

# Skeleton based gait recognition for long and baggy clothes

*Abrar Alharbi\**, *Fahad Alharbi*, and *Eiji Kamioka*

Graduate School of Engineering and Science, Shibaura Institute of Technology, Tokyo, Japan

**Abstract.** Human gait is a significant biometric feature used for the identification of people by their style of walking. Gait offers recognition from a distance at low resolution while requiring no user interaction. On the other hand, other biometrics are likely to require a certain level of interaction. In this paper, a human gait recognition method is presented to identify people who are wearing long baggy clothes like Thobe and Abaya. Microsoft Kinect sensor is used as a tool to establish a skeleton based gait database. The skeleton joint positions are obtained and used to create five different datasets. Each dataset contained different combination of joints to explore their effectiveness. An evaluation experiment was carried out with 20 walking subjects, each having 25 walking sequences in total. The results achieved good recognition rates up to 97%.

## 1 Introduction

The employment of modern technology to serve human security is considered one of the most essential services that should be provided to humanity with fewer errors. Due to the technological evolution, the security services are becoming better in accuracy and quality, which has increased its usage in securing critical places such as airports and stations. Security systems based on video cameras have been commonly used because of many advantages and large coverage domain. However, its results are affected by several factors such as lighting conditions, the angle of view, video quality, and the changes in the personal appearance. Therefore, biometric systems such as fingerprint, iris, voice, and facial recognition techniques have been widely used as powerful tools to identify and recognize a person [1]. Gait has been considered as a promising and unique biometric for person identification. Gait recognition is defined as automatic identification of an individual based on the style of walking. Gait recognition systems can be divided into either model-based or model-free approaches. Model-based systems generate the gait signature from modeling and tracking different parts of the body as segments or joint positions over time [2]. The model-free approach extracts gait features directly from the motion or shapes for recognition [3]. Implementing model-free approach is easier than model-based and less computationally expensive. However, the model approach free approach lacks the precision in identifying a subject's gait when silhouettes change due to poor visibility, varied carrying conditions, bulky clothing and occlusions [4]. The goal of

---

\* Corresponding author: [mg17002@shibaura-it.ac.jp](mailto:mg17002@shibaura-it.ac.jp)

this work is to design a method for identifying people who are wearing long and baggy clothes like Thobe and Abaya. Figure 1 shows the Abaya which is a traditional female cloth in Saudi Arabia as well as the Thobe which is the male version cloth. In this study, an experiment was conducted to find out which joints provide reliable information for recognizing the subjects while their body is obscured. The method used in the experiment will be described and the result also will be presented in this paper.



**Fig. 1.** An example of long and baggy is shown above. In (a) a traditional female cloth in Saudi Arabia (Abaya) and in (b) Traditional male cloth in Saudi Arabia (Thobe).

## 2 Related work

Many researchers have used Microsoft Kinect sensor for gait analysis and recognition in the past few years. For example, Ahmed *et al.* [5] presented a method for gait recognition from two sets of dynamic features. The first set is Horizontal Distance Features that is based on the distances between each pair of (Ankles, knees, hands, shoulders). The second set is Vertical Distance Features that provide significant information of human gait extracted from the height of (hand, shoulder, and ankles) to the ground. Their results reach 92% on their own database of 20 walking subjects with 10 walk cycles per subject on average. Dikovski *et al.* [6] evaluated a broad spectrum of geometric features and classifiers. Static body parameters like joint angles and inter-joint distances were aggregated and combined to construct seven different feature datasets. Multilayer perceptron, support vector machine with sequential minimal optimization and J48 algorithms were used for classification on these datasets to find out the role of different types of features and body parts in the recognition process. On their database of 15 walkers where each contributed with about 2 full gait cycles, the 89.8% recognition rate (10-fold cross-validation by MLP) was reached using a descriptor of 71 geometric features. Kumar & Babu [7] proposed an unrestricted gait recognition algorithm which uses trajectory covariance of joint points. The covariance matrices between these joint point trajectories formed the gait model. By computing the minimum dissimilarity measure between the gait models of the training data and the testing data the subject identity was classified. The recognition rate of 97.5% (2-fold cross-validation by 1-NN) is achieved for their dataset of 20 walking subjects, 10 gait samples each. Preis *et al.* [8] defined 13 biometric features, 11 static body features, and two dynamic features. Only four static features of them were sufficient to correctly identify a person with a recognition rate of 91% which was achieved by the Naïve Bayes classifier for their dataset of nine walking subjects. Balazia *et al.* [9] extracted 21 joint angles from 2 signature pose to form the gait pattern descriptor. Subject's identity was classified by the 1-NN classifier. The accuracy of 92.4% is reached on their database of 4188 walk samples of 48 people. Andersson *et al.* [10] used the skeleton joints to extract anthropometric

