

Project Proposal Template

Project Management and Software Development for Medical Applications

General Info

Project Title: Deep learning based grasp estimation for artificial hands

Contact Person: Federico Tombari

Contact Email: tombari@in.tum.de

Project Abstract

Computer vision-based assistive technology solutions can revolutionize the quality of care for people with sensorimotor disorders. The goal of this work is to enable trans-radial amputees to use a simple, yet efficient, computer vision system to grasp and move common household objects with a myoelectric prosthetic hand. That is, the hand is able to see the object via an RGBD camera and output an appropriate grasp accordingly. A deep learning-based artificial vision system will be developed to augment the grasp functionality of a commercial prosthesis in a way that objects are classified with regards to the grasp pattern without explicitly identifying the object category. The goal of this project is to implement a software in which an appropriate grasp for a target object is proposed through the RGBD stream. The software outputs the segmentation of target object in a scene, the depth information and the proposed grasp which can later be used by the artificial hand.

Background and Motivation

Hand amputation hampers upper-limb amputees and people with congenital motor deficit significantly in doing their normal activities. Prosthetic hands can help these people in carrying out their daily activities to a great extent. Current commercial prosthetic hands are typically

controlled via the myoelectric signals, that is the electrical activity of muscles recorded from the skin surface of the stump. These prosthetic hands have several limitations in providing the acceptable dexterity for their users and as a result the control of these systems is still limited to one or two degrees of freedom.

As intermediate solutions, alternative modalities have been adopted to replace or augment the EMG signals, such as vision[1-7] and accelerometry signals[8,9]. Specifically, in the case of using computer vision, it was shown that object shapes can be quantised such that appropriate grasp types and sizes can be determined[2,3]. Benefiting from the flexibility that a deep learning structure offers, an inexpensive vision-based system suitable for use in artificial hands was developed in [1]. The only modality of augmentation of the hand prosthesis was color images. A single RGB image was taken and given to a CNN architecture in order to estimate a grasp type. However, this work lacked the relevant depth data which causes the system to be sensitive to the distance from the object. In addition, the segmentation of cluttered scenes in 2D can be problematic. To overcome these limitations, we are going to benefit from depth data to facilitate the segmentation and improve the grasp estimation.

Student's Tasks Description

The target for the student is to establish a (semi-) automatic software tool to estimate a grasp type for the object which the RGBD camera is targeting. To this end, still RGBD images of household objects will be captured. To be able to estimate a grasp type, a multimodal CNN [10,11] or a point cloud-based network such as PointNet [12] can be trained to classify four different types of grips for different objects.

In the second step – for which also a graphical



user interface is needed – the captured data is processed, segmented, visualized and given to the trained deep learning structure and an appropriate grasp type is estimated and presented. After this project, the student is able to understand and implement an algorithm presented in a recent publication in the field of deep learning and computer vision. Moreover, he student will learn about working with 3D images and RGBD sensors.

Technical Prerequisites

The student should have

- Basic knowledge of MATLAB / C++
- Basic skills in linear algebra
- Basic knowledge of image processing
- 2D/3D computer vision expertise
- Experience with GUI-Programming (MATLAB / Qt)
- Basic knowledge of machine learning

References

- [1] Ghazaei, G., Alameer, A., Degenaar, P., Morgan, G. and Nazarpour, K., 2017. Deep learning-based artificial vision for grasp classification in myoelectric hands. *Journal of Neural Engineering*, 14(3), p.036025.
- [2] Došen S and Popović D B 2011 Transradial prosthesis: artificial vision for control of prehension *Artif. Organs* **35** 37–48
- [3] Došen S, Cipriani C, Kostić M, Controzzi M, Carrozza M C and Popović D B 2010 Cognitive vision system for control of dexterous prosthetic hands: experimental evaluation *J. Neuroeng. Rehabil.* **7** 42
- [4] Marković M, Došen S, Cipriani C, Popović D B and Farina D 2014 Stereovision and augmented reality for closed-loop control of grasping in hand prostheses *J. Neural Eng.* **11** 046001
- [5] Krausz N E, Lenzi T and Hargrove L J 2015 Depth sensing for improved control of lower limb prostheses *IEEE Trans. Biomed. Eng.* **62** 2576–87
- [6] Marković M, Došen S, Popović D B, Graimann B and Farina D 2015 Sensor fusion and computer vision for context-aware control of a multi degree-of-freedom prosthesis *J. Neural Eng.* **12** 066022
- [7] Ghazaei G, Alameer A, Degenaar P, Morgan G and Nazarpour K 2015 An exploratory study on the use of convolutional neural networks for object grasp classification *The 2nd IET Int. Conf. on Intelligent Signal Processing* p 5
- [8] Atzori M, Gijsberts A, Castellini C, Caputo B, Hager A-G M, Elsig S, Giatsidis G, Bassetto F and

Müller H 2014 Electromyography data for non-invasive naturally-controlled robotic hand prostheses *Sci. Data* **1** 140053

[9] Krasoulis A, Vijayakumar S and Nazarpour K 2015 Evaluation of regression comparison for the continuous decoding of finger movement from surface EMG, accelerometry *IEEE Neural Engineering and Int. Conf.* pp 631–4

[10] Schwarz, M., Schulz, H. and Behnke, S., 2015, May. RGB-D object recognition and pose estimation based on pre-trained convolutional neural network features. In *Robotics and Automation (ICRA), 2015 IEEE International Conference on* (pp. 1329-1335). IEEE.

[11] Eitel, A., Springenberg, J.T., Spinello, L., Riedmiller, M. and Burgard, W., 2015, September. Multimodal deep learning for robust RGB-D object recognition. In *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on* (pp. 681-687). IEEE.

[12] Qi, C.R., Su, H., Mo, K. and Guibas, L.J., 2016. Pointnet: Deep learning on point sets for 3d classification and segmentation. *arXiv preprint arXiv:1612.00593*.