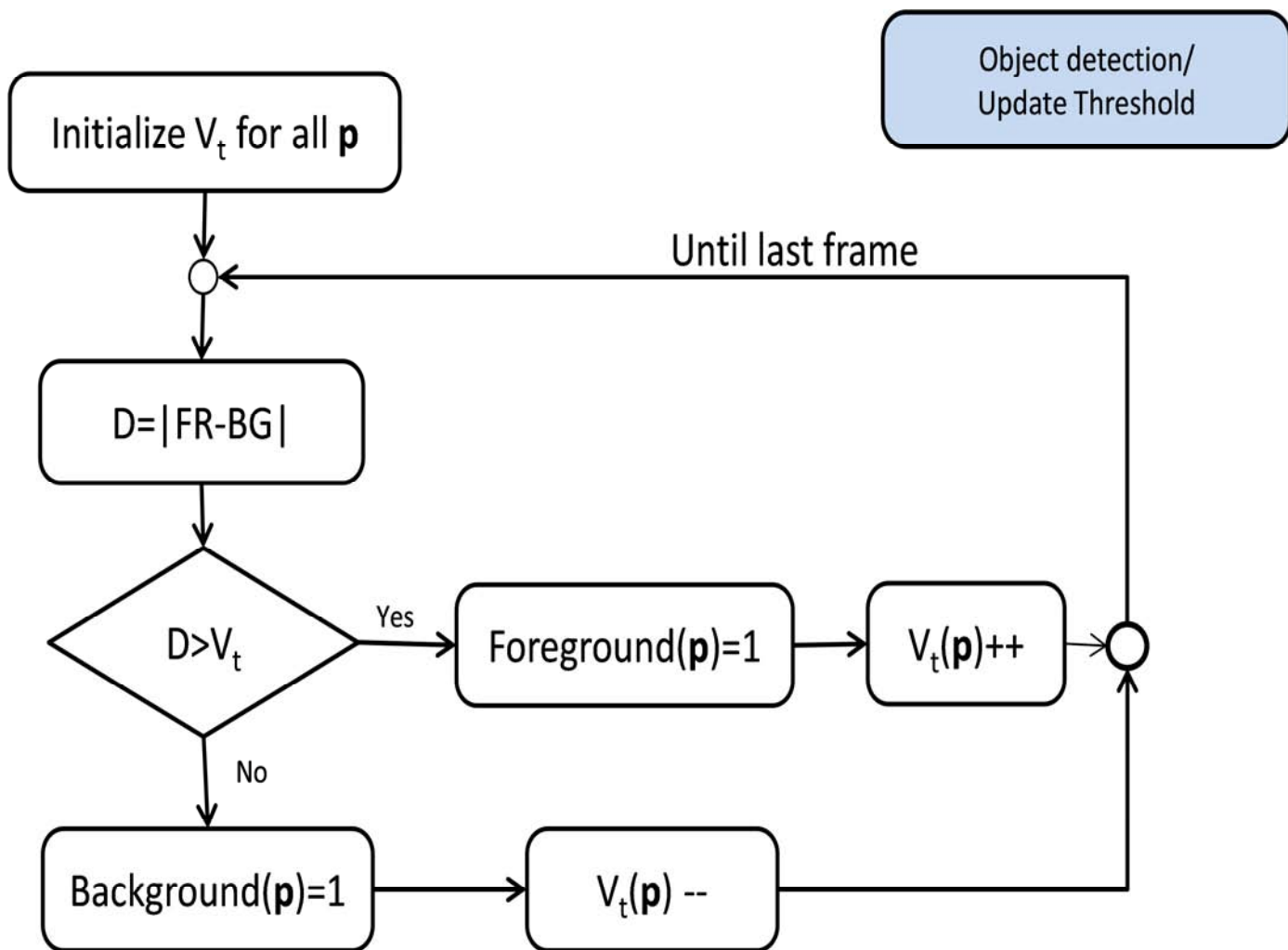
Ο αλγόριθμος IS τηρεί δύο διαφορετικά υπόβαθρα, ένα φωτεινό και ένα σκοτεινό (B_{light} , B_{dark}), καθώς και το τρέχον υπόβαθρο (BG), το οποίο επιλέγεται ως ένα από τα δύο. Το τρέχον υπόβαθρο, το B_{light} και το B_{dark} τίθενται ίσα με το πρώτο στιγμιότυπο. Για κάθε ένα από τα επόμενα στιγμιότυπα:

- Το τρέχον υπόβαθρο επικαιροποιείται με βάση έναν αλγόριθμο φίλτρου ενδιάμεσης τιμής.
- Αν υπάρχει σημαντική μεταβολή στην εντροπία μεταξύ του τρέχοντος και του προηγούμενου στιγμιότυπου, τότε:
 - Αν η μέση τιμή του φωτεινού υπόβαθρου είναι μεγαλύτερη από αυτήν του BG , τότε το B_{light} τίθεται ίδιο με το BG
 - Αν η μέση τιμή του φωτεινού υπόβαθρου είναι μικρότερη από αυτήν του BG , τότε το B_{dark} τίθεται ίδιο με το BG
 - Αν η μέση τιμή του τρέχοντος υπόβαθρου είναι μικρότερη από αυτήν του τρέχοντος στιγμιότυπου, τότε το BG τίθεται ίδιο με το B_{light} , αλλιώς το BG τίθεται ίδιο με το B_{dark}
- Τμηματοποίηση του προσκηνίου βάσει κατωφλίωσης.

Ο αλγόριθμος αυτός περιγράφεται διαγραμματικά στο σχήμα 19. Το δε βήμα της τμηματοποίησης αναλύεται στο σχήμα 20.



Σχήμα 19. Λογικό διάγραμμα αλγορίθμου τμηματοποίησης IS.



Σχήμα 20. Λογικό διάγραμμα του βήματος της τμηματοποίησης (IS) βάσει κατωφλίωσης.

3.4. Ανίχνευση Ανθρώπινης Σιλουέτας σε βίντεο

3.4.1. Αντικειμενική συνάρτηση

Η αντικειμενική συνάρτηση ποσοτικοποιεί την σύμπτωση της απεικόνισης I_M του μοντέλου του ανθρώπου από το βαθμονομημένο μοντέλο της κάμερας με την τμηματοποιημένη σιλουέτα I_S από το τρέχον στιγμιότυπο του βίντεο.

Η αντικειμενική συνάρτηση ορίζεται ως ακολούθως:

$$f(p) = - \sum_{image} I_M \cap I_S \quad (12)$$

όπου: το σύμβολο \cap συμβολίζει το λογικό τελεστή ΚΑΙ, ενώ η άθροιση γίνεται σε όλα τα εικονοστοιχεία της εικόνας. Το αρνητικό πρόσημο έχει εισαχθεί με σκοπό η συνάρτηση αυτή να ελαχιστοποιηθεί. Έτσι, η αντικειμενική συνάρτηση ορίζεται σαν το πλήθος των κοινών μη μηδενικών εικονοστοιχείων της I_M και της I_S .

