

ΠΡΟΣ

- 1) Όλα τα μέλη ΔΕΠ του Τμήματος Επιστήμης Υπολογιστών
- 2) Τους εκπροσώπους των Μεταπτυχιακών φοιτητών του Τμήματος Επιστήμης Υπολογιστών
- 3) Την Επταμελή Εξεταστική Επιτροπή
- 4) Όλα τα μέλη της Πανεπιστημιακής Κοινότητας

Πρόσκληση σε Δημόσια Παρουσίαση της Διδακτορικής Διατριβής του

κ. Σιγάλα Μάρκου

Την Δευτέρα 15 Ιουνίου 2015 και ώρα 12:00 στην αίθουσα Τηλεδιάσκεψης Κ206 του Τμήματος Επιστήμης Υπολογιστών του Πανεπιστημίου Κρήτης στο Ηράκλειο, θα γίνει η δημόσια παρουσίαση και υποστήριξη της Διδακτορικής Διατριβής του υποψηφίου διδάκτορος του Τμήματος Επιστήμης Υπολογιστών κ. Σιγάλα με θέμα:

“Full-body pose tracking under severe occlusions - the Top View Reprojection approach”

“Παρακολούθηση της πόζας ανθρώπινου σώματος υπό συνθήκες έντονων επικαλύψεων - προσέγγιση με χρήση Προβολής Κάτοψης”

ΠΕΡΙΛΗΨΗ

Η ανάκτηση και παρακολούθηση της πόζας του αρθρωτού ανθρώπινου σώματος, χωρίς την χρήση οπτικών σημαδιών, είναι ένα απαιτητικό πρόβλημα με σημαντικές θεωρητικές και πρακτικές πτυχές. Η πρόσφατη εισαγωγή καμερών βάθους (RGB-D cameras) χαμηλού κόστους είχε ως αποτέλεσμα την εμφάνιση ενός πλήθους από νέες ενδιαφέρουσες και αποδοτικές προσεγγίσεις στο πρόβλημα. Ωστόσο, παρά την αξιοσημείωτη πρόοδο, η εκτίμηση της πόζας του σώματος σε ρεαλιστικά και πολύπλοκα περιβάλλοντα παραμένει ένα ανοιχτό και άκρως ενδιαφέρων ερευνητικό θέμα.

Στην παρούσα διατριβή σχεδιάσαμε, αναπτύξαμε και αξιολογήσαμε μια μεθοδολογία για την ανάκτηση και παρακολούθηση της πόζας ολόκληρου του

σώματος χωρίς τη χρήση οπτικών σημαδιών, σε ακολουθίες από RGB-D δεδομένα. Η προτεινόμενη μεθοδολογία βασίζεται στη μοντελοποίηση του ανθρώπινου σώματος, και μπορεί να εφαρμοστεί χωρίς περιορισμούς στο σενάριο δράσης των χρηστών. Πιο συγκεκριμένα, οι χρήστες μπορούν να μπαίνουν και να βγαίνουν ελεύθερα από τη σκηνή, να μετακινούνται, να ενεργούν και να αλληλεπιδρούν με άλλους χρήστες ή το ίδιο το περιβάλλον. Η έρευνά μας επικεντρώνεται κυρίως στην αντιμετώπιση προβλημάτων που προκύπτουν από επικαλύψεις, είτε μεταξύ μερών του σώματος του ίδιου χρήστη, είτε μεταξύ διαφορετικών χρηστών. Ταυτόχρονα, επιχειρούμε να αντιμετωπίσουμε επιπλέον σημαντικά θέματα, όπως οι μεγάλες διαφορές των ανθρώπινων σωμάτων ή η χωρίς περιορισμούς εκκίνηση της παρακολούθησης.

Προς αυτή την κατεύθυνση, εισαγάγαμε την καινοτόμα έννοια της Προβολής Κάτοψης (ΠΚ - Top View Reprojection) για κυλινδρικά αντικείμενα, η οποία μπορεί να ορίσει μονοσήμαντα την πόζα του κυλινδρικού αντικείμενου βασιζόμενη σε συγκεκριμένα ποσοτικά χαρακτηριστικά της Κάτοψής του, δηλαδή της όψης η οποία είναι ευθυγραμμισμένη με τον κύριο άξονα του κυλίνδρου. Δεδομένου αυτού, το πρόβλημα της εκτίμησης της πόζας ενός κυλινδρικού αντικείμενου μεταφράζεται στην εκτίμηση της αντίστοιχης Κάτοψης. Η χρησιμοποιούμενη διατύπωση παραμένει ανεπηρέαστη από παράγοντες όπως θορυβώδη ή ελλιπή δεδομένα.

