

# Human Gender and Motion Analysis with Ellipsoid and Logistic Regression Method

Md Israfil Ansari<sup>1,\*</sup>, Jaechang Shim<sup>2</sup>

## Abstract

This paper is concerned with the effective and efficient identification of the gender and motion of humans. Tracking this nonverbal behavior is useful for providing clues about the interaction of different types of people and their exact motion. This system can also be useful for security in different places or for monitoring patients in hospital and many more applications. Here we describe a novel method of determining identity using machine learning with Microsoft Kinect. This method minimizes the fitting or overlapping error between an ellipsoid based skeleton.

**Key Words:** Gender Detection, Motion Analysis, Ellipsoid Skeleton, Non-verbal Communication.

## I. INTRODUCTION

This paper is concerned with the effective and efficient categorization of humans in various fields and filtering them for the purpose of data mining and personal service. This will help us to know the various positions and actions of humans which can be implemented in a wide range of fields, e.g., for advertisement, security etc. Mostly two types of motion tracking are used, the marker based methods and marker-less methods. The marker based method [1] is accurate but it is expensive. The marker-less method [2] is cheap compared to the marker based because it uses multi-view depth cameras. Here we are going to use the marker-less method with Microsoft Kinect V2. The marker-less method is cheap but its speed is quite slow when it comes to tracking a person in fast motion. Usually machine learning based methods face the pose tracking problem as a per-pixel labeling problem and it is solved by various probabilistic methods like Gaussian Process (GP)[3], random decision tree [4], Markov Random Fields (MRF), etc. But such methods fail when given fewer data sets. The main motivation for gender classification of

humans based on their skeletons is that the skeleton can't be changed compared to human face.

## II. IMPLEMENTATION

### 2.1. Using Kinect to detect Human Gesture

A number of studies have been done to analyze body language or gesture of humans using automated systems. Chaira, Michele, virgin and Marco (2014) used GANT (Gaze Analysis Technique for human identification) to show its potential use for gender and age (younger or older than 30 years) categorization.

The Kinect sensor is a horizontal bar connected to a small base with a motorized pivot and is designed to be positioned lengthwise above or below the video display. The device features an "RGB camera, depth sensor and multi-array microphone running proprietary software",

The Microsoft Kinect V2 [5] uses an infrared emitter and sensor to capture body movements by isolating the X, Y and Z coordinates of 25 nodes roughly representing the joints of the body as shown in Figure 1.

The Kinect gives a unique opportunity to study gestures. It is inexpensive compared to 3D or other automated systems. It is a portable and unobtrusive device and the Kinect V2

---

Manuscript received June 16, 20, 2016; Revised July 10, 2016; Accepted July 28, 2016. (ID No. JMIS-2016-0007)

Corresponding Author (\*): Israfil Ansari, Dept. of Computer Science, Andong National University, Republic of Korea, +82-54-820-5645, israfila3@hotmail.com.

<sup>1,2</sup>Department of Computer Science, Andong National University, Kyungbook, Andong, Republic of Korea,  
E-mail: [israfila3@hotmail.com](mailto:israfila3@hotmail.com), [jcshim@andong.ac.kr](mailto:jcshim@andong.ac.kr)

can capture movements of 6 people at once from a range of 4 to 12 feet.

An optimized algorithm runs under 5 ms per frame (200 frame per second). It works frame by frame across dramatically differing body shapes and size.

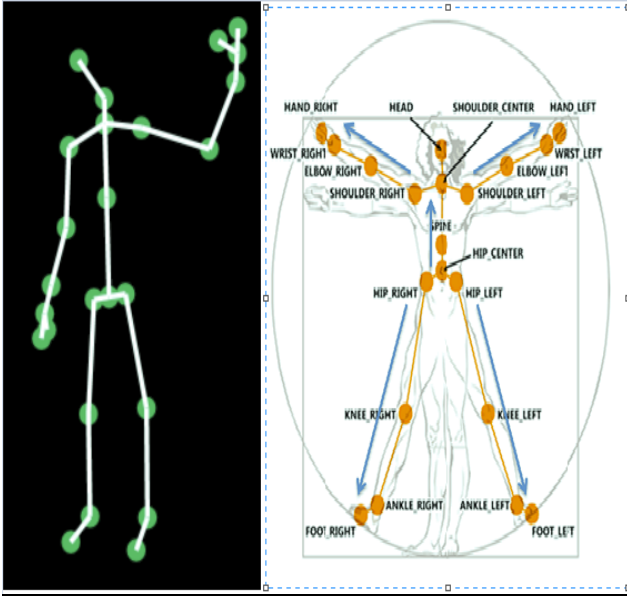


Fig. 1. Illustration of Kinect V2 data output in the form of a wireframe.

