# Real Time Extraction of Body Soft Biometric from 3D Videos

Carmelo Velardo, Jean-Luc Dugelay
EURECOM
2229 Route des Cretes
06560 Sophia Antipolis
{velardo,dugelay}@eurecom.fr

## ABSTRACT

In this technical demonstration, we show the application of our research on body soft biometrics. Exploiting a 3D video sensor we are able to extract semantic information that describes subjects standing in front of a camera. Semantic analysis and tracking is performed, series of anthropometric measures are extracted and used to compute subjects' height, weight, and gender information. Possible applications of such research fall in the medical domain for monitoring elderly people; in the gaming industry for automatic avatar creation, or in smart billboards which collects demographics of the public interested by the commercial. Our algorithm allows the estimation of all these parameters in real time without requiring the computational complexity of a 3D model fitting approach.

## Keywords

Soft biometrics, Kinect, user profiling, anthropometric measures

## 1. INTRODUCTION

The explosion in the controller-free market of the gaming industry led to the proliferation of smart cameras that are able to extract more than just a 2D video. Sensors like time of flight cameras and Microsoft Kinect [1] allow to perceive the world exploiting all the 4 dimensions. In our demonstration the Microsoft Kinect sensor, originally developed for gaming and natural interaction, will be exploited to fill the gap between biometric theory and practice. Using off the shelves components and this new cheap sensor, we are able to extract semantic information from one or more users standing in front of the camera (see fig. 1). We convert this information into a soft biometric signature that profiles the user. In our case we will be able to extract the height, weight, and gender information from the user thanks to the application of our theoretical study based on anthropometric measures [2, 3].

**Figure 1: The setup of our demo. A person is standing in front of the sensor triggering the estimation phase by assuming the calibration pose.**

In the following we will briefly describe the soft biometric characteristics that we are able to extract and the techniques that we developed to estimate the anthropometric measures from the 3D video.

## 2. BODY SOFT BIOMETRICS

Soft biometrics are traits that help to characterize people in predefined categories. Examples of soft biometric traits are the color of the hair or the eyes, the ethnicity, the stature, and the gait. They lack of the robustness which is necessary to discriminate among individuals. Nevertheless they can be extracted at a distance, do not require users' cooperation, and can enhance biometric system's capabilities. Furthermore, preliminary studies show potential in the identification application scenarios. We can identify three areas of interest: the soft biometrics related to the body, the ones extracted from the face, and the ones belonging to the accessories used in everyday life. To the first belongs the height, the weight, and the gender, that are the principal targets of this study. Using such characteristics human are able to classify body into three main *phenotypes* (Ectomorph, Endomorph, and Mesomorph) which help them during the recognition process. As example physical description is of utmost importance in many cases of missing persons or for crimes investigations. In this work we introduce an automatic system that, by elaborating a 3D video stream, is able to extract semantic information related to one or more subjects facing the sensor. More precisely we extract the information about the height, the weight, and the gender of the person. Hereafter we will present a description of the soft biometric traits analyzed in this study.

### 2.1 Height

During his extensive study on how to apply projective geometry to computer vision, Criminisi first explored the possibility of extracting geometrical information from 2D images. In [4] he gives details on how to reliably measure height of people standing in the field of view of a non calibrated camera. In our case we take advantage of the complexity of the Kinect capturing device that provides us with 3D information from the real world. The head and feet location of people facing the sensor are automatically extracted by our system that estimates their stature.

### 2.2 Weight

Extracting weight information from images is not straightforward. Being an image a projective representation of the real world, it lacks measuring capabilities (it's a projection) and proportions can be lost (due to camera distortions). In [2] we already formulated a possible solution to the problem. A series of anthropometric measures is chosen to be representative of human's body. Combining this research outcome with the Kinect capabilities, we are able to extract anthropometric measures from a 3D video and apply our body mass estimator for each person that trigger our system.

