

# Reconstruction of 3D Human Body Pose for Gait Recognition

Hee-Deok Yang and Seong-Whan Lee\*

Department of Computer Science and Engineering,  
Korea University, Anam-dong, Seongbuk-gu,  
Seoul 136-713, Korea  
{hdyang, swlee}@image.korea.ac.kr

**Abstract.** In this paper, we propose a novel method to reconstruct 3D human body pose for gait recognition from monocular image sequences based on top-down learning. Human body pose is represented by a linear combination of prototypes of 2D silhouette images and their corresponding 3D body models in terms of the position of a predetermined set of joints. With a 2D silhouette image, we can estimate optimal coefficients for a linear combination of prototypes of the 2D silhouette images by solving least square minimization. The 3D body model of the input silhouette image is obtained by applying the estimated coefficients to the corresponding 3D body model of prototypes. In the learning stage, the proposed method is hierarchically constructed by classifying the training data into several clusters recursively. Also, in the reconstructing stage, the proposed method hierarchically reconstructs 3D human body pose with a silhouette image. The experimental results show that our method can be efficient and effective to reconstruct 3D human body pose for gait recognition.

## 1 Introduction

There has been a growing interest in improving automatic person identification. Established methods of automatic person identification range from fingerprint to face, iris and gait recognition. Gait recognition is an attractive biometric for passive surveillance system. Unlike other biometrics, gait recognition can be measured at a distance.

Most of gait recognition systems consist of two main stages: a feature extraction stage and a recognition stage and features for gait recognition are extracted from silhouette image [2, 6, 9, 10]. Collins et al. [4] used silhouette corresponding to certain poses such as the double-support and mid-stance poses. Classification of person is achieved by comparing their silhouette sequence. He and Debrunner [6] computed a quantized vector of Hu moments from the subject's silhouette image and used them for recognition by using an HMM (Hidden Markov Model). In Murray's research [11], ankle rotation and spatial displacements were shown to have individual consistency in repeated. To extract features for gait recognition, a number of researchers have been developed for estimating and reconstructing 2D or 3D body pose [1, 3, 7, 12].

---

\* To whom all correspondence should be addressed.

In this paper, we propose a solution to the problem of reconstructing 3D human body pose for gait recognition using a hierarchical linear learning method. The proposed method is related to machine learning models [2, 6] that use a hierarchical method to reduce the complexity of the learning problem by splitting it into several simpler ones.

## 2 Top-Down Learning

### 2.1 3D Human Model

The 3D human modeling includes human body representation and kinematics. For human body representing, we build a 3D human model consists of body segments, joints, and a perfect coordination among them. It has 17 body parts with 17 joints and 37 DOF(Degree of Freedom) and has 5 additional joints which are end-effectors used to calculate angle of each body segment in inverse kinematics [1]. Our 3D human model and the corresponding tree structure are shown in Fig. 1.

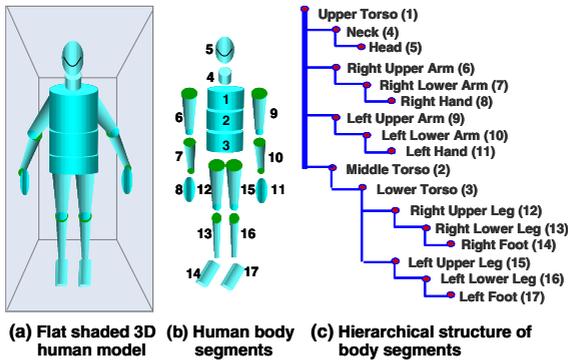
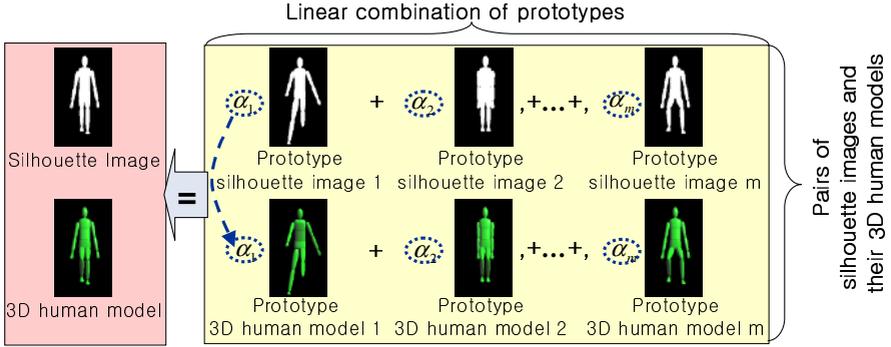