**2.2. Use of Ellipsoid method in this system**

An ellipsoid is a closed quadratic surface that is A three dimensional analogue of an ellipse. The standard equation of the ellipsoid method is

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

The Kinect gives an animated skeleton which can be defined as set of line segments and connection. Line segments are defined as bone and connections as the joints of the skeleton. The problem with line segments is that they overlap in different positions. To solve this issue, we introduce the ellipsoid method.

$$S = (B, j) \tag{1}$$

B = Collection of bones  
 J = collection of joints.

The equation of an arbitrary ellipsoid in Cartesian coordinate system is:

$$\| SR (x - p) \|^2 = 1. \tag{2}$$

Here x represents the variable of arbitrary point on an ellipsoid surface and p as the center of the ellipsoid. S and R as 3 \* 3 scaling and rotation matrix respectively. From Eq. (1) we can determine that an ellipsoid can be made by its center p with scaling matrix S and rotation

matrix R. Therefore, the collection of ellipsoidal bones can be represented as:

$$B = \{(p_j, R_j, S_j) \mid \text{for each bone } j\} \tag{3}$$

The bones in our ellipsoidal base skeleton will be connected through constraint vectors. The constraint vector is defined in the local coordinate system of the ellipsoidal bone, which is aligned with the three axes of the ellipsoid. If two bones, centered at p1 and p2 as in Figure 2, are connected at joint q, their constraint vectors v1 and v2 should point from p1 and p2 to q correspondingly.

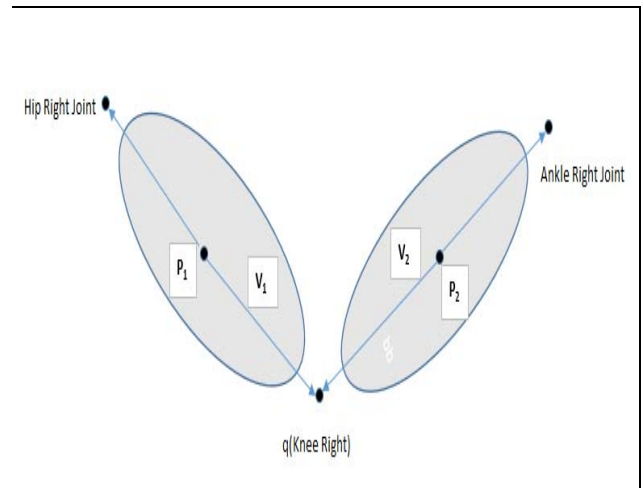


Fig. 2. Illustrate of Ellipsoid skeleton.

We can compare Figure 2 with particular part of Figure 1 from Hip Right joint to Ankle Right joint with Middle Knee Right joint. Here we are trying to make a gap between two joints which will remove the overlapping error and help us to track the actual movement of that particular area. This concept was carried out so that we can fit any shape at the tracked skeleton. Here shape means giving some meaning look to the tracked region which will be helpful to represent 3D dimension of particular person or in any other required dimension's.

**2.2. Gender and Motion Identification**

There is an enormous number of studies that have been done on human observers in real life to detect gender by facial recognition, voice tone, speech patterns and gestures. Among these gestures is the salient feature which helps to differentiate gender of same culture.

The Motion Capturing method provided in the Kinect SDK [16] is a typical machine learning based method. It infers the body part to which each depth image pixel belongs to a random decision forest trained with large and highly varied depth image sets. Traditionally Iterative Closest Point (ICP) was used for motion tracking, but due to its sensitivity to initial poses and proneness to local medium, Maximum A Posteriori (MAP)[6] was adopted.

However, beyond these culturally specific differences, there are differences in gesture and posture which can help

to distinguish men and women cross – culturally. The Kinect is similar to the point-light display (Johansson 1973) which consists of coordinates that indicate joint position; the Kinect also provides accurate information without the user wearing an obtrusive setup with lights. Here we propose machine learning for gender recognition with a “logistic regression” algorithm [7]. It helps to recognize two classes male and female and make a perfect filtration. Figure 3 illustrate the logistic regression filtration

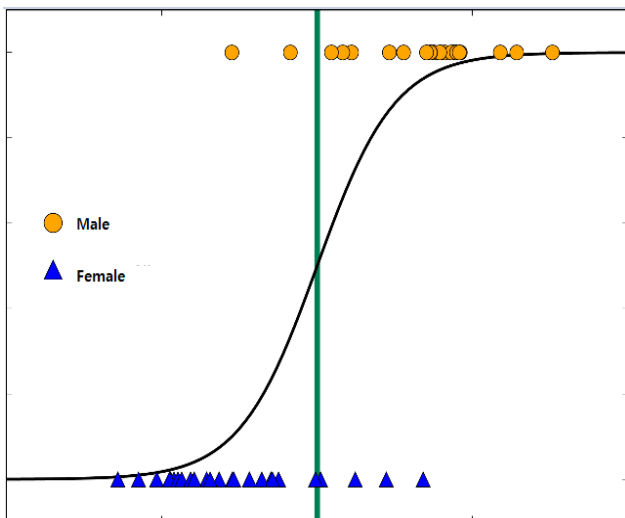


Fig. 3. Illustrate of logistic regression.

