Sitting Posture Quality Evaluation Using Depth Camera

Jozef GÁBORÍK, Matej LEŠKO, Jakub MAČINA, Jozef STAŇO*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
imaginecup2015@googlegroups.com

It turns out, that nowadays people are sitting huge part of the day. Most of this time, they are sitting incorrectly. Sitting in a wrong posture on a long term basis causes *health problems* - ranging from pain in back, neck, shoulders to more serious disorders [1]. Impact is huge, according to WHO paper [2], the most common health problem is low back pain with the lifetime prevalence estimated 60-70% in industrialized countries.

We propose a *real-time posture tracking application* on desktops that notifies user in case of a wrong posture. It is important to alert users when they are sitting wrong, because people have tendency to forget about their posture when working with computers and started to slouch or lean forward to the screen.

We are using information from depth and RGB camera to evaluate user's posture while working with computer. To make our application works correctly it is necessary to set up a camera to capture an upper part of the user's body. The most important part of our application is *evaluation of user sitting posture quality*. Therefore, we developed three methods for extracting features *from depth and RGB images* to gain the highest accuracy of detecting quality of user posture.

A preprocess is required in order to extract features from the image. Firstly, we find all contours in this image. *Contour* is a curve along boundary of pixels with the same intensity or depth. We filter and process only contour with maximum area, because it is our *region of interest* (ROI) as it is very likely to be a user. In the next step, statistical properties called *moments* are calculated. These moments are very useful for describing objects represented by contours. From these image moments the center of the contour known as centroid is calculated.

The first method detects depths at different parts of a user's body. This *depth comparison* approach can recognize several posture states comparing the differences with the user calibrated image using specified threshold. To find points on a user's body which are used for comparison, we use similar technique like ray casting in a

* Supervisor: Jakub Šimko, Institute of Informatics and Software Engineering

Spring 2015 PeWe Workshop, April 11, 2015, pp. 13–14.

computer graphics. Given a started point and an angle, algorithm is moving in a specific angle and finds boundary of a region. Boundary of a region is set when difference from previous pixel depth exceeds a specified threshold.

Histogram of oriented gradients (HOG) [3] is a feature descriptor for an object classification commonly used in a computer vision. We are extracting HOG features for RGB images in a full size. For bounding box of user's upper body we use the same approach too. The upper body bounding box is found by haarcascade classifier. Moreover, we try this approach for depth images. For depth images, we can also crop region of interest from image and scale it to a constant image size of 320x240 pixels. This leads to smaller and constant feature vector and invariant to other contours in an image. In the both cases we trained *deep neural network* for a posture classification.

The last method computes other depth features from an image region of interest, which is the user's body. These features are then fed into neural network for posture classification. All values are depth and region size independent. We create an 11 dimensional *vector of features*, which are calculated from ROI bounding box, convex hull, fitting ellipse and centroid. Other features are mean depth, depth ratio at different points of the image and ratios from the histogram of depths.

In our experiment, we used RGB and depth photos of four women and six men while sitting in front of a computer. These photos are from variety of places, camera placements and participants are captured in a different posture. Experiments show that depth comparison method gained the best results with accuracy over 93%. Some of the methods that use neural networks are good, but they are depended on a quality of training samples. In the future, we would like to add other features in an image extraction, e.g. for detecting position of arms. We would like to validate our solution on bigger data set containing even more people with various body types and camera placements.

Extended version was published in Proc. of the 10th Student Research Conference in Informatics and Information Technologies (IIT.SRC 2014), STU Bratislava, 219-224.

Acknowledgement. This contribution was created with the support of the Ministry of Education, Science, Research and Sport of the Slovak Republic within the Research and Development Operational Programme for the project University Science Park of STU Bratislava, ITMS 26240220084, co-funded by the ERDF.

References

[1] Melrose, A.S., Graveling, R.A., Cowie, H., Ritchie, P., Hutchison, P., Mulholland, R.M.: Better Display Screen Equipment (DSE) work-related ill health data, 2007.

[2] Duthey, B.: Background Paper 6.24 Low back pain, 2013.

[3] Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on. Volume 1., IEEE, 2005, pp. 886–893.