

GAIT IDENTIFICATION USING THE 3D PROTRUSION TRANSFORM

Dimosthenis Ioannidis, Dimitrios Tzovaras and Konstantinos Moustakas

e-mail: {djoannid, tzovaras, moustak }@iti.gr

Informatics and Telematics Institute/CERTH, P.O. Box 361, 57001 Thessaloniki, Greece

ABSTRACT

The present paper presents a novel approach for gait identification using 3D data and Krawtchouk moments to generate the descriptor feature vectors. The gait sequence is captured by a stereoscopic camera and the resulting 2.5D data are processed to generate a 3D hull of the captured silhouette. The *3D Protrusion Transform* is then proposed that generates a silhouette image containing protrusion information. Finally, the descriptor vector of the extended silhouette is calculated using the Krawtchouk moments. Experimental evaluation illustrates that the proposed scheme is highly efficient in identifying gait sequences when compared to state of the art approaches.

Index Terms— Gait recognition, geodesic distance, Krawtchouk moments, protrusion map.

1. INTRODUCTION

During the latest years there has been a growing interest in the identification of the humans based on their way of walking. The results of this research field can be directly applied for surveillance, identity verification and in medical applications as well. Most of the recent gait analysis methods can be divided into two main categories; model based and feature based methods.

Model-based approaches, study static and dynamic body parameters of the human locomotion. In the literature many different shapes have been used to assemble the human body. In [1], a combination of different shapes is used for resembling the human body. In [2], a gait recognition method has been proposed based on Procrustes shape analysis. Furthermore, there are model-based approaches that compute gait dynamics parameters from gait sequences like stride length, stride speed and cadence [3].

On the other hand, feature based techniques do not rely on the assumption of any specific model of the human body for gait analysis. Initially, the binary map of the moving subjects is computed and a feature vector is extracted from the silhouette sequences using several techniques. Noise removal and normalization procedures follow so as to generate, scaling, translation and in some cases rotation invariant descriptors. In the final stage, the feature-based approaches use a matching method for finding the similarity

between two input gait sequences. The methods proposed in the past include simple temporal correlation, full volumetric correlation on partitioned subsequent silhouette frames [4], linear time normalization [5] and dynamic time warping [6]. For the calculation of the distance between the gait feature vectors different techniques can be applied. In most cases, Euclidean distance was used as a metric for distance calculation, but there are also reports on using procrustes distance [2] and symmetric group distances [7].

The present paper proposes a novel gait identification and authentication method based on the use of novel 2D and 3D features of the image silhouette sequence based on the weighted Krawtchouk moments that are well known for their compactness and discriminating power. The 3D data that are obtained by a stereo camera are transformed using the novel 3D Protrusion Transform (3D-PT). The proposed algorithm is tested and evaluated in two datasets and was also compared to the state-of-the-art methods in gait analysis and recognition.

2. PRE-PROCESSING OF GAIT SEQUENCES

2.1. Binary silhouette extraction

In order to analyze the human movement, the walking subject silhouette needs to be extracted from the input image sequence.

Initially, the background is estimated using a temporal median filter on the image sequence, assuming static background and moving foreground. Next, the silhouettes are extracted by comparing each frame of the sequence with the background. The areas where the difference of their intensity from the background image is larger than a predefined threshold are considered as silhouette areas. The generated silhouette images are noisy. Therefore morphological filtering, based on anti-extensive connected operators [8] is applied so as to denoise the silhouette sequences. Finally, potential shadows are removed by analyzing the sequence in the HSV color space [9].

2.2. Generating 3D geodesic silhouettes

Using the aforementioned techniques a binary silhouette sequence \tilde{B}^{sil} is generated as illustrated in Figure 1a. In the proposed framework 2.5D information is available since the gait sequence is captured by a stereoscopic camera. Using

Delaunay triangulation on the 2.5D data a 3D triangulated hull of the silhouette is generated that is further processed using the proposed 3D Protrusion Transform.

