

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ

ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ

ΠΑΡΟΥΣΙΑΣΗ / ΕΞΕΤΑΣΗ ΜΕΤΑΠΤΥΧΙΑΚΗΣ ΕΡΓΑΣΙΑΣ

Γαλανάκης Γεώργιος

Μεταπτυχιακός Φοιτητής

Τμήμα Επιστήμης Υπολογιστών, Πανεπιστήμιο Κρήτης

Επόπτης Μεταπτ. Εργασίας: Καθηγητής, Κ. Στεφανίδης

Τετάρτη, 4 Μαρτίου 2015, 11:00

Αίθουσα «Σ. Ορφανουδάκης», Ινστιτούτο Πληροφορικής, ΙΤΕ

Παρακολούθηση, επανα-ταύτιση και αλληλεπίδραση με χρήση αισθητήρων RGB-D σε περιβάλλοντα Διάχυτης Νοημοσύνης

Περίληψη

Τα τελευταία χρόνια, η εμφάνιση των αισθητήρων RGB-D (χρώμα, βάθος) έχει ενισχύσει προσεγγίσεις Υπολογιστικής Όρασης που υποστηρίζουν φυσική αλληλεπίδραση με μη παρεμβατικούς τρόπους. Τέτοιοι τύποι αλληλεπίδρασης είναι χρήσιμοι σε περιβάλλοντα Διάχυτης Νοημοσύνης (AmI), σκοπός των οποίων είναι να κρύψουν την παρουσία της τεχνολογίας από τους χρήστες, είτε να την ενσωματώσουν με ομαλό τρόπο στο περιβάλλον. Σε αυτό το πλαίσιο, η παρούσα εργασία εξετάζει τα προβλήματα της παρακολούθησης ατόμων (person tracking), επανα-ταύτισης ατόμων (person re-identification) και αναγνώρισης νευμάτων του κεφαλιού.

Στην παρούσα εργασία, το σύστημα παρακολούθησης ατόμων χρησιμοποιεί την είσοδο από βίντεο προκειμένου να ανιχνεύσει και να παρακολουθήσει άτομα στη διάρκεια του χρόνου. Το σύστημα παρακολούθησης χρησιμοποιεί πολλαπλές κάμερες εφόσον διευκολύνουν την κάλυψη ευρύτερων περιοχών· επιπλέον δεν απαιτείται κάποια εξεζητημένη μέθοδος προκειμένου να αντιμετωπιστούν οι επικαλύψεις. Συνηθισμένες προκλήσεις εμφανίζονται όταν πολλαπλά άτομα βρίσκονται στον ίδιο χώρο, έρχονται κοντά είτε αλληλεπιδρούν. Το προτεινόμενο σύστημα παρακολούθησης πολλαπλών ατόμων λαμβάνει είσοδο από πολλαπλούς αισθητήρες RGB-D οι οποίοι μοιράζονται ένα κοινό οπτικό πεδίο, κατασκευάζει

μια τρισδιάστατη αναπαράσταση της σκηνής και εν συνεχεία ανιχνεύει και παρακολουθεί άτομα. Προκειμένου να διευθετηθούν οι πιθανές αμφιβολίες κατά τη διάρκεια της ανάθεσης τροχιών, το σύστημα συνδυάζει την πρόβλεψη θέσης με χαρακτηριστικά της εμφάνισης. Η εμφάνιση ενός ατόμου απεικονίζεται σε μια τρισδιάστατη αναπαράσταση η οποία αποτυπώνει την χωρική διάταξη των χρωμάτων. Το προτεινόμενο σύστημα αξιολογείται μέσω σύνθετων συνόλων δεδομένων και συγκρίνεται με δύο προηγούμενες προσεγγίσεις παρακολούθησης.

Το εύρος κάλυψης του συστήματος παρακολούθησης ατόμων περιορίζεται σε μία σκηνή, επομένως μια επέκταση της λειτουργικότητάς του είναι η επανα-ταύτιση ενός ατόμου που επισκέπτεται πολλαπλές σκηνές. Η παρούσα εργασία παρουσιάζει μια αρχιτεκτονική η οποία συνδυάζει πληροφορία από πολλαπλά συστήματα παρακολούθησης ατόμων· κάθε σύστημα παρακολουθεί μια ξεχωριστή σκηνή. Το κύριο συστατικό της είναι ένα σύστημα επανα-ταύτισης το οποίο διατηρεί μια βάση δεδομένων. Η βάση δεδομένων αποτελείται από τρισδιάστατες αναπαραστάσεις οι οποίες διατηρούνται για ένα άτομο κατά τη διάρκεια της παρακολούθησης. Μόλις ένα άτομο εισέλθει σε μια σκηνή, τεχνικές βασισμένες σε χαρακτηριστικά της εμφάνισης εφαρμόζονται, προκειμένου να αντιστοιχηθεί σε μια εγγραφή της βάσης δεδομένων. Για να μετρηθεί η ακρίβεια του συστήματος έγινε μια αξιολόγηση βασισμένη σε ήδη αποθηκευμένα δεδομένα. Τα αναφερόμενα αποτελέσματα δείχνουν επαρκή ακρίβεια, η οποία παρ' όλα αυτά ελαττώνεται όταν οι αναπαραστάσεις της εμφάνισης λαμβάνονται κάτω από σημαντικά διαφορετικές συνθήκες φωτισμού.