Fig. 1. The proposed 3D human model

### 2.2 3D Gesture Representation

In order to reconstruct 3D human body pose from continuous silhouette images, we used a learning based approach. If we have sufficiently large amount of pairs of a silhouette image and its 3D body model as prototypes of 3D body models, we can reconstruct an input 2D silhouette image by a linear combination of prototypes of 2D silhouette images. Then we can obtain its reconstructed 3D body model by applying the estimated coefficients to the corresponding 3D body model of prototypes as shown in Fig. 2. Our goal is to find an optimal parameter set  $\alpha$  which best reconstructs a given silhouette image. To make various prototypes of 2D silhouette images and theirs 3D body models, we generate data using the 3D human model described in Sec. 2. 1.



**Fig. 2.** Basic idea of the proposed method

The silhouette image is represented by a vector  $s = (s'_1, \dots, s'_n)^T$ , where  $n$  is the number of pixels in the image and  $s'$  is the intensity value of a pixel in the silhouette image. The 3D body model is represented by a vector  $p = ((x_1, y_1, z_1), \dots, (x_q, y_q, z_q))^T$ , where  $x, y$  and  $z$  are the position of body joint in the 3D world and  $q$  is the number of joints in 3D human model. Eq. (1) explains training data.

$$S = (s_1, \dots, s_m), P = (p_1, \dots, p_m) \tag{1}$$

where  $m$  is the number of prototypes.

A 2D silhouette image is represented by a linear combination of a number of prototypes of 2D silhouette images  $\tilde{S}$  and its 3D body model  $\tilde{P}$  represented by estimated coefficients to the corresponding 3D body model of prototypes by such as:

$$\tilde{S} = \sum_{i=1}^m \alpha_i s_i, \tilde{P} = \sum_{i=1}^m \alpha_i p_i \tag{2}$$

### 2.3 Hierarchical Statistical Model

In order to reduce search area, we construct our algorithm hierarchically. Given a set of silhouette images and their 3D body models for training, we classify them into several clusters. A set of cluster is built in which each has similar shape in 2D silhouette image space. Then, for each of the cluster, we divide it into several sub-clusters recursively. To divide training data into sub-clusters, we apply K-means algorithm.

The lower level is the mean value of the data in higher-cluster. In our model, the value of the first level is the mean value of each cluster in the second level, the values of the second and third level are the mean value of each cluster in the higher level respectively, and all leaf nodes are about 100,000 and each cluster in the third level has about 10~60 leaf nodes. Each cluster in the hierarchical model is presented by one gait representation described by Eq. (2).

### 2.4 Reconstruction of 3D Human Body Pose

To reconstruct 3D human body pose, we calculate the inverse matrix of  $S$  in Eq. (2). The inverse  $S^{-1}$  of a matrix  $S$  exists only if  $S$  is square. However, a matrix  $S$  is not square. In this case, we apply least square minimization. The pseudo inverse is defined such as:

$$S^+ = (S^T S)^{-1} S^T \tag{3}$$

And, the solution  $\alpha S = \tilde{S}$  can be rewritten such as  $\alpha = S^+ \tilde{S}$ .

After calculating  $\alpha S = \tilde{S}$ , we solve a set of coefficients of prototypes by Eq. (3). The silhouette image and the position of each segment of 3D human model calculated such as:

$$\tilde{S}_i = \sum_{k=1}^m \alpha_k s_k, \quad \tilde{P}_i = \sum_{k=1}^m \alpha_k p_k \tag{4}$$

## 3 Gait Recognition

For recognizing gait, we compute 5 angles as shown in Fig. 3. The trajectory of the above features used to recognize gait of subject.

$$F = (\theta_1, \theta_2, \dots, \theta_c) \tag{5}$$

where  $c$  is the number of angle used as features.