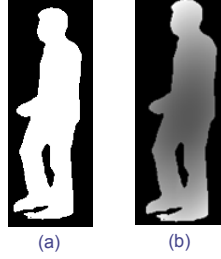


Figure 1. Extracted silhouettes, a) binary silhouette, b) geodesic silhouette

Initially, the triangulated version of the 3D silhouette is generated. Adjacent pixels of the silhouette are grouped into triangles. Next, the dual graph $G = (V, E)$ of the given mesh is generated [10], where V and E are the dual vertices and edges. A dual vertex is the center of mass of a triangle and a dual edge links two adjacent triangles. The degree of protrusion for each dual vertex results from equation (1):

$$p(\mathbf{u}) = \sum_{i=1}^N g(\mathbf{u}, \mathbf{v}_i) \cdot \text{area}(\mathbf{v}_i) \quad (1)$$

where $p(\mathbf{u})$ is the protrusion degree of dual vertex \mathbf{u} , $g(\mathbf{u}, \mathbf{v}_i)$ is the geodesic distance of \mathbf{u} from dual vertex \mathbf{v}_i and $\text{area}(\mathbf{v}_i)$ is the area of triangle that corresponds to the dual vertex \mathbf{v}_i .

Let us define $G_k^{rSil}(\mathbf{u})$ a function that refers to the dual vertices, to be given by:

$$G_k^{rSil}(\mathbf{u}) = p(\mathbf{u}) \cdot \tilde{B}_k^{Sil}(\mathbf{u}) \quad (2)$$

The 3D-PT for the silhouette image, denoted as $G_k^{Sil}(x, y)$, is simply a weighted average of the dual vertices that are adjacent to the corresponding pixel (x, y) , i.e.:

$$G_k^{Sil}(x, y) = \sum_{i=1}^8 G_k^{rSil}(\mathbf{u}) \cdot w(x, y, \mathbf{u}) \quad (3)$$

$$\hat{G}_k^{Sil}(x, y) = m + G_k^{Sil}(x, y) \cdot (255 - m) \quad (4)$$

where $i=1, \dots, 8$ denotes the number of adjacent pixels (x, y) to be weighted, $w(x, y, \mathbf{u})$ is the weighting function and $\hat{G}_k^{Sil}(x, y)$ represent the geodesic silhouette image at frame k , as illustrated in Figure 1b, which takes values in the interval of $[m, 255]$. The higher the intensity value of a pixel in Figure 1b, the higher its protrusion degree. In the proposed approach m was selected to be equal to 60.

2.3. Normalization

In the final step of the preprocessing stage and before feature extraction, the preprocessed binary (\tilde{B}^{Sil}) or 3D silhouette images (\hat{G}^{Sil}) are scaled and aligned to the center [5],[6]. In the proposed paper all silhouette gait sequences are normalized to the same resolution as in [4] and the silhouette is aligned to the center of the frame.

3. GAIT SEQUENCE FEATURE EXTRACTION

In the present paper, the use of descriptors based on the weighted Krawtchouk moments is proposed. In all cases, the input to the feature extraction system is assumed to be either the binary silhouettes (\tilde{B}_k^{Sil}) or the 3D-distributed silhouettes (\hat{G}_k^{Sil}) when the 3D-PT is used.

For almost all recent approaches on gait analysis, after feature extraction the original gait sequence cannot be reconstructed. In the suggested approach, the use of a new set of orthogonal moments is proposed based on the discrete classical weighted Krawtchouk polynomials [11]. The orthogonality of the proposed moments assures minimal information redundancy. In most cases, Krawtchouk transform is used to extract local features of images [11]. The Krawtchouk moments Q_{nm} of order $(n+m)$ are computed using the weighted Krawtchouk polynomials for a silhouette image (binary or 3D) with intensity function $Sil(x, y)$ by [11]:

$$Q_{nm} = \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} \bar{K}_n(x; p1, N-1) * \bar{K}_m(y; p2, M-1) \cdot Sil(x, y) \quad (5)$$

$$\bar{K}_n(x; p, N) = K_n(x, p, N) \sqrt{\frac{w(x, p, N)}{\rho(n, p, N)}} \quad (6)$$

where \bar{K}_n , \bar{K}_m are the weighted Krawtchouk polynomials, and $(N-1) \times (M-1)$ represents the pixel size of the silhouette image $Sil(x, y)$. A more detailed analysis of Krawtchouk moments and their computational complexity is presented in [11].