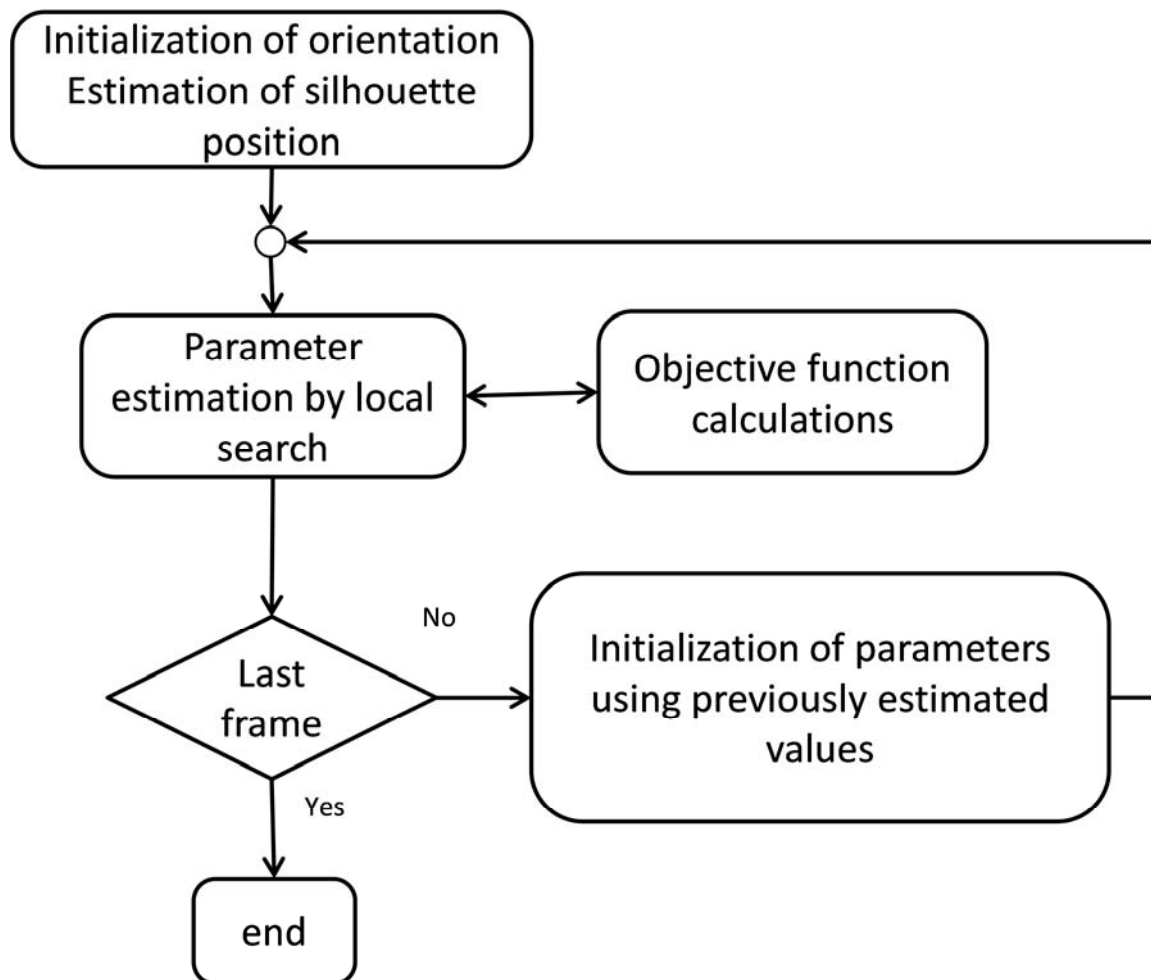
3.4.2. Εύρεση στάσης σιλουέτας

Ο αλγόριθμος εύρεσης της στάσης της σιλουέτας, όπως παρουσιάζεται διαγραμματικά στο σχήμα 21, περιγράφεται ως εξής:

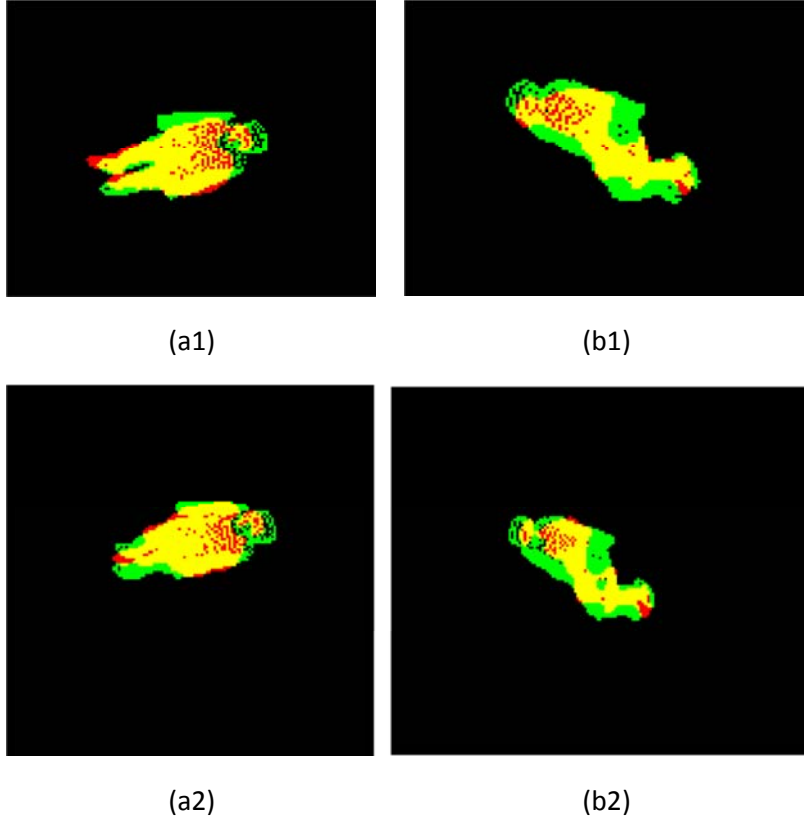
- Τυχαία αρχικοποίηση του προσανατολισμού και αρχικοποίηση της πραγματικής θέσης με τον τρόπο που περιγράφηκε στην παράγραφο 4.1.3 «Υπολογισμός Θέσης Τμηματοποιημένης Σιλουέτας».
- Για κάθε επόμενο στιγμιότυπο:
 - Προσδιορισμός των παραμέτρων του μοντέλου με χρήση μεθόδου τοπικής βελτιστοποίησης (συνάρτηση `fminsearch` του περιβάλλοντος ανάπτυξης MATLAB) της αντικειμενικής συνάρτησης.

- Νέα αρχικοποίηση παραμέτρων συνάρτησης ελαχιστοποίησης με βάση τα αποτελέσματα του προηγούμενου βήματος.

Με τη χρήση του τοπικού άπληστου αλγόριθμου βελτιστοποίησης επιτυγχάνουμε προσδιορισμό των παραμέτρων με 20-30 υπολογισμούς της αντικειμενικής συνάρτησης για κάθε στιγμιότυπο, το σύνολο των οποίων λαμβάνει χώρα σε περίπου 900msec σε ένα τυπικό υπολογιστικό μηχάνημα σε περιβάλλον MATLAB (Intel i5 laptop with 4 GB Ram). Αποτελέσματα της εύρεσης της στάσης με τη χρήση της παραπάνω διαδικασίας παρουσιάζονται στο σχήμα 22 (α), (β). Με κόκκινο χρώμα απεικονίζεται η εικόνα I_S και με πράσινο η I_M , ενώ με κίτρινο χρώμα απεικονίζεται η τομή τους ($I_M \cap I_S$).



Σχήμα 21. Διαγραμματική απεικόνιση αλγορίθμου εύρεση στάσης ανθρώπινης σιλουέτας.



Σχήμα 22. Οπτικοποίηση διαφόρων αποτελεσμάτων εύρεσης ανθρώπινης στάσης. Στις εικόνες (a1) και (a2) φαίνεται η «όρθια» στάση και στις (b1) και (b2), η «καθιστή».