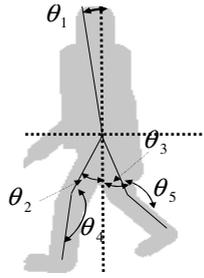
Given  $s$  subjects and each subject represents feature sequences of subject's gait. Let  $D_{i,j}$  be the  $j$ th  $F$  in class  $I$  and  $N_i$  the number of  $F$  in the  $i$ th class. We can easily obtain the mean  $m_d$  of full data set and the global covariance matrix such as:

$$m_d = \frac{1}{N_t} \sum_{i=1}^s \sum_{j=1}^{N_i} D_{i,j} \tag{6}$$

By applying PCA for training data, we can obtain eigenvalues  $\lambda_1, \lambda_2, \dots, \lambda_k$  and the associated eigenvectors  $e_1, e_2, \dots, e_k$ . We construct transform matrix  $E = [e_1, e_2, \dots, e_k]$  to project an  $D_{i,j}$  into a point  $P_{i,j}$  in the eigenspace. A gait sequence will be represented as a trajectory in the eigenspace. The projection average  $C_i$  of each training sequence in the eigenspace is given by

$$P_{i,j} = [e_1 \ e_2 \ \dots \ e_k]^T D_{i,j} \tag{7}$$

$$C_i = \frac{1}{N_t} \sum_{j=1}^{N_i} P_{i,j} \tag{8}$$



**Fig. 3.** The space domain features

## 4 Experimental Results and Analysis

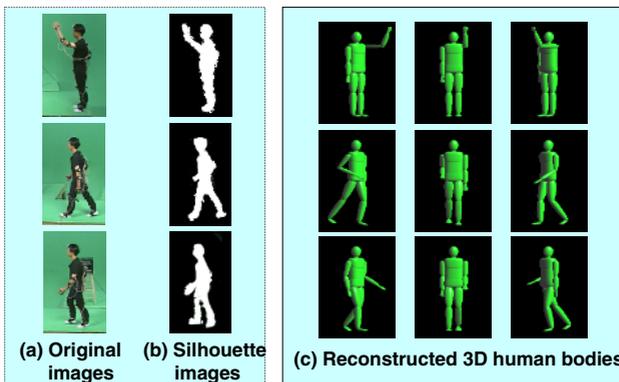
### 4.1 Experimental Results

For training the proposed method, we generated approximately 100,000 pairs of silhouette images and their 3D human models. We use a perspective projection transform to achieve silhouette and depth images.

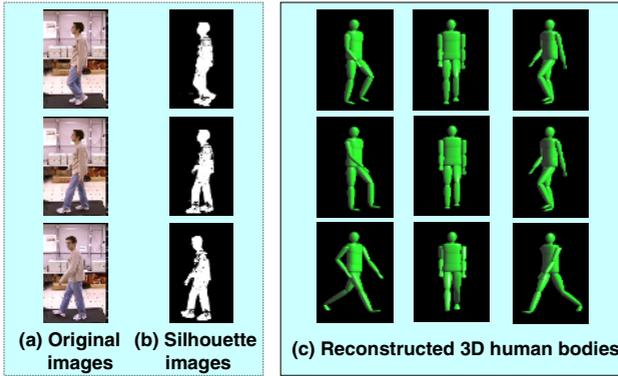
For testing the performance of our gait recognition method, we use two data sets. One is Georgia Tech Database [14] and the other is CMU Motion of Body(BoMo) database [5]. For testing the reconstruction performance of 3D body pose, we used the KU Gesture Database [8, 15].

### 4.2 Experimental Results

Fig. 4. shows the reconstructed results from the Georgia Tech. Fig. 4(a) and Fig. 4(b) are source image and silhouette image respectively. Fig. 4(c) represents silhouette image, left side, front, right side view of reconstructed 3D human body respectively. Fig. 5. shows the reconstructed results from the CMU Mobo database.



**Fig. 4.** Examples of the reconstructed 3D human body pose with the Georgia Tech database



**Fig. 5.** Examples of the reconstructed 3D human body pose with the CMU MoBo database

For recognizing subject, we select the normalized Euclidean distance measure. The accumulated distance between the test sequence  $P(t)$  and projection averages  $C(i)$  obtained such as:

$$d(i) = \sum_{t=1}^T \left\| \frac{P(t)}{\|P(t)\|} - \frac{C(i)}{\|C(i)\|} \right\|^2 \tag{9}$$

The classification result is accomplished by choosing the minimum of  $d(i)$ . The sequences are divided into two training and testing sets, by the ratio 4:1. The result of four experiments is shown in Table 1.