Η παρούσα εργασία παρουσιάζει επίσης ένα σύστημα το οποίο αναγνωρίζει νεύματα του κεφαλιού, αναλύοντας την κίνησή του. Ένα σύστημα που εκτιμά τη στάση του κεφαλιού, βασισμένο σε δεδομένα βάθους, χρησιμοποιείται για το σκοπό αυτό. Για τον χαρακτηρισμό της κίνησης του κεφαλιού, διεξήχθησαν προκαταρκτικά πειράματα, υποδεικνύοντας ιδιότητες και ανατομικά όρια της κίνησης. Η προτεινόμενη προσέγγιση ανιχνεύει ένα σύνολο από απλά νεύματα. Η μέθοδος αξιολογείται στο πλαίσιο ενός διαλόγου ανθρώπου-υπολογιστή, αποδίδοντας ανταγωνιστικά αποτελέσματα αναγνώρισης σε σχέση με σύγχρονες εναλλακτικές μεθόδους.

Η αξιολόγηση των προαναφερθέντων συστημάτων υποδήλωσε επιτυχή υιοθέτηση των RGB-Δαισθητήρων για την διευθέτηση των διαφορετικών προβλημάτων για την παρακολούθηση ατόμων, επανα-ταύτιση ατόμων και αναγνώριση νευμάτων του κεφαλιού. Πιθανές μελλοντικές προεκτάσεις σχετίζονται με την ενσωμάτωση των παραπάνω μεθόδων σε καθημερινά περιβάλλοντα.

Galanakis Georgios

M.Sc. Thesis

Computer Science Department

University of Crete

Master's Thesis Supervisor: Professor Constantine Stephanides

Wednesday ,4/03/2015, 11:00

Room "S. Orfanoudakis, ITE

Tracking, re-identification and interaction using RGB-D sensors in Ambient Intelligence environments

Abstract

In recent years, the emergence of RGB-D sensors has reinforced the formulation of Computer Vision approaches which support natural interaction in unobtrusive ways. Such types of interaction are useful in Ambient Intelligence environments, whose purpose is to either entirely hide the presence of technology from users or to smoothly integrate technology within the environment. In this context, this thesis addresses the problems of person tracking, person re-identification and recognition of head gestures.

In this thesis, the person tracker is a system which analyzes video input to detect and monitor the positions of multiple persons over time. The tracker uses multiple cameras as they facilitate coverage of wider areas; moreover, a sophisticated method to cope with occlusions is not required. Common challenges appear when multiple persons are in the same place, come close or even interact with each other. The proposed multi-person tracking system receives input from multiple RGB-D sensors which share a common field of view, constructs a 3D representation of the scene and, then, detects and tracks persons. To address possible ambiguities during the assignment of tracks, the system combines position prediction and appearance characteristics. A person's appearance is projected in a 3D representation that captures the spatial arrangement of colors. The proposed system is evaluated using complex datasets and compared against two former tracking approaches.

The coverage of the person tracker is limited to a single scene, thus an expansion of its functionality is the re-identification of a person who visits multiple scenes. This thesis presents an architecture that combines information from multiple person trackers; each one monitors a distinct scene. Its main component is a re-identification system which maintains a database. This database consists of the 3D representations of persons which are maintained during the tracking. As soon as a person enters the scene, techniques based on appearance characteristics are applied in order to match with a database entry. To measure the accuracy of the system, an evaluation based on offline data was performed. Reported results indicate adequate accuracy, which however decreases when the appearance representations are obtained under significantly different lighting conditions.

This thesis also presents a system which recognizes head gestures by analyzing a person's head motion. To this purpose, a head pose estimator, based on depth data, is utilized. To characterize the head motion, preliminary experiments were conducted, indicating properties and anatomical limits of this motion. The proposed approach detects a set of primitive

gestures. The method is evaluated within the context of a human-computer dialog, yielding competitive recognition results to state-of-the-art approaches.

The evaluation to the aforementioned systems indicated successful adoption of RGB-D sensors to address the pursued problems of person tracking, person re-identification and recognition of head gestures. Discussion upon future work concerns the integration of the above approaches in everyday environments.