Krawtchouk moments can be used to extract local information of the images by varying the parameters N and M . Parameter N can be used to increase the extraction of silhouette image in the horizontal axis. Larger N provides more information on the silhouette image in the horizontal axis, whereas the parameter M extracts local information of the silhouette image in the vertical axis. For the experiments, values for $N=R/15$ and $M=C/3$ were used, where R and C denote the number of rows and columns of the silhouette image, respectively.

Krawtchouk transform is proposed for feature extraction, due to its very high discriminative power. Krawtchouk transformation is scale and rotation dependent. However, silhouette sequences are pre-scaled and aligned to the center, thus the Krawtchouk transform is unaffected by scaling. Furthermore, the input gait sequences are captured in a near fronto-parallel view and thus rotation does not affect the results of the Krawtchouk transform.

4. SIGNATURE MATCHING

The following notations are used in this section: the term gallery is used to refer to the set of reference sequences, whereas the test or unknown sequences to be verified or identified are termed probe sequence. In this paper, the gait cycle is detected using a similar approach to [6], using

autocorrelation of the input periodic signal. Instead of measuring only the sum of the foreground pixels in a temporal manner, the time series of the width of the silhouette sequence was also calculated. Then, the mean value of these signals formed the final gait period of the current gait sequence.

Each probe sequence is initially partitioned into several full gait cycle segments and the distance between each segment and the gallery sequence is computed separately. This approach can be considered as a brute-force attempt to match a pattern of segmented feature vectors (segmentation using gait cycle) by shifting them over a gallery sequence vector. The main purpose of this shifting is to find the minimum distance (or maximum similarity) between the probe and the gallery sequence.

Let $\mathbf{F}_{P,T}, \mathbf{F}_{G,T}$ represent the feature vectors of the probe with N_P frames and the gallery sequence with N_G frames respectively, and T denote the Krawtchouk transform. The probe sequence is partitioned into consecutive subsequences of T_p adjacent frames, where T_p is the estimated period of the probe sequence. Also, let the k -th probe subsequence be denoted as $\mathbf{F}_{(P,T)}^k = \{\mathbf{F}_{P,T}^{kT_p}, \dots, \mathbf{F}_{P,T}^{(k+1)T_p}\}$ and the gallery sequence of N_G frames be denoted as $\mathbf{F}_{G,T} = \{\mathbf{F}_{G,T}^1, \dots, \mathbf{F}_{G,T}^{N_G}\}$. Then, the distance metric between the k -th subsequence, and the gallery sequence:

$$Dist_T(k) = \min_l \sum_{i=0}^{T_p-1} \sqrt{\sum_{x=0}^{S-1} (\mathbf{F}_{P,T}^{i+kT_p}(x) - \mathbf{F}_{G,T}^{i+l}(x))^2}, \quad k = 0, \dots, m-1 \quad (7)$$

where $l = 0, \dots, N_G - 1$, S denotes the size of a Probe/Gallery feature vector \mathbf{F} , and $m = N_P / T_p$ represents the number of probe subsequences.

The above equation (7) indirectly supposes that the probe and gallery sequences are aligned in phase. After computing all distances between probe segments and gallery sequences of feature vectors, the median, [4] of the distances is taken as the final distance $D_T(Probe, Gallery)$ between the probe and the gallery sequence.

$$D_T = Median(Dist_T(1), \dots, Dist_T(m)) \quad , \quad m = \frac{N_P}{T_p} \quad (8)$$

where m denotes the number of distances calculated between the probe subsequences and the whole gallery sequence. In the above equation (8), smaller distance means a closer match between the probe and the gallery sequence.

5. EXPERIMENTAL RESULTS

The proposed methods was evaluated on two different databases: a) the publicly available HumanID ‘‘Gait Challenge’’ dataset [16], and b) the proprietary large indoor HUMABIO dataset.

Since the HumanID ‘‘Gait Challenge’’ dataset includes only monoscopic image sequences it cannot be used to evaluate the proposed scheme using the 3D-PT. However,

the Krawtchouk descriptor efficiency on binary silhouettes was evaluated using this database, so as to generate comparative results with state-of-the-art approaches. In an identification scenario, a score vector for a given probe gait sequence is calculated, that contains the distance of the probe sequence from all the gallery sequences that exist in a database. The gallery sequence that exhibits the minimum distance from the probe sequence is identified as the correspondent sequence to the probe sequence.