**Table 1.** Performance for CMU MoBo database in terms of PI at rank 1, 2, 5, 10

Train vs Test	PI (%) (at rank)			
	1	2	5	10
Slow vs Slow	60	65	80	100
Fast vs Fast	55	60	75	90
Incline vs Incline	50	50	60	85
Slow vs Fast	30	40	55	75

## 5 Conclusion and Further Research

In this paper, we proposed an efficient method to reconstruct 3D human body pose for gait recognition from monocular image sequence using top-down learning. Human body pose is represented by a linear combination of prototypes of 2D silhouette images and their corresponding 3D body models in terms of the position of a predetermined set of joints. With the 2D silhouette images and their corresponding 3D body models, we can estimate optimal coefficients for a linear combination of prototypes of the 2D silhouette images and their corresponding 3D body models by solving least square minimization.

The performance of the presented method shows that reconstructing 3D human body pose from visual features obtained from a single image is possible and the extracted feature by model 3D human body model is effective for gait recognition.

## Acknowledgements

This research was supported by the Intelligent Robotics Development Program, one of the 21st Century Frontier R&D Programs funded by the Ministry of Commerce, Industry and Energy of Korea.

## References

1. Benabdalkader et al.: Gait Recognition using Image Self-Similarity. *EURASIO Journal on Applied Signal Processing*, Vol. 4, (2004) 1-14
2. Bissacco, A. et al.: Recognition of Human Gaits. *Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Hawaii, USA, (Dec 2001) 52-57
3. Bowden, R., Mitchell, T. A., Sarhadi, M.: Non-linear Statistical Models for 3D Reconstruction of Human Pose and Motion from Monocular Image Sequences. *Image and Vision Computing*, Vol. 18, No. 9, (2000) 729-737
4. Collins, R., Gross, R., Shi, J.: Silhouette-based Human Identification from Body Shape and Gait. *Proc. of the IEEE International Conference on Automatic Face and Gesture Recognition*, Washington D.C., USA, (May 2002) 351-356
5. Gross, R., Shi, J.: The CMU Motion of Body(MOBO) Database. Technical report CMU-RI-TR-01-18, Robotics Institute, Carnegie Mellon University, (June 2001)
6. He, Q., Debrunner, C.: Individual Recognition from Periodic Activity using Hidden Markov Models. *Proc. of the IEEE Workshop on Human Motion*, Austin Texas, USA, (Dec. 2000) 47-527
7. Heap, T., Hogg, D.: Improving Specificity in PDMs using a Hierarchical Approach. *Proc. of 8th British Machine Vision Conference*, Colchester, UK, (Sep. 1997) 590-599
8. Hwang, B.-W., Kim, S. Lee, S.-W. "Full-Body Gesture Database for Analyzing Daily Human Gestures," *Proc. of 1st Int'l Conf. on Intelligent Computing*, Hefei, China, (Aug. 2005) 611-620.
9. Kale, A. et al.: Identification of Humans using Gait. *IEEE Trans. on Image Processing*, Vol. 13, (2004) 1163-1173
10. Little, J., Boyd, J.: Recognizing People by Their Gait: the Shape of Motion. *Videre*, Vol. 1, No. 2, (1998) 1-32
11. Murray, M.: Gait as a Total Pattern of Movement. *American Journal of Physical Medicine*, Vol. 46, No. 1, (1967) 290-332
12. Rosales, R., Sclaroff, S.: Specialized Mapping and the Estimation of Human Body Pose from a Single Image. *Proc. of IEEE Workshop on Human Motion*, Texas, USA (Dec. 2000) 19-24
13. Yam, C. et al.: Automated Person Recognition by Walking and Running via Model-based Approaches. *Pattern Recognition*, Vol. 37, No. 5, (2004) 1057-1072
14. The Gait Recognition Database at Georgia Tech, <http://www.cc.gatech.edu/cpl/projects/hid/images.html>.
15. The KU Gesture Database, <http://gesturedb.korea.ac.kr/>.