4. Αποτελέσματα

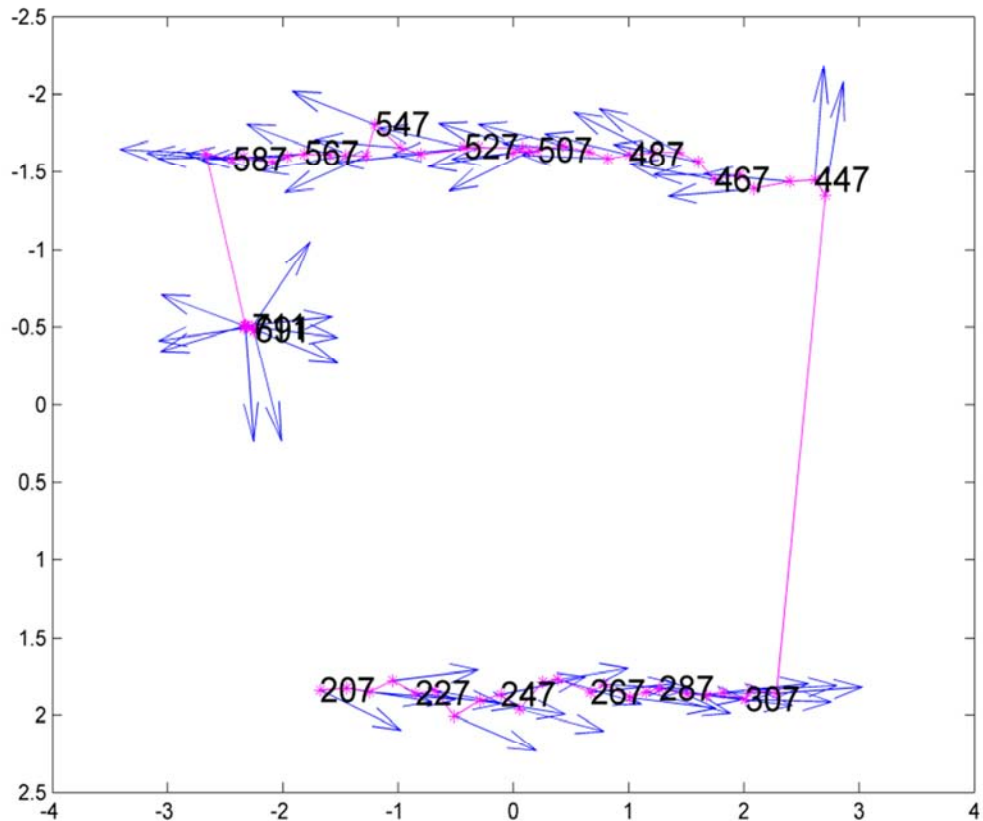
Λάβαμε ένα μικρό αριθμό από βίντεο ανθρώπινης δραστηριότητας σε χώρο στον οποίο έχει εγκατασταθεί η fisheye κάμερα. Ως ανθρώπινη δραστηριότητα θεωρούμε τις διαφορετικές φάσεις του βαδίσματος, καθώς και τη στάση του όρθιου και καθιστού. Πιο συγκεκριμένα, λήφθηκαν:

- δύο βίντεο στα οποία η ανθρώπινη δραστηριότητα είναι το βάδισμα
- δύο στα οποία ελέγχεται η στάση «όρθιος» - «καθιστός».

Με χρήση των αλγορίθμων που περιγράφηκαν προχωρήσαμε σε αυτόματη αναγνώριση της θέσης, της στάσης και της γωνίας προσανατολισμού των ατόμων που καταγράφηκαν.

4.1. Αποτελέσματα στα οποία η ανθρώπινη δραστηριότητα είναι το βάδισμα

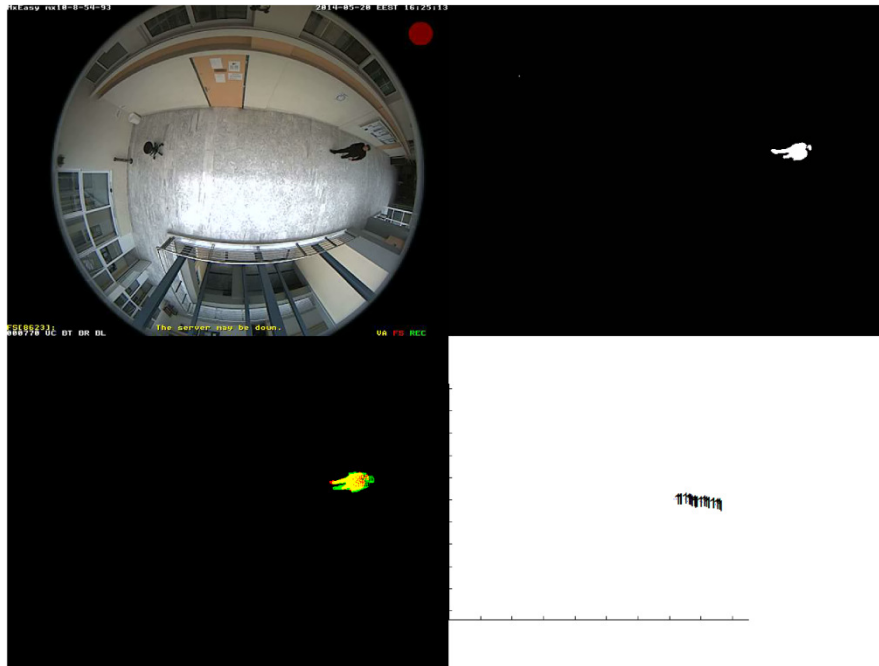
Το σχήμα 23 δείχνει τις προσδιορισμένες θέσεις και τον προσανατολισμό της ανθρώπινης σιλουέτας, όπως προσδιορίστηκαν με την παραπάνω μέθοδο, για κάθε στιγμιότυπο του βίντεο 1. Ο άνθρωπος κινείται από αριστερά προς τα δεξιά στο κάτω μέρος του σχήματος και με αντίθετη φορά στο πάνω μέρος. Περίπου στο στιγμιότυπο 691, ο άνθρωπος περιστρέφεται στατικά. Μεταξύ των στιγμιότυπων 307 και 447 ο αλγόριθμος τμηματοποίησης αποτυγχάνει παροδικά, γεγονός το οποίο ο αλγόριθμος αντιπαρέρχεται ελέγχοντας το πλήθος των λευκών εικονοστοιχείων της I_s και επαναπροσδιορίζοντας τη θέση της σιλουέτας.



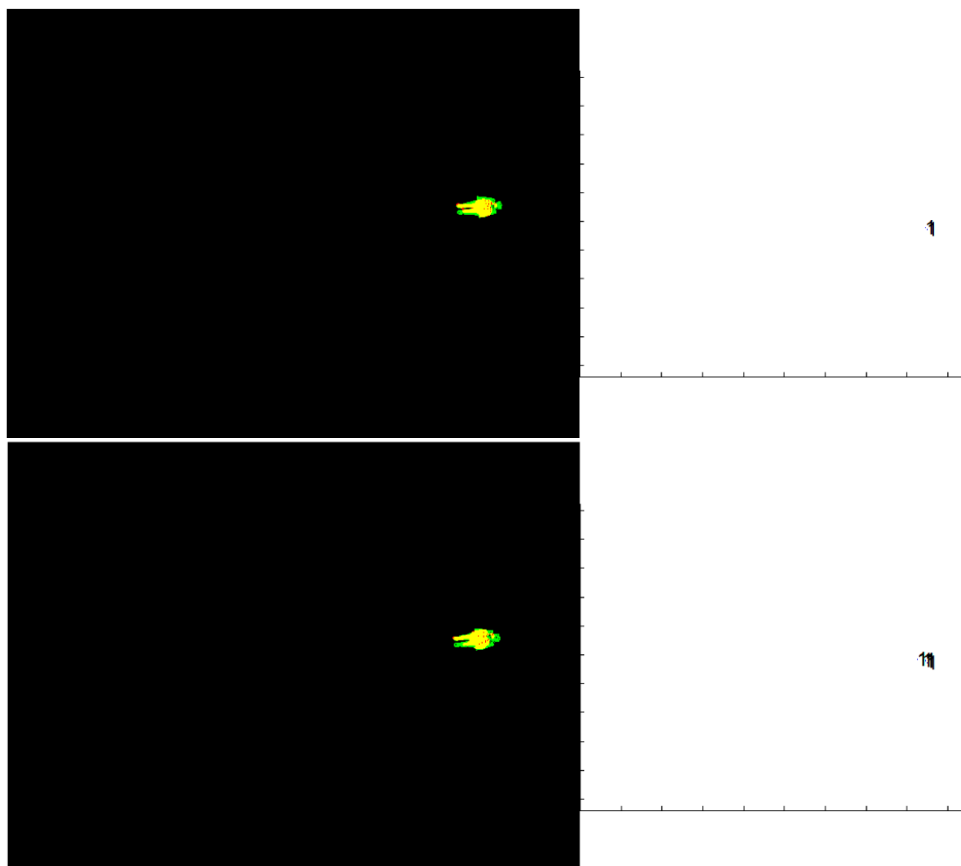
Σχήμα 23. Αποτελέσματα του προσδιορισμού θέσης και προσανατολισμού της τμηματοποιημένης ανθρώπινης σιλουέτας. Τα στιγμιότυπα που χρησιμοποιήθηκαν απαριθμούνται.