Για να επωφεληθούμε από την έννοια της ΠΚ, αναπαριστούμε το ανθρώπινο σώμα ως ένα μοντέλο βασισμένο σε κυλίνδρους, αποτελούμενο από 11 μέρη. Το σώμα αντιμετωπίζεται ομοιόμορφα μέσα από το πλαίσιο της ΠΚ, ακολουθώντας μια τεχνική τοπικής βελτιστοποίησης. Τα μέρη, τα οποία αναπαρίστανται ως κύλινδροι, εξετάζονται σε μια από-πάνω-προς-τα-κάτω σειριακή δομή, εκκινώντας από το κεφάλι. Για κάθε μέρος του σώματος, δημιουργείται ένα σύνολο από υποθέσεις πόζας, το οποίο και παρακολουθείται στο χρόνο από ένα Φίλτρο Σωματιδίων (Particle Filter). Για να αξιολογήσουμε κάθε υπόθεση, χρησιμοποιούμε μια καινοτόμα μετρική η οποία λαμβάνει υπόψη την εικονική Κάτοψη του αντίστοιχου μέρους σώματος. Αυτό, σε συνδυασμό με την πληροφορία βάθους, μπορεί να αντιμετωπίσει επιτυχώς δύσκολες και αμφισβητήσιμες περιπτώσεις, όπως αυτές που προκύπτουν από έντονες επικαλύψεις, είτε μεταξύ μερών του ίδιου χρήστη, είτε μεταξύ διαφορετικών χρηστών.

Για την αξιολόγηση της εν λόγω μεθοδολογίας, διεξήχθησαν εκτεταμένα πειράματα τα οποία στοχεύουν σε ρεαλιστικά σενάρια, με αύξοντα βαθμό δυσκολίας και με κυμαινόμενο πλήθος χρηστών να αλληλεπιδρούν μεταξύ τους. Επιπροσθέτως συγκρίναμε την απόδοση της προτεινόμενης μεθόδου με την απόδοση των επικρατέστερων σύγχρονων μεθόδων, χρησιμοποιώντας δεδομένα διαθέσιμα από τον παγκόσμιο ιστό και δικά μας δεδομένα τα οποία εμπεριέχουν πληροφορία για την πραγματική πόζα των χρηστών. Τόσο τα ποιοτικά όσο και τα ποσοτικά αποτελέσματα που παρουσιάζονται πιστοποιούν την αποδοτικότητα της μεθόδου μας.

Επόπτης Διδακτορικής Διατριβής: Καθηγητής Παναγιώτης Τραχανιάς

ABSTRACT

Marker-less articulated human body pose recovery and tracking is a challenging problem of great importance, with strong theoretical and practical implications. The recent introduction of low-cost depth cameras triggered a number of interesting new works, pushing forward the state of the art. However, despite the remarkable progress, estimating the body pose in realistic, complex scenarios is still an open research task.

In this thesis we propose and develop a markerless model-based method to recover and track the full body pose, from RGB-D sequences, in arbitrary scenarios where users can freely enter or leave the scene, move, act and interact with other users or the environment. Our research focuses mainly on the problem of handling occlusions, either across body parts belonging to the same user, or across different users. At the same time, we attempt to tackle additional important issues encountered in the problem at hand, such as dealing with the large diversity of human bodies or the unconstrained initialization of tracking.

Towards this goal, we introduced the novel concept of Top View Reprojection (TVR) of cylindrical objects, which uniquely defines the pose of a cylinder based on certain quantitative appearance properties of its Top View, i.e. the view aligned with the cylinder's main axis. Based on this, the problem of estimating the pose of a cylindrical object becomes that of estimating the corresponding Top View. Interestingly, the developed formulation of TVR remains unaffected from factors such as noisy or missing data.

Capitalizing on the TVR concept, we represent the human body by a cylinder-based model, consisting of 11 body parts. The body is uniformly treated within the TVR framework following a local optimization technique; body parts, represented as cylinders, are examined in a top-to-bottom sequential order, starting from the head. For each body part a set of hypotheses is generated and tracked over time by a Particle Filter (PF). To evaluate each hypothesis, we employ a novel metric that considers the virtual Top View of the corresponding body part. The latter, in conjunction with regular depth information, effectively copes with difficult and ambiguous cases, such as severe inter- and intra-person occlusions.

For evaluation purposes, we conducted several series of experiments addressing realistic scenarios of gradually increased difficulty, involving varying number of users interacting with each other. We further compared the performance of the proposed method against that of state-of-the-art approaches using public or own-collected datasets with ground truth annotation. The presented quantitative and qualitative results attest for the effectiveness of our approach.

Supervisor: Professor Panagiotis Trahanias