2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
September 24–28, 2017, Vancouver, BC, Canada

Human Upper-Body Pose Estimation using Fully Convolutional Network and Joint Heatmap

Jungmo Koo, Seunghee Lee, Hyungjin Kim, Kwangyik Jung, Taekjun Oh, *Student Member, IEEE*, and Hyun Myung, *Senior Member, IEEE*

***Abstract*— In this paper, we applied a fully convolutional deep learning algorithm to robustly estimate the human pose. Based on semantic segmentation using FCN (Fully Convolutional Network), we estimate human upper body joints. It can be seen that the algorithm works well for various objects being held in hand, and even when the joints are not visible.**

I. Introduction

Recently, human pose recognition has been applied to various fields. Applications that use human pose in areas such as robotics, virtual reality, and device control are being developed. In the past, various feature detection methods such as HOG [1], SURF [2], have been used to recognize human pose. However, pose detection based on human-defined features was not robust. So we propose an algorithm that can operate on various people and environments using a deep learning method. The ultimate goal of this work is to teach robots how to clean objects on the table. Therefore, using only the upper body information, the joint should be correctly estimated even if the user takes various objects. For this purpose, FCN (Fully Convolutional Network) [3] is utilized to train features of each human joint and to estimate its position.

II. METHODS

A. Training

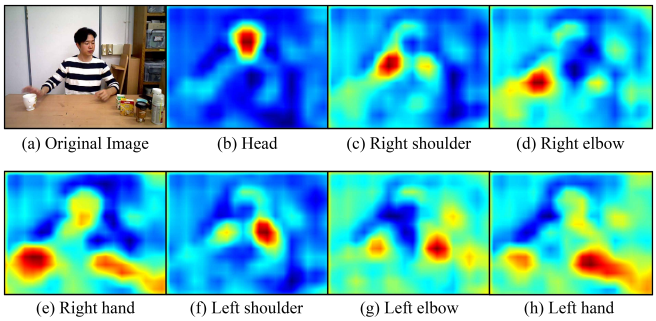
First, we used a scenario to move objects on the table to create a data set, and captured color images. Each joint of human was labeled to make data set. The network architecture FCN-AlexNet is used and fine-tuned using a pre-trained model by PASCAL-VOC data set. The solver used was Adam and the learning rate policy was set to sigmoid decay.

B. Detection

Based on the heat map of each joint as shown in Fig. 1, segmentation is performed by selecting the region with the highest confidence. After that, the central moment of the region is obtained and points of the joint are estimated.

III. RESULT AND FUTURE WORKS

As shown in Fig. 2, the proposed FCN is robust in various environments and poses. Also, as shown in Table I, the distance error to the ground truth is less than 7.3 cm. In the future, we will make more data sets and estimate the human hand pose for the robot's end-effector.



(a) Original Image (b) Head (c) Right shoulder (d) Right elbow

(e) Right hand (f) Left shoulder (g) Left elbow (h) Left hand

Fig. 1. Heat map of each joint. The higher the confidence, the closer to red. The lower the confidence, the closer to blue.

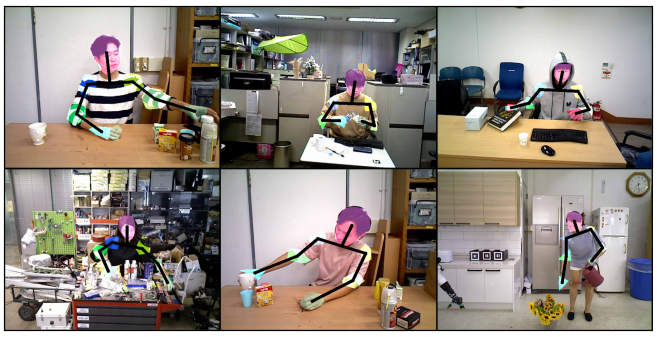


Fig. 2. Results of pose detection. Connection of 7 joints.

TABLE I. ERROR OF EACH JOINT

HEAD	*RIGHT SHOULDER*	*RIGHT ELBOW*	*RIGHT HAND*
6.35	43.15	56.43	50.05
LEFT SHOULDER	***LEFT ELBOW***	***LEFT HAND***	
33.11	72.09	55.57	

Unit : mm

References

[1] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection," 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR05).

[2] H. Bay, T. Tuytelaars, and L. V. Gool, "SURF: Speeded Up Robust Features," Computer Vision – ECCV 2006 Lecture Notes in Computer Science, pp. 404–417, 2006.

[3] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015.

*This work was supported in part by the Technology Innovation Program, 10045252, Development of robot task intelligence technology, supported by the Ministry of Trade, Industry, and Energy (MOTIE, Korea); and the students are supported by Korea Minister of Ministry of Land, Infrastructure and Transport (MOLIT) as U-City Master and Doctor Course Grant Program.

All authors are with the Urban Robotics Laboratory (URL), Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Korea; (Tel : +82-042-350-5670; E-mail: {jungmokoo, seunghee.lee, hjkim86, ankh88324, buljaga and hmyung}@kaist.ac.kr)