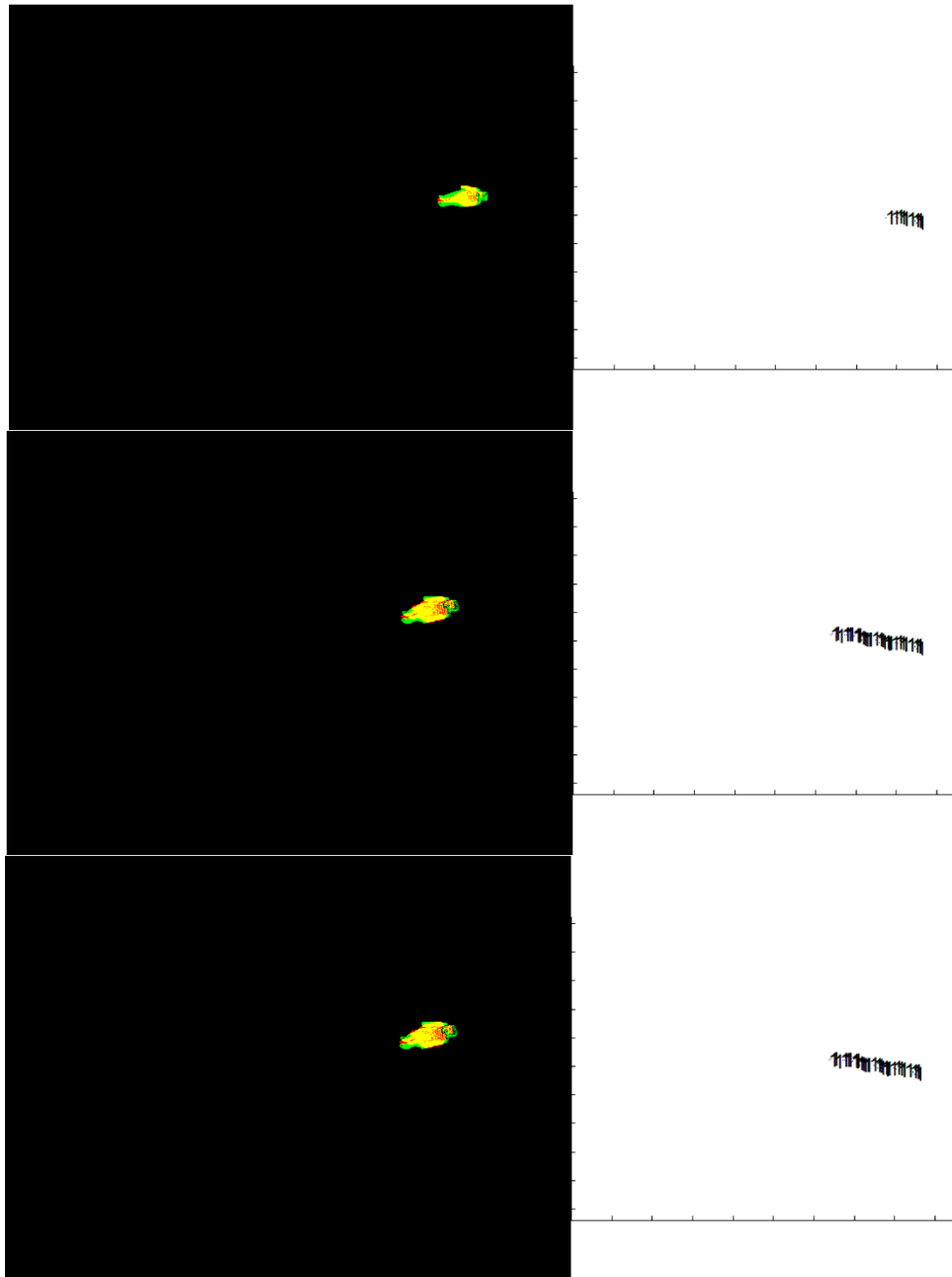
Στο σχήμα 24 φαίνονται τα διάφορα στάδια του προσδιορισμού θέσης και της πορείας: το αρχικό στιγμιότυπο (πάνω αριστερά), τη τμηματοποιημένη εικόνα I_S (πάνω δεξιά), το αποτέλεσμα του συνδυασμού της I_S και I_M (κάτω αριστερά) και τέλος η γραφική παράσταση της τροχιάς σε πραγματικές συντεταγμένες.

Στο επόμενο σχήμα (σχήμα 25) φαίνονται το αποτέλεσμα του συνδυασμού της I_S και I_M και η γραφική παράσταση της τροχιάς του ανθρώπου σε πραγματικές συντεταγμένες από ενδεικτικά στιγμιότυπα του βίντεο.



Σχήμα 24. Τα διάφορα στάδια του προσδιορισμού θέσης και τροχιάς.





Σχήμα 25. Αποτέλεσμα συνδυασμού I_S και I_M και γραφική παράσταση τροχιάς σε πραγματικές συντεταγμένες από ενδεικτικά στιγμιότυπα του βίντεο.

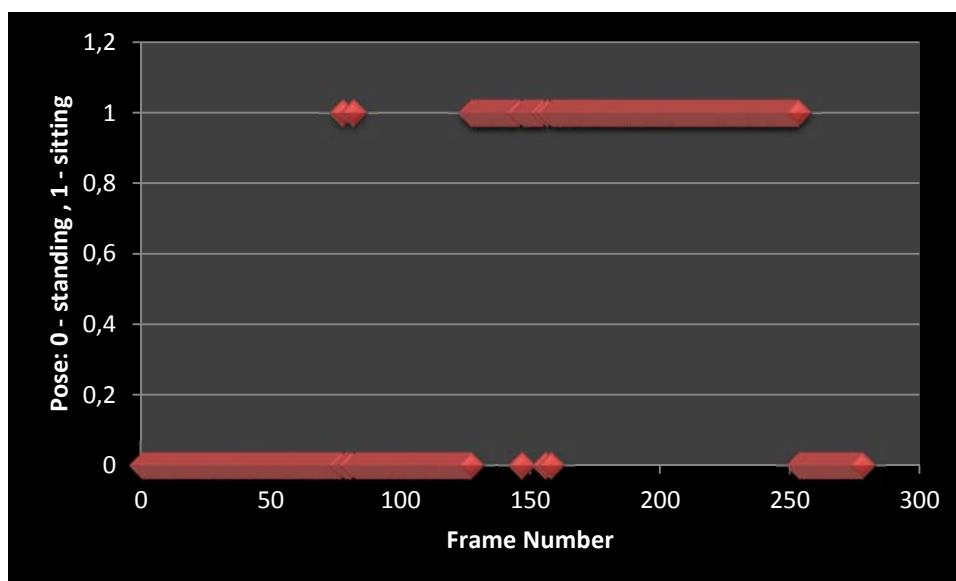
4.2. Αποτελέσματα στα οποία ελέγχεται η στάση «όρθιος» - «καθιστός»

Ο παρακάτω πίνακας (πίνακας 2) συνοψίζει την ακρίβεια, την ευαισθησία και την ειδικότητα του αυτόματου καθορισμού της στάσης (όρθιος/καθιστός) για τα δύο βίντεο που λήφθηκαν. Θεωρούμε σαν αληθώς θετική –TP (αρνητική –TN) αναγνώριση το πλήθος των εικονοστοιχείων στα οποία αναγνωρίστηκε σωστά η στάση «καθιστός» («όρθιος»).