In the USF’s Gait Challenge Database, the gallery sequences were used as the systems knowledge base and the Probe sequences as the ones that should be recognized by comparing their descriptors to the gallery set. The available gallery sequences include (C, G) cement or grass surface, (A, B) shoe type A or B and (L, R) two different view points. In the performed experiments we used the set GAR as the gallery. The probe set is defined using seven experiments A-G of increasing difficulty. Experiment A differs from the gallery only in terms of the view, B of shoe type, C of both shoe type and view, D of surface, E surface and shoe type, F of surface and viewpoint and G of all surface, shoe type and viewpoint.

For evaluation of the proposed approach, Cumulative Match Scores (CMS) are reported at ranks 1 and 5. Rank 1 performance illustrates the probability of correctly identifying subjects in the first place of the ranking score list and the rank 5 illustrates the percentage of correctly identifying subjects in one of the first five places.

Table 1 illustrates rank 1 and 5 results of the proposed approach on binary silhouettes (KR) compared to the approaches (CMU) [12], (LTN-A) [5], and (BASE) [4]. It is obvious that the proposed approach based on Krawtchouk moments performs better in almost all experiments.

Table 1. Comparative results for the Krawtchouk transform on binary silhouettes (the number of subjects in each set is reported in squared brackets)

Probe Set	Rank 1				Rank 1-5			
	KR	CMU	LTN-A	BASE	KR	CMU	LTN-A	BASE
A (GAL) [71]	96	87	89	79	100	100	99	96
B (GBR) [41]	85	81	71	66	93	90	81	81
C (GBL) [41]	76	66	56	56	89	83	78	76
D (CAR) [70]	30	21	21	29	63	59	50	61
E (CBR) [44]	27	19	26	24	66	50	57	55
F (CAL) [70]	20	27	15	30	49	53	35	46
G (CBL) [44]	21	23	10	10	48	43	33	33

Figure 2 illustrates detailed results on the identification rate of the proposed method on binary silhouettes when compared to the baseline algorithm [4] for two experiments. From the above results it can be concluded that the proposed method based on Krawtchouk moments outperforms the state-of-the-art approaches. It should be noted that until now all results refer to processing binary silhouettes in the ‘‘Gait Challenge’’ database.

Unfortunately the proposed 3D-PT approach cannot be directly compared to the state-of-the-art approaches since it reports also depth information. However, experimental results have been extracted for the 3D-PT in the HUMABIO database and are compared with the results that are extracted using the Krawtchouk descriptor without using the 3D information (i.e. on binary silhouettes), that outperform the state-of-the-art approach as aforementioned.

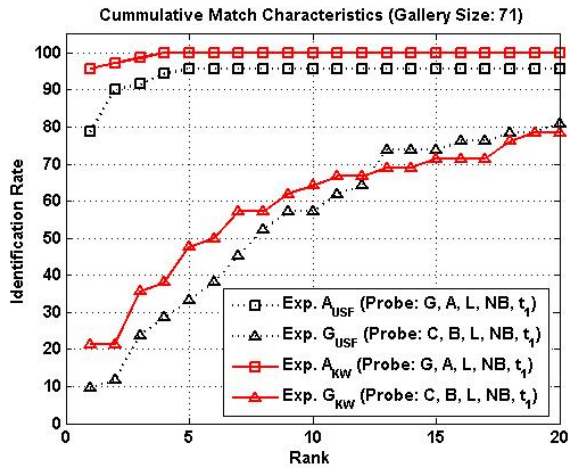


Figure 2. Identification rate of the proposed method on binary silhouettes (KW) for two experiments (A, G) compared to the baseline algorithm [4].

The proprietary HUMABIO gait database consists of sequences of 75 subjects. The available sequences include (C) normal surface, (CL, PA) shoe type classic or slipper, (BF, NB) carrying a briefcase or not and (H) when the subject wears a hat. In this paper two experiments on this database are demonstrated. The experiment A addresses the difference with hat and the experiment B addresses the briefcase difference between the gallery and the probe sequence. In all cases the gallery sequence is consisted of a classic shoe type (CL) with no briefcase (NB) and not wearing a hat (C-CL-NB).

Figure 3 illustrates detailed results on the identification rate of the 3D-PT when compared to the algorithm that uses the Krawtchouk descriptors on binary silhouettes. As illustrated an increased identification rate can be expected when using the 3D-PT.