### III. Conclusion

In this paper, a novel feature descriptor specialized for human motion tracking and gender classification with the Microsoft Kinect V2 and logistic regression algorithm is introduced. The proposed descriptor is very much compact and more reliable compared to other systems. The main idea for this proposed system is to introduce compact and low price products which can be installed at any place.

In Mather and Murdoch’ 1994 [8] paper they state that gender classification of humans in the point-light display is dependent on dynamic cues of lateral body sway in shoulders and hips.

Our machine learning algorithm with the Microsoft Kinect V2 performs significantly in determining human gender with their various gestures. We hope that it will be useful in tracking nonverbal behaviors to facilitate the presence of humans and their social interaction.

This system can be helpful in security services in various organization, for e.g. in airport where a person’s motion will be tracked and this system can alert if any unwanted motion has occurred.

Another implementation can be possible i.e. human recognition with the help of skeleton tracking. We believe that this can be more effective as compared to other system because a person can change the outer part of the body but the skeleton can’t be changed.

In the future we would like to determine the age

along with gestures and gender of a human which will help us to understand deeply a particular person. In the current system we have focused only on human beings but we would like to interact with different creatures with this system.

### Acknowledgement

My deepest gratitude and warmest affection go to my advisor Prof. Shim Jaechang, for his generous advice, inspiring guidance, encouragement and professional assistance.

### References

- [1] A. J. Stoddart, P. Mrazek, D. Ewins, and D. Hynd “Marker based motion capture in biomedical applications,” *Motion Analysis and Tracking*, 1999.
- [2] Gait, “Recognition Based on Marker-less 3D Motion Capture,” *10th IEEE International Conference on Advanced Video and Signal Based Surveillance*, 2013.
- [3] C. Stoll, N. Hasler, J. Gall, H.-P. Seidel, and C. Theobalt, “Fast articulated motion tracking using a sums of gaussians body model,” in *Computer Vision (ICCV), 2011 IEEE International Conference on*, pp. 951-958, 2011.
- [4] K. Eguro , R. Bittner, and A. Forin “Random decision tree body part recognition using FPGAs” in *22nd International Conference on Field Programmable Logic and Applications (FPL)*, pp. 330-337, Aug 2012.
- [5] <https://developer.microsoft.com/en-us/windows/kinect/hardware>, June 2016.
- [6] G. Chantas, N. Galatsanos, and A. Likas “Maximum a posteriori image restoration based on a new directional continuous edge image prior” in *IEEE International Conference on Image Processing 2005*, pp. 941-944, Sep. 2005.
- [7] T. P. Minka “Algorithm for maximum-likelihood logistic regression”, 2003.
- [8] G. Marther and L. Murdoch “Gender discrimination in biological motion displays based on dynamics cues”, *Biological Science*, pp. 273-279, 1994.
- [9] Y. Zhu and K. Fujimura, “Constrained optimization for human pose estimation from depth sequences,” *ACCV 2007*, pp. 408-418, 2007.
- [10] R. Birdwhistell “kinesics and context”, 1970.
- [11] P. Zhang, K. Siu, J. Zhang, C. K. Liu, and J. Chai, “Leveraging depth cameras and wearable pressure sensors for full-body kinematics and dynamics capture,”

*ACM Trans. Graph.*, vol. 33, no. 6, pp. 221:1–221:14, 2014.

- [12] C. Stoll, N. Hasler, J. Gall, H.-P. Seidel, and C. Theobalt, “Fast articulated motion tracking using a sums of gaussians body model,” in *Computer Vision (ICCV), 2011 IEEE International Conference*, pp. 951–958, 2011.
- [13] S. Sridhar, A. Oulasvirta, and C. Theobalt, “Interactive markerless articulated hand motion tracking using RGB and depth data,” in *IEEE International Conference on Computer Vision*, pp. 2456–2463, 2013.

## Authors



**Md Israfil Ansari** received the BCA degree in Computer Application from Punjab Technical University, Jalandhar, India in 2011 and the MCA degree in Computer Application from Visvesvaraya Technological University, Bangalore, India in 2015. His previous position was a

Software Developer as an Intern at Xcelvations Consultancy Pvt Ltd, Hyderabad, India. Currently studies for Ph.D. at Department of Computer Engineering, Andong National University, Korea. He has interests in the areas of computer vision, pattern recognition, Deep Learning.



**Jaechang Shim** received the B. S. degree in 1987 and Ph.D. degree from Kyungpook National University, Korea, in 1997. From 1997 to 1999, he worked as a researcher in IBM T. J. Watson Research Center. From 2005 to 2006, he was a visiting Fellow Professor at Princeton University. Since 1994, he has been a Professor of the Andong National

University, Korea. He received the best Industrial-Related Paper Award form International Association for pattern Recognition in 1998. His research interests include computer vision, image processing, and pattern recognition.