Πίνακας 2. Αποτελέσματα αναγνώρισης στάσης «όρθιος» - «καθιστός» .

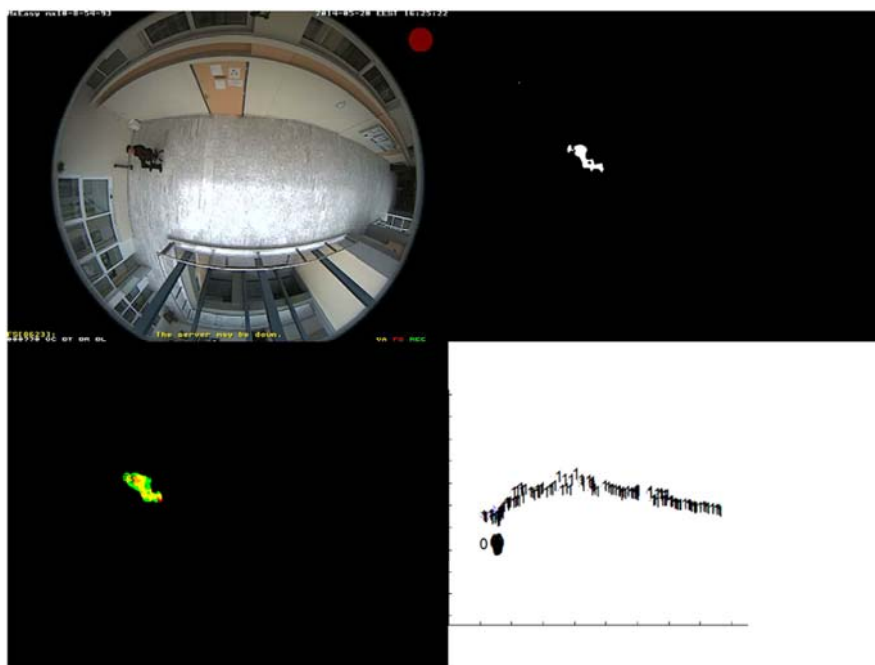
	TP	TN	FN	FP	Accuracy	Sensitivity	Specificity
Video 1	110/278	112/278	23/278	33/278	0.7986	0.8271	0.7724
Video 2	98/164	54/164	5/164	7/164	0.9268	0.9515	0.8852

Στο σχήμα 26 δίνεται η γραφική παράσταση του αποτελέσματος του προσδιορισμού της στάσης «όρθιος» - «καθιστός» συναρτήσεως των στιγμιότυπων του βίντεο. Τα στιγμιότυπα που αντιστοιχούν σε μηδενική τιμή στον Υ άξονα κατοπτρίζουν την όρθια στάση και εκείνα της μοναδιαίας τιμής, την καθιστή στάση του ανθρώπου.



Σχήμα 26. Γραφική παράσταση προσδιορισμού στάσης «όρθιος» - «καθιστός» συναρτήσεως των στιγμιότυπων του βίντεο.

Στο σχήμα 27 φαίνονται τα διάφορα στάδια του προσδιορισμού θέσης και στάσης: το αρχικό στιγμιότυπο του βίντεο (πάνω αριστερά), τη τμηματοποιημένη εικόνα I_S (πάνω δεξιά), το αποτέλεσμα του συνδυασμού της I_S και I_M (κάτω αριστερά) και τέλος η γραφική παράσταση της τροχιάς σε πραγματικές συντεταγμένες. Οι τιμές 0 και 1 που αναγράφονται στα στιγμιότυπα αντιστοιχούν στην καθιστή και όρθια στάση, αντίστοιχα. Διακρίνεται η αναγνώριση ($I_M \cap I_S$ και ανάθεση μηδενικής τιμής) της πραγματικά καθιστής στάσης (βλ. I_S) στο συγκεκριμένο στιγμιότυπο.



Σχήμα 27. Τα διάφορα στάδια του προσδιορισμού θέσης, τροχιάς και στάσης.

Η διαδικασία εύρεσης στάσης βασίζεται στη σύγκριση της βέλτιστης τιμής («FmaxO» για την όρθια στάση, «FmaxK» για την καθιστή) της αντικειμενικής συνάρτησης (παράγραφος 4.4.1) για τις δύο πιθανές στάσεις, όπως περιγράφηκε στην παράγραφο 4.4.2. Για το σκοπό αυτό απαιτείται ο προσδιορισμός της θέσης στο πραγματικό σύστημα συντεταγμένων (x_{real}, y_{real}), καθώς και του προσανατολισμού (θ_{rot}). Στον παρακάτω πίνακα (πίνακας 3) παρατίθενται οι σχετικές αριθμητικές τιμές οι οποίες υπολογίστηκαν για κάποια στιγμιότυπα ενός από τα βίντεο που

λήφθηκαν, καθώς και το αποτέλεσμα του προσδιορισμού της στάσης (οι τιμές 0 , 1 αντιστοιχούν στις στάσεις «όρθιος» , «καθιστός»).

Πίνακας 3. Αριθμητικές τιμές ενδεικτικών στιγμιотύπων.

Frame	θ_{rot}	x_{real}	y_{real}	FmaxO	FmaxK	Αποτέλεσμα
394	5.548837	-2.01565	-0.57925	-715	-499	0
395	5.335974	-2.0639	-0.55903	-699	-457	0
396	5.703916	-2.17352	-0.57948	-676	-477	0
397	5.994467	-2.22654	-0.54301	-539	-380	0
398	5.623489	-2.27045	-0.56711	-498	-374	0
399	5.32511	-2.25072	-0.50151	-491	-381	0
400	6.018961	-2.3709	-0.47492	-473	-342	0
401	5.859846	-2.45964	-0.38399	-543	-409	0
402	5.888832	-2.48987	-0.3819	-597	-453	0
403	5.796247	-2.46161	-0.34846	-667	-490	0
404	5.452914	-2.42155	-0.30813	-732	-555	0
405	5.592719	-2.50736	-0.31102	-772	-583	0
406	5.330917	-2.48494	-0.26573	-760	-600	0
407	5.607521	-2.54519	-0.26548	-722	-589	0
408	5.508148	-2.58419	-0.266	-730	-593	0
409	5.504697	-2.59538	-0.25011	-679	-566	0
410	5.044474	-2.55022	-0.2069	-642	-523	0
411	5.103683	-2.53969	-0.15921	-666	-506	0
412	5.962107	-2.74183	-0.20161	-615	-488	0
413	5.126295	-2.58175	-0.14169	-659	-460	0
414	5.990709	-2.74144	-0.1885	-615	-458	0
415	6.21891	-2.97195	-0.23802	-601	-439	0
416	6.219349	-2.95934	-0.24202	-611	-440	0
417	6.238721	-3.01209	-0.26865	-604	-435	0
418	6.238139	-2.66428	-0.20176	-579	-425	0
419	5.808436	-2.6544	-0.21816	-569	-405	0
420	5.808436	-2.79277	-0.28073	-571	-429	0
421	5.549132	-2.74823	-0.29191	-578	-427	0
422	2.831932	-3.00656	-0.34907	-603	-638	1
423	3.146934	-2.66484	-0.27483	-589	-620	1
424	3.214266	-2.61222	-0.2926	-587	-618	1
425	2.926164	-2.68193	-0.32878	-592	-624	1
426	3.499456	-2.63761	-0.3101	-582	-619	1
427	3.442461	-2.60322	-0.27601	-548	-599	1
428	3.450215	-2.54076	-0.26814	-544	-583	1
429	3.692203	-2.57618	-0.333	-527	-602	1
430	3.776578	-2.58488	-0.3577	-519	-600	1
431	3.46757	-2.52581	-0.26432	-517	-550	1