6. CONCLUSIONS

In this paper a novel feature-based gait recognition framework was presented that uses the 2.5D information of the captured sequence captured by a stereo camera. This information is initially transformed into a 3D hull and then the 3D protrusion transform is proposed to generate the “geodesic” silhouette. The feature vector of the gait sequence is generated, by calculating the Krawtchouk moments of the new enhanced silhouette. Experimental results demonstrate the efficiency of the proposed method when compared to state of the art approaches.

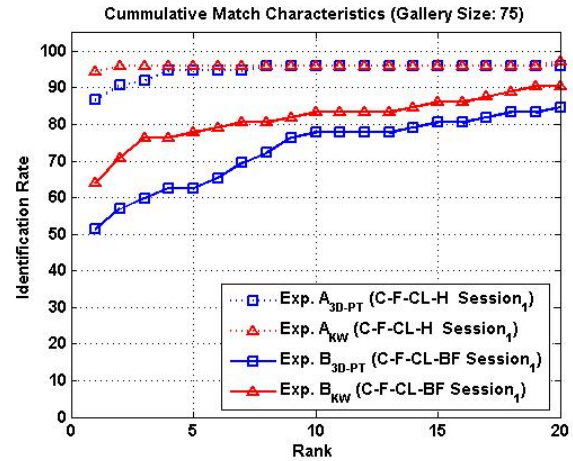


Figure 3. Identification rate of the 3D-PT method (3D-PT) for two experiments (A, B), compared to the algorithm that uses the Krawtchouk descriptors on binary silhouettes (KW).

7. REFERENCES

- [1] D.K. Wagg and M.S. Nixon, “On automated model-based extraction and analysis of gait,” in Proc. IEEE. Int. Conf. Automatic Face and Gesture Recognition, Seoul, Korea, May 2004, pp.11-16.
- [2] L. Wang, H. Ning, W. Hu and T. Tan. “Gait recognition based on procrustes shape analysis,” In International Conference on Image Processing (ICIP),” pages III: 433–436, 2002.
- [3] L. Wang, H. Ning, T. Tan and W. Hu, “Fusion of static and dynamic body biometrics for gait recognition,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol.14, no.2, pp. 149-158, Feb. 2004.
- [4] S. Sarkar, P.J. Phillips, Z. Liu, I. Robledo-Vega, P. Grother and K.W. Bowyer, “The Human ID Gait Challenge Problem: Data Sets, Performance, and Analysis,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 27, no. 2, pp. 162-177, Feb. 2005.
- [5] N.V. Boulgouris, K.N. Plataniotis and D. Hatzinakos, “Gait Recognition Using Linear Time Normalization,” *Pattern Recognition*, vol. 39, pp. 969-979, 2006.
- [6] N.V. Boulgouris, K.N. Plataniotis and D. Hatzinakos, “Gait recognition using dynamic time warping,” *2004 IEEE 6th Workshop on Multimedia Signal Processing*, vol., no.pp. 263- 266, 29 Sept.-1 Oct. 2004.
- [7] Y. Liu, R. Collins, and Y. Tsing. “Gait sequence analysis using frieze patterns,” In European Conference on Computer Vision, pp. 657–671, 2002.
- [8] P. Salembier and F. Marqués, “Region-based representations of image and video: Segmentation tools for multimedia services,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 9, no 8, pp. 1147-1169, Dec 1999.
- [9] R. Cucchiara, C. Grana, M. Piccardi, A. Prati and S. Sirotti, “Improving shadow suppression in moving object detection with HSV color information,” *IEEE Proceedings of the Intelligent Transportation Systems*, pp.334-339, 2001.
- [10] H.S. Lin, H.M. Liao, and J. Lin, “Visual Saliency-Guided Mesh Decomposition,” Proc. IEEE Int’l Workshop Multimedia Signal Processing (MMSP), 2004.
- [11] P.T. Yap, R. Paramesran and O. Seng-Huat, “Image analysis by Krawtchouk moments,” *IEEE Transactions on Image Processing*, vol.12, no.11pp. 1367- 1377, Nov. 2003.
- [12] R. Collins, R. Gross, and J. Shi, “Silhouette-based human identification from body shape and gait,” In IEEE Conf Automatic Face and Gesture Recognition, pages 351–356, 2002.