information and calculate lower joint (hips, knees, and ankles) angles as gait kinematic parameters. The mean and standard deviation of the body segment were also used as attributes. Walker identity was classified K-Nearest Neighbor. On their Kinect database of 160 individuals of about five walks each, the method achieves (10-fold cross-validation) 80% recognition rate. Faisal *et al.* [11] introduced a dynamic time warping (DTW) kernel based rank level fusion that takes a collection of joint relative distance (JRD) or joint relative angle (JRA) sequences as parameters and computed a dissimilarity measure between the training samples and the unknown samples. The JRD–JRA fusion achieves a recognition rate of 92 % on their database of 20 walking subjects with three walk cycles per subject.

### 3 Proposed method

The purpose of this study is to identify people who are wearing long baggy clothes like Thobe and Abaya using gait-based skeletal data from Kinect V2. The main difficulty is derived from extracting the human gait characteristic features to be used for identification, especially when the person body limbs are obscured, and the gait shape is hidden. The proposed gait recognition system is executed in three steps: First, building a database by recording skeleton information using the Kinect sensor. Second, extracting the gait features by detecting the skeleton information and obtaining the joint positions. Finally, training and recognizing using four different classifiers, Support vector machine (SVM), J48 decision tree algorithm, K-Nearest Neighbor (KNN) and Multilayer Perceptron (MLP) artificial neural network

#### 3.1 Data acquisition

There was not any existing database that delivers the skeleton information from Kinect sensor for people wearing long baggy clothes. Therefore, a database for this study was created through an experiment. The environment for recording the skeleton information is as follows: The Kinect sensor was placed parallel to the walking line at the height of 0.26 meters from the ground level and 3m away from the walking line. Twenty different people wearing either Thobe or Abaya joined the experiment. The recordings were done in a lab room at daytime. Each person walked 25 times from right to left in front of the Kinect sensor at their normal speed. All the skeleton information provided by the Kinect sensor were collected from each frame to create the database.

#### 3.2 Feature extraction

The Kinect sensor and its SDK provide a 3D skeleton model of the entire human body [12]. The skeleton consists of 25 joints. The estimated skeleton structure provided by the Kinect and joint names are shown in figure 2. The Kinect provides approximately 30 skeleton frames per second. For each frame, the joint position of each joint is expressed in the X, Y and Z coordinates. In this study, the skeleton joints position in XYZ-coordinates is extracted. However, only the Y-coordinate is used as input feature because the distance between the Kinect sensor and the walking line was fixed as well as the direction of walking which means that the X and Z coordinates are irrelevant. Twenty subjects joined the experiment; each subject had 25 walk sequences which resulted in 500 gait samples obtained for a side view. For each sequence, 32 frames were chosen based on the estimated time to generate one gait cycle on average. The Skeleton joints for each subject were

extracted in a column vector of 25 by 32. Recall that, thirty samples were allocated for each subject, which resulted in 800 by 500.

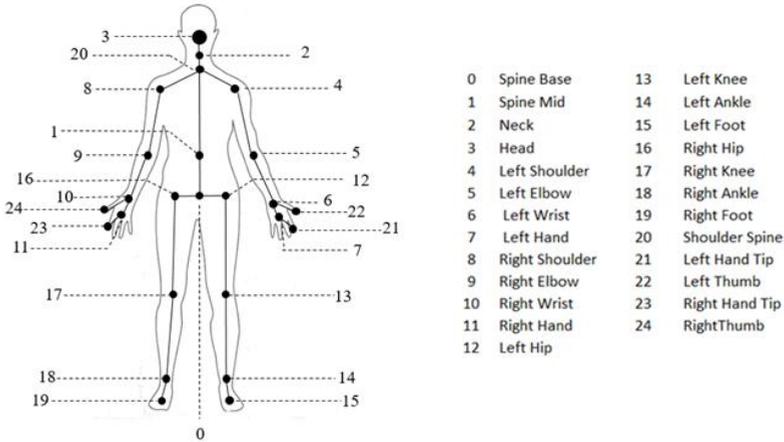


Fig. 2. The skeleton structure of a subject as provided by Microsoft Kinect SDK 2.0.

### 4 Experimental setup and design

From the skeleton information explained in the previous section, five different datasets were created. The first dataset contains all the skeleton joints positions per walking cycle. The second dataset contains the joint position of (Spine Mid, Neck, Head, Shoulders, Elbows, Wrists, Ankles, Feet, Shoulder Spine). These joints were chosen because they can represent the motion of the human body [13]. However, the hip and knee joints were not considered because their positions are difficult to detect. The third dataset is similar to the second one excluding the joints position on the right side of the body in order to focus on the side facing the Kinect. The fourth data set is constructed from the joints of the upper body part, and the final fifth dataset contained the joints of the lower body part. The final number of joints used in each dataset is shown in table 1.