432	3.232959	-2.51604	-0.3758	-526	-575	1
433	3.659539	-2.48115	-0.31752	-544	-598	1
434	3.880198	-2.56601	-0.37813	-552	-627	1
435	3.83949	-2.61996	-0.40797	-551	-629	1
436	3.80303	-2.56816	-0.37034	-552	-610	1
437	3.923909	-2.53718	-0.39867	-542	-623	1
438	4.067428	-2.5376	-0.43265	-541	-644	1
439	4.063326	-2.54835	-0.45655	-455	-618	1
440	4.080812	-2.55358	-0.45465	-484	-613	1
441	4.080812	-2.43208	-0.3866	-489	-461	0
442	4.186327	-2.55615	-0.46925	-427	-582	1
443	4.174411	-2.54748	-0.47985	-477	-594	1
444	4.075447	-2.56293	-0.46714	-445	-587	1
445	4.21097	-2.54123	-0.45519	-451	-577	1
446	4.147143	-2.59998	-0.49801	-456	-590	1
447	3.923984	-2.23606	-0.31879	-430	-439	1
448	4.021901	-2.55881	-0.45821	-475	-621	1
449	4.062196	-2.57359	-0.46091	-497	-623	1
450	3.659116	-2.43696	-0.4014	-492	-424	0
451	3.961489	-2.52349	-0.40788	-502	-599	1
452	3.678039	-2.42721	-0.38644	-501	-451	0
453	4.070126	-2.58927	-0.44217	-490	-612	1
454	3.856993	-2.49535	-0.43485	-493	-588	1
455	3.930851	-2.40068	-0.35552	-495	-526	1
456	3.649782	-2.39212	-0.36537	-494	-482	1
457	4.055981	-2.59533	-0.44999	-477	-595	1
458	4.055981	-2.37953	-0.36739	-490	-483	1
459	4.059022	-2.55633	-0.40706	-481	-572	1
460	4.054716	-2.59972	-0.44679	-472	-598	1
461	4.077678	-2.59334	-0.44839	-452	-585	1
462	4.095	-2.59167	-0.43964	-477	-595	1
463	3.735567	-2.52308	-0.43478	-474	-560	1
464	3.735567	-2.40624	-0.38404	-475	-429	1
465	3.924211	-2.51655	-0.418	-490	-577	1
466	3.924211	-2.39273	-0.36591	-490	-457	1
467	4.038208	-2.58722	-0.4504	-476	-583	1
468	4.091439	-2.62669	-0.46102	-483	-593	1
469	4.057878	-2.60397	-0.44667	-493	-600	1
470	4.098734	-2.59496	-0.45284	-500	-597	1
471	4.020706	-2.62191	-0.45529	-483	-601	1
472	3.925604	-2.44227	-0.33707	-494	-495	1
473	4.088142	-2.59658	-0.4602	-494	-598	1
474	3.986474	-2.54764	-0.4281	-491	-584	1
475	4.003585	-2.58365	-0.44395	-487	-592	1

476	4.024857	-2.62753	-0.48142	-490	-593	1
477	4.090848	-2.60332	-0.45981	-491	-600	1
478	4.461585	-2.47913	-0.38739	-497	-522	1
479	3.953773	-2.57985	-0.46152	-492	-591	1
480	4.082439	-2.57547	-0.44243	-492	-594	1
481	4.107351	-2.58439	-0.44344	-476	-583	1
482	3.780158	-2.54768	-0.44977	-482	-566	1
483	3.780158	-2.48904	-0.36818	-496	-410	1
484	4.059561	-2.57102	-0.4227	-488	-604	1
485	3.658388	-2.46282	-0.37996	-514	-403	1
486	3.957061	-2.39751	-0.40104	-497	-571	1
487	4.045006	-2.4885	-0.46187	-531	-646	1
488	4.04574	-2.55795	-0.50385	-523	-644	1
489	3.972095	-2.42355	-0.46572	-559	-680	1
490	3.929776	-2.52851	-0.51563	-540	-692	1
491	3.838638	-2.39095	-0.48549	-521	-670	1
492	3.452212	-2.3528	-0.54002	-492	-568	1
493	3.172075	-2.28608	-0.562	-523	-578	1
494	3.082358	-2.34436	-0.54911	-529	-554	1
495	3.038788	-2.27543	-0.57051	-527	-557	1
496	3.01084	-2.32565	-0.59708	-566	-600	1
497	3.331219	-2.35868	-0.5552	-558	-598	1
498	2.900962	-2.43968	-0.64937	-557	-579	1
499	3.35123	-2.57346	-0.60634	-545	-603	1
500	2.798178	-2.52765	-0.66038	-565	-577	1
501	10.07312	-2.52997	-0.52814	-555	-597	1
502	8.944446	-2.55654	-0.68352	-551	-576	1
503	6.228392	-2.42799	-0.57888	-531	-307	1
504	6.228392	-2.39856	-0.56904	-515	-454	1
505	6.228392	-2.36714	-0.5604	-517	-462	1
506	6.228392	-2.39896	-0.56355	-508	-468	1
507	3.513741	-2.51722	-0.569	-497	-571	1
508	3.082705	-2.59114	-0.6555	-483	-552	1
509	2.773548	-2.58223	-0.65882	-483	-549	1
510	3.629717	-2.51154	-0.5545	-495	-565	1
511	3.178312	-2.40416	-0.61312	-468	-567	1
512	3.175324	-2.3955	-0.62773	-478	-570	1
513	3.164204	-2.36574	-0.6163	-462	-561	1
514	3.565471	-2.4411	-0.53533	-444	-509	1
515	3.119967	-2.29966	-0.5743	-419	-502	1
516	2.930906	-2.52829	-0.66338	-426	-432	1
517	3.131415	-2.35245	-0.59209	-419	-488	1
518	2.72215	-2.39765	-0.64853	-417	-452	1
519	3.593679	-2.36515	-0.54365	-415	-467	1

