

**ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ**

**ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ**

**ΠΑΡΟΥΣΙΑΣΗ / ΕΞΕΤΑΣΗ ΜΕΤΑΠΤΥΧΙΑΚΗΣ ΕΡΓΑΣΙΑΣ**

**Ταβουλάρης Νικόλαος  
Μεταπτυχιακός Φοιτητής**

**Τμήμα Επιστήμης Υπολογιστών, Πανεπιστήμιο Κρήτης**

**Επόπτης Μεταπτυχιακής Εργασίας: Καθηγητής, Π. Τραχανιάς**

**Τετάρτη 1 Ιουλίου 2020 ,ώρα 14:00 μ.μ.**

**Τηλεδιάσκεψη ( μέσω του συστήματος e:Presence), Τμήμα Επιστήμης Υπολογιστών,  
Πανεπιστήμιο Κρήτης**

Διεύθυνση μετάδοσης (url): <http://video.ucnet.uoc.gr/live/show/282>

Κανάλι YouTube του Τμήματος

[https://www.youtube.com/channel/UC7uE3QiMTQjkrpByB\\_Gnt6Q/live](https://www.youtube.com/channel/UC7uE3QiMTQjkrpByB_Gnt6Q/live)

**“Ταυτόχρονος Προσδιορισμός της Πόζας και Χαρτογράφηση για Ανθρωποειδή  
Ρομπότ με χρήση Πυκνών Τεχνικών”**

**Περίληψη**

Ένα από τα δυσκολότερα προβλήματα της σύγχρονης ρομποτικής είναι το πρόβλημα του προσδιορισμού της πόζας του ρομπότ μαζί με την ταυτόχρονη χαρτογράφηση του περιβάλλοντός του (*Simultaneous Localization and Mapping - SLAM*). Το πρόβλημα αυτό συχνά αντιμετωπίζεται με χρήση οπτικών αισθητήρων όπως κάμερες βάθους ή χρώματος και αναφέρεται στην βιβλιογραφία ως ταυτόχρονος εντοπισμός της θέσης και χαρτογράφηση με οπτικά μέσα, *visual SLAM (vSLAM)*. Αντίστοιχα, ο αποκλειστικός εντοπισμός της πόζας ονομάζεται *Οπτική οδομετρία, Visual Odometry VO*. Οι αλγόριθμοι

vSLAM χωρίζονται κυρίως σε δύο ομάδες ανάλογα με το μέγεθος των δεδομένων που χρησιμοποιούνται: α) Στους πυκνούς αλγορίθμους, που χρησιμοποιούν μεγάλο όγκο από σημεία στο χώρο, και β) στους αραιούς αλγορίθμους που μειώνουν τον αριθμό των σημείων, επιλέγοντας μόνο κάποια χαρακτηριστικά σημεία του χώρου. Γενικά, το πρόβλημα του vSLAM είναι ιδιαίτερα δύσκολο λόγω υψηλού θορύβου στους αισθητήρες, θολότητα στις εικόνες εισόδου, σφάλματα στους προσεγγιστικούς αλγορίθμους, περιορισμένοι υπολογιστικοί πόροι κ.τ.λ. Το πρόβλημα όμως είναι ακόμα πιο δύσκολο όταν εφαρμόζεται σε ανθρωποειδή ρομπότ. Η δίποδη βάδιση των ανθρωποειδών ρομπότ προκαλεί ξαφνική επιτάχυνση της κίνησης της κάμερας μαζί με αυξημένη θολότητα στην εικόνα. Η παρούσα εργασία ξεκινάει με μια ανάλυση των υπάρχοντων συστημάτων vSLAM και των ιδιοτήτων τους. Έπειτα ακολουθεί μια σύγκριση καινοτόμων αλγορίθμων όπως το *3DMatch* και *3DSmoothNet*, οι οποίοι εξάγουν χαρακτηριστικά σημεία με χρήση νευρωνικών δικτύων. Οι αλγόριθμοι αυτοί μελετώνται σε σχέση με πιο συμβατικές τεχνικές όπως ο αλγόριθμος των *SIFT*. Εν συνεχεία παρουσιάζεται μια μέθοδος που επιτυγχάνει καλύτερη εκτίμηση της πόζας μαζί με πιο ακριβή χαρτογράφηση. Αυτό γίνεται με τον συνδυασμό δυο συστημάτων, του vSLAM συστήματος *Kinect Fusion* και του *State Estimation Robot Walking (SEROW)* που συνδυάζει οπτική οδομετρία με κινηματική. Επίσης παρέχεται μια χαμηλή σε υπολογιστικούς πόρους υλοποίηση που επιτυγχάνει εκτέλεση σε πραγματικό χρόνο σε ενσωματωμένες συσκευές όπως το *jetson tx2*. Τέλος μελετάται ο συνδυασμός πυκνών αλγορίθμων μαζί με αραιά χαρακτηριστικά σημεία σε μια ενοποιημένη αναπαράσταση γράφου.

**University of Crete**

**Computer Science Department**

**M.Sc. Thesis presentation / examination**

**Tavoularis Nikolaos**

**Master's Thesis Supervisor: Professor P. Trahanias**

**Wednesday, 1 July 2020, 14:00 p.m**

**Teleconference (will use the e: Presence system), Computer Science Department,  
University of Crete**

(url) : <http://video.ucnet.uoc.gr/live/show/282>

YouTube channel :

[https://www.youtube.com/channel/UC7uE3QiMTQjkrpByB\\_Gnt6Q/live](https://www.youtube.com/channel/UC7uE3QiMTQjkrpByB_Gnt6Q/live)

# “Visual Simultaneous Localization And Mapping for Humanoid Robots with Dense Techniques”

## Abstract

One of the hardest problems in modern robotics is the estimation of robot's pose while simultaneously creating the map of the environment, termed *Simultaneous Localization and Mapping (SLAM)*. This problem is often confronted with the use of visual sensors such as RGB and depth cameras and it is mentioned in the literature as *visual SLAM (vSLAM)*. In addition, solely determining the robot's pose is called *Visual Odometry (VO)*. vSLAM algorithms can be divided into two different groups, dense algorithms that have to deal with a huge amount of data and sparse algorithms that handle less data but they have to choose them meticulously. vSLAM is particularly hard due to sensor's noise, blurriness of the input image, approximations errors and limited resources. The problem is even harder when it is applied to humanoids. Humanoids bipedal movement, cause sudden acceleration of the camera's motion and extra blurriness in the image. This work starts with an evaluation of vSLAM systems and their properties. Then it follows an evaluation of novel feature extraction algorithms that use neural network in order to extract descriptors from 3D data. Algorithms such as *3DMatch* and *3DSmoothNet* were compared with more conventional methods such as *SIFT*. Next it proposes an algorithm that produces better pose estimation and map representation for humanoids. This is achieved with the combination of two systems, *Kinect Fusion SLAM* system and *State Estimation Robot Walking (SEROW)* which fuses VO with kinematics. Moreover it provides a lightweight implementation that runs on embedded devices such as *jetson tx2*. Finally this work studies the combination of dense methods with sparse features in a unified graph representation.