Table 1. Different datasets used.

Dataset	Joint
1	All joint
2	1,2,3,4,5,6,8,9,10,14,15,18,19,20
3	1,2,3,4,5,6,14,15,20
4	1,2,3,4,5,6,7,8,9,10
5	0,12,13,14,15,16,17,18,19

### 5 Results and discussion

The performance of the proposed approach was evaluated using a self-constructed Kinect-based side view gait database. In the database, Kinect is used to extract twenty-five different walking sequences for twenty subjects. The achieved results using ten-fold cross-validation in the experiments for each method are shown in table 2. The datasets are listed in the first column of the table, while the remaining columns show the performance of each classifier for every dataset. The MLP and SVM classifiers gave similar results except for

the first dataset the SVM achieved better results than the MLP due to its ability to work with high dimensional data. The second and the third datasets achieved the highest recognition rate which showed the effectiveness of the selected joints. However, the third dataset achieved better accuracy than the second one, because few right side joints were detected incorrectly. Finally, the fourth data set achieved better results than the fifth data set which indicated that upper body joints give more information than the lower body joints in the process of gait recognition for long and baggy clothes.

**Table 2.** Recognition rates from test combinations using ten-fold cross-validation.

Dataset	MLP	SVM	KNN	J48
1	82	90	89	81
2	95	95	93	89
3	97	97	94	83
4	92	91	89	80
5	88	85	84	73

Table 3 shows recognition rates obtained in the Comparison of the proposed method with some of the other studies of gait recognition based on Kinect sensors that have been discussed in the related work section using the Thobe and Abaya database.

**Table 3.** Comparison According to Recognition Rate.

Method	Recognition Rate
Ahmed [5]	64
Preis [8]	50
Proposed method	95

## 6 Conclusions and future work

In this paper, a human gait recognition method was introduced. The skeleton joints provided by the Kinect sensor were utilized to achieve proper recognition for the people who are wearing long and baggy clothes like Thobe and Abaya. The skeleton information of each person was detected and used to create five data sets. Based on the experiment on a database of 20 subjects and 25 walk sequences, a high recognition rate of 97% was achieved by using the joints of the spine, head, neck, left shoulder, left elbow, left ankle, and left foot. Therefore, the proposal confirmed the effectiveness of gait recognition for long and baggy clothes. In the future, the dataset will be expanded with more number of subjects and walking scenarios, and advanced techniques like deep learning will be used to get a better recognition accuracy.

## References

1. Xiao Q 2007 Technology review-biometrics-technology, application, challenge, and computational intelligence solutions *IEEE Computational Intelligence Magazine* vol 2 no 2 pp 5-25
2. Liu L, Jia W and Zhu Y 2009 Survey of gait recognition *Emerging Intelligent Computing Technology and Applications. With Aspects of Artificial Intelligence* pp 652-659

3. Han J and Bhanu B 2006 Individual recognition using gait energy image *IEEE transactions on pattern analysis and machine intelligence* vol **28** no 2 pp 316-22
4. Wang J, She M, Nahavandi S and Kouzani A 2010 A Review of Vision-Based Gait Recognition Methods for Human Identification *2010 International Conference on Digital Image Computing: Techniques and Applications* pp 320-7
5. Ahmed M, Al-Jawad N and Sabir A 2014 Gait recognition based on Kinect sensor *Proc. SPIE Real-Time Image and Video Processing* vol 9139 pp 1-10
6. Dikovski B Madjarov G and Gjorgjevikj D 2014 Evaluation of different feature sets for gait recognition using skeletal data from Kinect pp 1304 -1308
7. Kumar M S N and Babu R.V 2012 Human gait recognition using depth camera: a covariance-based approach pp 1-6
8. Preis J, et al., 2012 Gait recognition with Kinect *Int. Workshop on Kinect in Pervasive Computing*
9. Balazia M and Plataniotis K 2015 Human gait recognition from motion capture data in signature poses *IET Biometrics* vol **6** no 2 pp 129-37
10. Andersson V, Dutra R and Araujo R 2014 Anthropometric and human gait identification using skeleton data from Kinect sensor pp 60-1
11. Ahmed F, Paul P P and Gavrilova M 2015 DTW based Kernel and Rank Level Fusion for 3D Gait Recognition using Kinect vol 31 no 6 pp 915-24
12. Zhang Z 2012 Microsoft Kinect Sensor and Its Effect *IEEE MultiMedia* vol **19** no 2 pp 4-10
13. Johansson G 1973 Visual perception of biological motion and a model for its analysis *Perception Psychophys* vol 14 pp 201-211