520	3.660652	-2.41979	-0.54381	-437	-499	1
521	3.660652	-2.29114	-0.55654	-429	-425	1
522	3.108805	-2.32878	-0.58165	-411	-460	1
523	3.588485	-2.39054	-0.50815	-440	-513	1
524	3.634173	-2.40507	-0.52231	-436	-532	1
525	3.963736	-2.42528	-0.48518	-464	-551	1
526	5.208588	-2.46549	-0.52321	-483	-398	1
527	5.208588	-2.32196	-0.55763	-482	-455	1
528	5.208588	-2.39196	-0.4901	-475	-466	1
529	5.208588	-2.38867	-0.49096	-456	-438	1
530	4.860795	-2.35571	-0.40149	-495	-486	1
531	4.197221	-2.38023	-0.41246	-473	-527	1
532	3.740637	-2.52051	-0.46133	-508	-570	1
533	3.780582	-2.48688	-0.4664	-489	-569	1
534	3.806358	-2.42756	-0.44104	-501	-575	1
535	4.941699	-2.36371	-0.37875	-476	-442	1
536	3.94673	-2.4891	-0.44634	-503	-577	1
537	5.130359	-2.43902	-0.37759	-490	-420	1
538	5.130359	-2.40224	-0.37372	-479	-428	1
539	5.130359	-2.43018	-0.34732	-476	-427	1
540	3.950885	-2.44037	-0.38198	-476	-517	1
541	3.860302	-2.47235	-0.36136	-472	-502	1
542	3.886595	-2.50572	-0.37161	-463	-504	1
543	3.673844	-2.47981	-0.31844	-452	-465	1
544	3.577957	-2.48334	-0.32452	-451	-494	1
545	3.611979	-2.44496	-0.30085	-449	-492	1
546	3.601964	-2.44231	-0.30131	-431	-485	1
547	3.490963	-2.54469	-0.29867	-473	-506	1
548	5.709686	-2.56259	-0.30303	-482	-412	0
549	6.187539	-2.65549	-0.37893	-529	-421	0
550	6.187539	-2.55958	-0.29771	-503	-414	0
551	6.209816	-2.65339	-0.3312	-513	-427	0
552	6.104295	-2.62888	-0.2982	-505	-423	0
553	6.282061	-2.68115	-0.30112	-528	-419	0
554	6.2536	-2.68333	-0.29418	-525	-418	0
555	6.175683	-2.71998	-0.27975	-538	-427	0
556	6.149583	-2.73738	-0.27954	-560	-441	0
557	5.359582	-2.73152	-0.27833	-571	-441	0
558	6.158049	-2.85087	-0.26106	-604	-463	0
559	4.779473	-2.66476	-0.16188	-650	-423	0
560	4.939404	-2.66483	-0.14255	-629	-440	0
561	5.202781	-2.76507	-0.14394	-598	-435	0
562	5.55132	-2.93481	-0.06827	-587	-492	0
563	5.346352	-2.89574	-0.04069	-594	-487	0

5. Συμπεράσματα και Μελλοντική εργασία

Στην εργασία αυτή παρουσιάστηκε μία προσέγγιση στο πρόβλημα της αναγνώρισης θέσης και στάσης σιλουέτας σε βίντεο από fisheye καμερα οροφής κλειστού χώρου. Η σιλουέτα αποτελεί προβολή ενός 3D αντικειμένου (του ανθρώπινου σώματος) σε μία 2D εικόνα. Λόγω της συνεπακόλουθης απώλεια της πληροφορίας βάθους, το πρόβλημα του εντοπισμού της στάσης του ανθρώπου βάσει της σιλουέτας συχνά δεν επιδέχεται μοναδικής λύσης. Επισημαίνεται ότι η δυαδική φύση της σιλουέτας και η χρήση της κάμερας ευρυγώνιου φακού (fisheye), τοποθετημένης στην οροφή, επιτείνει το χαρακτηριστικό αυτό. Η προσέγγιση που περιγράφεται στην παρούσα εργασία περιέχει ένα αριθμό από νέα στοιχεία:

- Το γεωμετρικό μοντέλο του ανθρώπινου σώματος που χρησιμοποιήθηκε αποτελείται από γενικευμένους κυλίνδρους με κατά τμήματα ευθεία κεντρική καμπύλη και ελλειπτικές διατομές με μεταβλητού μήκους ημιάξονες. Το μοντέλο μπορεί να αναπαραστήσει διάφορες σχετικά απλές στάσεις του ανθρώπινου σώματος, όπως «όρθιος», «καθιστός» και τις φάσεις του βαδίσματος.
- Το υπολογιστικό μοντέλο της fisheye καμερας, μετά την βαθμονόμηση του, είναι ακριβές και επιτρέπει την εξομοιωμένη απεικόνιση της ανθρώπινης σιλουέτας σε διάφορες στάσεις.
- Η εξομοιωμένη απεικόνιση, καθώς και η τμηματοποίηση του βίντεο εισάγονται σε μία αντικειμενική συνάρτηση η οποία υπολογίζει την μεταξύ τους ταύτιση. Η βελτιστοποίηση της συνάρτησης αυτής είναι εφικτή σε με ρυθμό μικρότερο από 1 στιγμιότυπο /sec με χρήση τοπικής βελτιστοποίησης
- Ποιοτική και ποσοτική αποτίμηση των πειραματικών δοκιμών έδειξαν ότι η βελτιστοποίηση της αντικειμενικής συνάρτησης οδηγεί σε σωστές τιμές παραμέτρων του μοντέλου σε μεγάλο ποσοστό των στιγμιότυπων. Η θέση της κάμερας (οροφή) συχνά δυσχεραίνει την βελτιστοποίηση της συνάρτησης.

Κάθε ένα από τα στοιχεία αυτά επιδέχεται νέες προσεγγίσεις και υλοποιήσεις, οι οποίες αποτελούν αντικείμενο μελλοντικής εργασίας. Πιο συγκεκριμένα:

- Το γεωμετρικό μοντέλο που χρησιμοποιήθηκε μπορεί να αντικατασταθεί από ένα άλλο, στατιστικής φύσεως, το οποίο θα είναι σε θέση να αντιπροσωπεύσει μεγάλο πλήθος στάσεων, αλλά και σωματότυπων με σχετικά μικρό αριθμό παραμέτρων [40].
- η βελτιστοποίηση της αντικειμενικής συνάρτησης μπορεί να επιτευχθεί με μεθόδους ολικής βελτιστοποίησης, ισχυρότερες από την απλή τοπική βελτιστοποίηση που χρησιμοποιήθηκε. Διάφορες τεχνικές μπορούν να χρησιμοποιηθούν ώστε να διατηρείται ένας πληθυσμός λύσεων ο οποίος να έχει

δυνατότητα αποτελεσματικής εξερεύνησης του χώρου των λύσεων με μικρό αριθμό αποτιμήσεων της αντικειμενικής συνάρτησης.

- Η ανίχνευση της στάσης του ατόμου που βασίστηκε σε 3D μοντέλο του ανθρώπου μπορεί να συνδυαστεί με περιγραφείς που βασίζονται στη σιλουέτα, ή στις τιμές στα pixels του στιγμιότυπου της σιλουέτας.
- Το 3D μοντέλο του ανθρώπου, σε συνδυασμό με το βαθμονομημένο μοντέλο της fisheye κάμερας μπορεί να χρησιμοποιηθεί για τον υπολογισμό περιγραφέων την επιλογή των πιο κατάλληλων και την εκπαίδευση αλγορίθμων κατηγοριοποίησης στάσης, βάσει των περιγραφέων αυτών.
- Τέλος, το πλήθος των πιθανών στάσεων που μπορούν να αναγνωριστούν μπορεί να διευρυνθεί με την εφαρμογή των προαναφερθέντων βελτιώσεων.

Οι μελλοντικές αυτές επεκτάσεις της παρούσας εργασία βρίσκονται στην αιχμή της έρευνας στον χώρο της επεξεργασίας και ανάλυσης βίντεο, διάχυτου και μη επεμβατικού υπολογισμού, ρομποτικής και τεχνητής νοημοσύνης και θα μπορούσαν να οδηγήσουν σε εφαρμογές σε ποικίλους τομείς δραστηριοτήτων, όπως βιοϊατρική έρευνα, ασφάλεια και επιτήρηση χώρων, υποβοήθηση της διαβίωσης χρόνια πασχόντων, αυτόνομα συστήματα και ρομποτική.

