AN ANGULAR TRANSFORM OF GAIT SEQUENCES FOR GAIT ASSISTED RECOGNITION

Nikolaos V. Boulgouris, Konstantinos N. Plataniotis, and Dimitris Hatzinakos

The Edward S. Rogers Sr. Department of Electrical and Computer Engineering
University of Toronto, Ontario M5S 3G4, Canada
nikos@comm.toronto.edu

ABSTRACT

A new system is proposed for gait analysis and recognition applications. The new system is based on a denoising process and a new angular transform that are applied on binary silhouettes. Each human silhouette in a gait sequence is transformed into a low dimensional feature vector consisting of average pixel distances from the center of the silhouette. The sequence of feature vectors corresponding to a gait sequence is used for identification based on a minimum-distance criterion between test and reference sequences. By using the new system on the Gait Challenge database, improvements in recognition performance are seen in comparison to other methods of similar or higher complexity.

1. INTRODUCTION

The identification of humans based on their way of walking has recently emerged as a very attractive research area due to its applications on tele-surveillance and remote identification systems. Existing approaches for gait sequence analysis try to capture as much gait information as possible in order to use it for recognition purposes. Particularly, the techniques that deal with the gait recognition problem using only sequences of walking silhouettes are of much interest since they do not presume the availability of any further information, such as color or texture, which may not be available or extractable.

In [1], a baseline experiment was conducted aiming to serve as a reference experiment for gait recognition. A large set of gait sequences was divided into several disjoint sets. One set was kept aside as the system database set. The recognition decisions were taken by computing the similarity between the sequences in each of the remaining sets and the sequences in the database set. In [2], Hidden Markov Models (HMM) were used to train models for each gait sequence. Comparisons in the HMM parameter domain yielded improved recognition performance. In [3], identification using Principal Component Analysis (PCA) was performed using horizontal and vertical projections of silhouettes. In [4], a model-based approach was taken for improving the quality of extracted silhouettes. The silhouettes generated using the proposed technique were tested for gait recognition and the performance of the resulting scheme was shown to be superior in comparison to the baseline system in [1]. In [5], area-based metrics were used in order to derive a time-varying signal that was subsequently used for automatic gait recognition. Recently, an image analysis methodology which bears some similarity with the one proposed in this paper was presented [6]. In [6], a silhouette unwrapping technique was applied by calculating the distances of all contour pixels from the center of the silhouette. The collection of distances was used as a feature vector on which PCA was applied in order to reduce the dimensionality of the problem.

In this paper, we propose a novel methodology for the efficient processing of walking silhouettes which can be used in a gait recognition system. The proposed system bases its efficiency on a preprocessing of gait sequences and an angular transform which calculates a metric of the silhouette in angular slices of various orientations with respect to the center of the silhouette. In comparison to the method in [6], the present feature extraction technique has several advantages: it is robust to segmentation errors, it does not require detection of contour pixel positions, and it can be computed easily. By using the new system, the recognition performance on the Gait-challenge database [1] will be seen to improve over the methods in [1] and [6].

The structure of the paper is as follows: section 2 presents the preprocessing of gait sequences. In section 3, the angular transform for gait analysis is introduced. Section 4 describes the gait recognition system based on the new feature. Experimental results are presented in section 5 and finally, conclusions are drawn in section 6.

2. PREPROCESSING OF GAIT SEQUENCES

As in almost all recent approaches for gait recognition, we rely solely on binary silhouettes derived by background subtraction. However, the background subtraction process, which is essential in silhouette-based gait recognition, will usually be imperfect yielding inaccurate silhouettes. This is due to the fact that the colors of clothes, hair, or skin of a walking person may also exist in the background during capturing, and therefore, part of the background is likely to be misinterpreted as being part of the body during the background subtraction process [4]. Consequently, in most practical cases, there will be a need for denoising prior to the application of a gait recognition algorithm. In this work, we use a process of four steps for the denoising of gait sequences. These steps are detailed below:

- **Step 1:** ***Median filtering.*** Each silhouette is filtered using a 3×3 median filter in order to eliminate isolated errors.
- **Step 2:** ***Aligning.*** All silhouettes are aligned so that their centre is in the centre of the frame.

This work is supported by Bell University Labs at the University of Toronto.