### 2.3 Gender

Phenotypes refers to the *observable* characteristics which are embedded into our DNA (Genotype) and that define our human physical aspect. In the literature of gender classification many works concentrate their efforts for facial appearance analysis, others use gait sequences. Following our results in weight estimation, we use a trained Neural Network to classify the gender starting from the same set of anthropometric measures exploited for weight estimation.

## 3. ANTHROPOMEASURES EXTRACTION

Capturing the output of the Kinect with the OpenNI framework [5], and exploiting OpenCV [6] and PCL [7] open source libraries, we are able to extract the silhouette of people moving in front of the 3D sensor. A background subtraction and a blob tracking algorithm automatically segment and track each single user. Once a user is extracted, his/her bounding volume is extracted thanks to the cloud of points coordinates. A simple sorting allow us to extract the information about the minima and the maxima in the 3D point cloud. To automatically measure height, we look for in the upper-central part of the bounding volume where presumably the head is located. Here we extract the 3D information for the outermost points of the silhouette. A similar approach is performed for the lower part were the feet are located. Height is then measured as distance between these two extrema (head and feet locations). In order to stabilize and filter out the outliers a moving median approach is used; this processing step is needed due to the noise produced around the edges of the silhouette by the 3D sensor.

We leverage the outcome of our previous theoretical work on weight estimation [2], to estimate the body mass from a series of anthropometric measures (height, arm and leg lengths, arm and leg circumferences, and waist circumference) extracted directly from the 3D silhouette. In order to locate limbs of each user, we exploit the skeleton tracker embedded in the OpenNI framework that provides us with the location of joints of a stick figure.

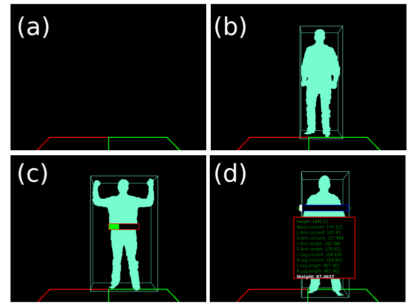


**Figure 2: Processing steps breakdown. (a) Floor is estimated and initialized, (b) the person is detected and his/her silhouette segmented, (c) calibration phase, (d) as mensuration ends the measured and the estimated values of height, weight, and gender are shown.**

The user has the possibility to trigger the automatic estimation assuming the calibration pose required by the skeleton tracker. Once the calibration pose is hold for few seconds our system immediately trigger the mensuration step which computes lengths and estimates circumferences of the limbs required by our system. To filter out along time the outliers that may affect all the anthropometric measures we adopted a similar solution to the one employed for the height estimation.

Once all the measures are extracted both gender and weight are estimated. The weight is directly shown by its value, while the gender is extracted as probability measure of being part of one of the two classes (Male/Female) and is shown as a sliding bar on the system's interface (see figure 2).

## 4. CONCLUSIONS

We present a demo that extracts semantic information for user in front of a Kinect camera. Possible applications are in the medical domain, in the gaming industry, and in automatic demographic collection. Guests of ACM Multimedia that will volunteer will be welcome to come and challenge our system's reliability. Only people assuming the required pose will be able to trigger the demo that will display their anthropometric measures, height, weight, and gender information.

## 5. REFERENCES


[1] Microsoft. http://www.xbox.com/kinect. [08-05-2011].

[2] C. Velardo and J.-L. Dugelay. Weight estimation from visual body appearance. In *BTAS'10, September*, 2010.

[3] A. Dantcheva, C. Velardo, A. D'Angelo, and J.-L. Dugelay. Bag of soft biometrics for person identification. *Multimedia Tools and Applications*, 2011.

[4] A. Criminisi, I. Reid, and A. Zisserman. Single view metrology. Springer, 2000.

[5] PrimeSense Inc. http://www.openni.org/. [08-05-2011].

[6] Willow Garage. http://opencv.willowgarage.com/. [08-05-2011].

[7] Radu Bogdan Rusu and Steve Cousins. 3d is here: Point cloud library (pcl). In *IEEE International Conference on Robotics and Automation (ICRA)*, 2011.