6. Αναφορές

1. R. Poppe, Vision-based human motion analysis: An overview, *Computer Vision and Image Understanding*, 108, (2007), 4–18.
2. R. Poppe, A survey on vision-based human action recognition, *Image and Vision Computing*, 28, (2010), 976–990.
3. T. B. Moeslund and E. Granum, A Survey of Computer Vision-Based Human Motion Capture, *Computer Vision and Image Understanding* 81, 231–268 (2001).
4. J. K. Aggarwal, M. S. Ryoo, Human Activity Analysis: A Review, *ACM Computing Surveys*, Vol. 43(3), 2011
5. D. Weinland, R. Ronfard, E. Boyer, Free viewpoint action recognition using motion history volumes, *Computer Vision and Image Understanding*, 103(2-3), 249-257, 2006.
6. M. Blank, L. Gorelick, E. Shechtman, M. Inari, R. Barsi, Actions as space-time shapes, in *IEEE International Conf. on Computer Vision*, 2005, pp 1395-1402.
7. C. Schuldt, I. Laptev, B. Caputo, Recognizing human action: a local svm approach, in *ICPR*, 2004, pp32-36
8. R. Gross, J. Shi, The cmu motion of body (mobo) database, Tech. Rep. CMU-RI-TR-01-18, Robotics Institute, Pittsburgh, PA (June 2001)
9. L. Sigal, M. Black, Human eva: Synchronized video and motion capture dataset for evaluation of articulated human motion, Tech. Rep. BrownUniversity 2006.
10. I. Laptev, M. Marszalek, C. Schmidt, B. Rozenfeld, Learning realistic human actions from movies, in *IEEE Conf on Computer Vision and Pattern Recognition*, 2008.
11. N. Howe, Boundary fragment matching and articulated pose under occlusion. In: *Proceedings of the International Conference on Articulated Motion and Deformable Objects (AMDO'06)*. Number 4069 in *Lecture Notes in Computer Science*, Port d'Andratx, Spain (July 2006) 271–280.
12. R. Poppe and M. Poel, Body-part templates for recovery of 2D human poses under occlusion, *Articulated Motion and Deformable Objects*, 2008 - Springer
13. R. Rosales and S. Sclaroff, Learning and synthesizing human body motion and posture, in *The Fourth International Conference on Automatic Face and Gesture Recognition, Grenoble, France, March 2000*.
14. M. Brand, Shadow puppetry, in *International Conference on Computer Vision, Corfu, Greece, September 1999*.

15. M. Brand. Shadow puppetry. In *Proceedings of the International Conference on Computer Vision (ICCV'99) - volume 2*, pages 1237–1244, Kerkyra, Greece, September 1999.
16. R. Rosales and S. Sclaroff. Learning body pose via specialized maps. In *Advances in Neural Information Processing Systems (NIPS) 14*, pages 1263–1270, Vancouver, Canada, December 2001.
17. N. Howe. Silhouette lookup for automatic pose tracking. In *Proceedings of the Conference on Computer Vision and Pattern Recognition Workshops (CVPRW'04)*, page 15, Los Alamitos, CA, June 2004.
18. J. O'Rourke, N. Badler, Model-based image analysis of human motion using constraint propagation, *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI) 2 (6) (1980) 522–536*.
19. K. Rohr, Towards model-based recognition of human movements in image sequences, *Computer Vision, Graphics, and Image Processing: Image Understanding 59 (1) (1994) 94–115*.
20. S. Pehlivan, 3D human pose search using oriented cylinders, 2009 IEEE 12th International Conference on Computer Vision Workshops (ICCV Workshops), pp.16-22.
21. D. Droschel and S. Behnke, 3D Body Pose Estimation using an Adaptive Person Model for Articulated ICP, ICIRIA, Aachen, Germany, 2011.
22. K.K. Delibasis, V.P. Plagianakos, I. Maglogiannis, Refinement of human silhouette segmentation in omni-directional indoor videos, *Computer Vision and Image Understanding, Vol 128, November 2014, Pages 65-83*.
23. N. Max 1983, Computer Graphics Distortion for IMAX and OMNIMAX Projection, *ProcNicograph, 83, Dec 1983 pp 137*
24. Greene N. 1986, Environment Mapping and Other Applications of World Projections, *IEEE Computer Graphics and Applications, November 1986, vol. 6(11), pp 21*.
25. <http://paulbourke.net/dome/fisheye/>
26. S. Shah and J. Aggarwal, Intrinsic parameter calibration procedure for a high distortion fish-eye lens camera with distortion model and accuracy estimation, *Pattern Recognition, 29(11), 1775- 1788, 1996*.
27. M.G. Epitropakis, D.K. Tasoulis, N.G. Pavlidis, V.P. Plagianakos, and M.N. Vrahatis, "Enhancing Differential Evolution Utilizing Proximity-based Mutation Operators", *IEEE Transactions on Evolutionary Computation, vol. 15, pp. 99-119, 2011*.
28. K. K. Delibasis, A. Kechriniotis , I. Maglogiannis, A novel tool for segmenting 3D medical images based on generalized cylinders and active surfaces, *computer methods and programs in biomedicine 111 (2013) 148–165*.
29. <http://www.3dmodelfree.com/models/20966-0.htm>
30. <http://www.youtube.com/watch?v=FESHyf8yUhQ>

31. J. Bergmann, C. Alexiou, I. Smith, Procedural differences directly affect timed up and go times *J. Am. Geriatr. Soc.*, 2009, 57(11):2168-9.
32. J. Willems, G. Debar, B. Bonroy, B. Vanrumste and T. Goedemé 2009, How to detect human fall in video? An overview, In *Proceedings of the positioning and context-awareness international conference* (Antwerp, Belgium, 28 May, 2009), POCA '09.
33. R. Cucchiara, C. Grana, M. Piccardi, and A. Prati, Detecting moving objects, ghosts, and shadows in video streams. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25, 10, (2003), 1337-1442.
34. N. McFarlane and C. Schofield, Segmentation and tracking of piglets in images, *Mach. Vision Appl.* 8, 3, (May. 1995), 187-193.
35. C. Wren, A. Azarhayejani, T. Darrell and A. Pentland, Pfunder: real-time tracking of the human body, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19, 7, (October. 1997), 780-785.
36. T. Bouwmans, F. El Baf, B. Vachon, Background Modeling using Mixture of Gaussians for Foreground Detection - A Survey, *Recent Patents on Computer Science* 1, 3 (2008) 219-237.
37. C. Stauffer and W. Grimson 1999, Adaptive background mixture models for real-time tracking, In *Proceedings of the conference on computer vision and pattern recognition* (Ft. Collins, USA, June 23-25, 1999), CVPR '99. IEEE Computer Society, New York, NY, pp. 246-252.
38. F. C. Cheng, S. C. Huang and S. J. Ruan, Implementation of Illumination-Sensitive Background Modeling Approach for Accurate Moving Object Detection, *IEEE Trans. on Broadcasting*, vol. 57, no. 4, pp.794-801, 2011.
39. A. Christodoulidis, K. K. Delibasis, I. Maglogiannis, Near real-time human silhouette and movement detection in indoor environments using fixed cameras, in *the 5th ACM International Conference on Pervasive Technologies Related to Assistive Environments*, Heraklion, Crete, Greece, 2012.
40. N. Hasler, C. Stoll, M. Sunkel, B. Rosenhahn, H.-P. Seidel: A Statistical Model of Human Pose and Body Shape, *Computer Graphics Forum* (Proc. Eurographics 2009), Munich, Germany, March 2009.
41. K. Kottari, K. Delibasis, V. Plagianakos, I. Maglogiannis, Fish-Eye Camera Video Processing and Trajectory Estimation Using 3D Human Models, *Artificial Intelligence Applications and Innovations*, 385-394, Rhodes, Greece, 2014.

