ΕΛΛΗΝΙΚΟ ΜΕΣΟΓΕΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ

ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ

ΠΡΟΣΚΛΗΣΗ ΣΕ ΔΗΜΟΣΙΑ ΠΑΡΟΥΣΙΑΣΗ

ΜΕΤΑΠΤΥΧΙΑΚΗΣ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ

Τίτλος

«Αναγνώριση Ακολουθιών Πόζας Σώματος για Αλληλεπίδραση Ανθρώπου με Ρομπότ»

του **Γερμενή Ευάγγελου,** μεταπτυχιακού φοιτητή του

Διατμηματικού Προγράμματος Μεταπτυχιακών Σπουδών
"Προηγμένα Συστήματα Παραγωγής, Αυτοματισμού και Ρομποτικής"

Επιβλέπων: Δρ. ΜΑΚΡΗΣ ΑΛΕΞΑΝΔΡΟΣ (ΙΠ-ΙΤΕ)

Η παρουσίαση θα πραγματοποιηθεί την **Τετάρτη 14 Δεκεμβρίου 2022 στις 13:30** μέσω της ηλεκτρονικής πλατφόρμας Google Meet, η οποία επιτρέπει την εξ' αποστάσεως παρακολούθηση της διαδικασίας από κάθε ενδιαφερόμενο μέλος της Ακαδημαϊκής Κοινότητας του Πανεπιστημίου.

Σύνδεσμος παρακολούθησης της παρουσίασης μέσω Google Meet:
https://meet.google.com/uor-xbnb-dwy

ΠΕΡΙΛΗΨΗ ΤΗΣ ΕΡΓΑΣΙΑΣ

Πρόσφατα, η έννοια της αλληλεπίδρασης ανθρώπου-ρομπότ (Human Robot Collaboration) έχει κερδίσει το ερευνητικό ενδιαφέρον. Η συνεργασία ανθρώπου-ρομπότ επιτρέπει στους ανθρώπους να εργάζονται παράλληλα με τα ρομπότ ως ομάδα, προκειμένου να επιτύχουν από κοινού αποτελεσματικά μια συγκεκριμένη δραστηριότητα.

Η αναγνώριση της ανθρώπινης δράσης-κίνησης παίζει κύριο ρόλο σε αυτή την προσπάθεια συνεργασίας μεταξύ των ανθρώπων και των μηχανών και χρησιμοποιείται ως η κύρια διεπαφή μεταξύ των δύο. Η αναγνώριση των ανθρώπινων ενεργειών-κινήσεων κατατάσσεται ως ένα δύσκολο ερευνητικά πρόβλημα λόγω της αρθρωτής γεωμετρίας του ανθρωπίνου σώματος και της εγγενούς πολυπλοκότητας που εμπεριέχει η ανθρώπινή κινησιολογία. Οι πρόσφατες εξελίξεις στην βαθιά μάθηση και στα νευρωνικά δίκτυα σε συνδυασμό με αλγορίθμους εκτίμησης ακολουθιών ανθρώπινης πόζας άνοιξαν το δρόμο για την ταχεία πρόοδο αυτού του ερευνητικού τομέα. Ένα από τα πιο δύσκολα βήματα στην αναγνώριση ακολουθιών πόζας σώματος είναι η εξαγωγή μιας συγκεκριμένης ανθρώπινης πόζας από δεδομένα βίντεο και η ταξινόμησή της.

Η παρούσα διατριβή προτείνει μια μέθοδο αναγνώρισης πόζας σώματος με χρήση δεδομένων εισόδου βίντεο RGB-D. Η μέθοδος βασίζεται σε ένα βαθύ νευρωνικό δίκτυο (OpenPose) για να πάρει μια αρχική δισδιάστατη (2D) θέση του σώματος. Έπειτα αφού χρησιμοποιηθεί συνδυαστικά η πληροφορία βάθους από τον αισθητήρα (κάμερα), υλοποιείται μια μέθοδος εκτίμησης της τρισδιάστατης (3D) ανθρώπινης πόζας για τη μετατροπή των 2D σημείων του σώματος σε 3D. Επιπλέον, εφαρμόζεται ένας αλγόριθμος δυναμικής χρονικής παραμόρφωσης (Dynamic Time Warping) σε συνδυασμό με έναν ταξινομητή πλησιέστερου γείτονα (nearest neigbhor) για την ταξινόμηση χρονοσειρών πόζας του ανθρώπινου σώματος.

Τα κύρια πλεονεκτήματα της προτεινόμενης μεθόδου είναι ότι στοχεύει προς την αναγνώριση ανθρωπίνων ενεργειών με χαμηλή χρονική υστέρηση για εφαρμογές πραγματικού χρόνου (Human Robot Collaboration) και ευελιξία λόγω της προσέγγισης μάθησης που δεν απαιτεί πολλά παραδείγματα εκπαίδευσης (One-Shot learning). Η κύρια ιδέα της μεθόδου βασίζεται στο να επωφεληθεί από τα καλά αποτελέσματα του αλγορίθμου OpenPose που βασίζεται στη βαθιά μάθηση, ενώ ταυτόχρονα αποφεύγεται το απαιτητικό σε δεδομένα έργο της εκπαίδευσης ενός νευρωνικού δικτύου.

Η απόδοση ταξινόμησης αυτής της μεθόδου εξετάζεται σε μια δημοσία βάση δεδομένων University of Berkley-MHAD και σε μια βάση δεδομένων που δημιουργήθηκε για το σκοπό αυτής της διατριβής στο Εργαστήριο Υπολογιστικής Όρασης και Ρομποτικής (CVRL) του ΙΠ-ΙΤΕ.