- **Step 3:** ***Temporal processing.*** Temporal lines [7] are used for the correction of artifacts. A temporal line consists of all pixels lying in the same position of all frames in a sequence. We apply two simple rules for combatting pixel misclassification during background subtraction. Specifically,
  - if the number of background pixels on a temporal line is 90% or higher of the total number of pixels on the line, then all pixels are considered foreground pixels.
  - If the number of consecutive background pixels is above 25% of the total number of pixels on the line, then all pixels on the line are considered background pixels.
- **Step 4:** ***Averaging.*** Averaging is performed in order to deal with errors that cannot be corrected using the three previous steps i.e. persistent errors in moving areas of the silhouette. Averaging silhouettes was also used in order to denoise [4] or summarize a gait sequence into a number of characteristic templates [8, 9]. In the present work, the gait sequence is processed on a frame by frame basis by identifying the $N_d = 4$ most similar frames, i.e. the indices $l_1^t, \ldots, l_{N_d}^t$ of the frames with which the currently processed frame t has the smallest Euclidean distance. After the most similar frames are identified, the pixels $\hat{s}_t[i,j]$ in the tth frame are re-estimated as

$$\hat{s}_t[i,j] = \lfloor \frac{1}{N_d+1}(s_t[i,j] + \sum_{m=1}^{N_d} s_{l_m^t}[i,j]) \rfloor \quad (1)$$

  where (i,j) is the pixel position and $\lfloor \cdot \rfloor$ denotes rounding to the nearest integer (zero or one). The above recalculation of silhouettes using (1) is applied to the entire sequence of silhouettes in a gait sequence.

3. ANGULAR GAIT ANALYSIS

We assume that each gait sequence is composed of several binary silhouettes $s[i,j]$. Let

$$s[i,j] = \begin{cases} 1 & \text{if } (i,j) \text{ belongs to the foreground} \\ 0 & \text{otherwise} \end{cases}$$

In order to apply our approach, first the area center (i_c, j_c) of each silhouette is computed as:

$$i_c = \frac{1}{N}\sum_{i,j} i \cdot s[i,j], \quad j_c = \frac{1}{N}\sum_{i,j} j \cdot s[i,j]$$

where N is the number of foreground pixels, given by

$$N = \sum_{i,j} s[i,j]$$

Once the center of the silhouette is calculated, we define a new coordinate system $x-y$, whose origin is at the center of the silhouette. We propose an angular transform:

$$\mathcal{A}(\theta) = \frac{1}{N_\theta} \sum_{(x,y)\in\mathcal{F}_\theta} s[x,y]\sqrt{x^2+y^2} \quad (2)$$

where $\mathcal{F}_\theta$ is the set of the pixels in the circular sector $(\theta - \frac{\Delta\theta}{2}, \theta + \frac{\Delta\theta}{2})$ and N_θ is the number of pixels in $\mathcal{F}_\theta$. For each θ, the transform coefficient expresses the average distance of foreground pixels (in the direction defined by θ) from the center of the silhouette. In practice, since there is an infinite number of angles θ, the angular transform is computed in slices of $\Delta\theta$. The angle step $\Delta\theta$ determines the level of detail of the transform. The transform is graphically illustrated in fig. 1.

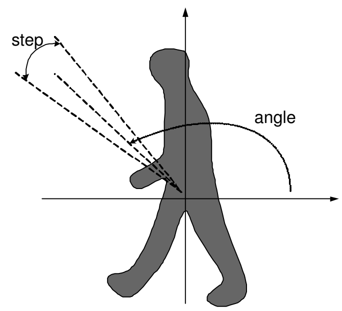


Fig. 1. Graphical representation of the proposed transform.

The proposed angular transform has a very convenient scaling property. Suppose that a silhouette is scaled by α in both horizontal and vertical dimensions. It can be shown that the transform coefficients $\tilde{\mathcal{A}}(\theta)$ of the scaled sequence are related to the transform coefficients of the original silhouettes as:

$$\tilde{\mathcal{A}}(\theta) = \alpha\mathcal{A}(\theta)$$

The above equation means that if the silhouette is scaled by α in each dimension, the transform $\tilde{\mathcal{A}}(\theta)$ of the scaled silhouette is equal to $\alpha\mathcal{A}(\theta)$. This property is useful since it allows scaling of the silhouettes in the transform domain by directly scaling the transform coefficients.

Although in the preprocessing stage, we aligned the silhouettes for the purpose of denoising, in general, the use of the proposed transform obviates the need to align the silhouettes to the center of the frame as most gait recognition methods demand. This is due to the fact that the proposed angular transform is inherently translation invariant, since it is always calculated with respect to the center of the silhouette.

Another desirable feature of the proposed algorithm is related to the fact that the averaging that takes place in each direction of θ is practically implicitly equivalent with a low-pass filtering procedure which endows the resulting transform with robustness to segmentation errors. This is very important since, in practice, the background subtraction process, that is usually performed automatically without human intervention, will always generate imperfect silhouettes.

4. APPLICATION ON GAIT RECOGNITION

The gait recognition system we implemented is based on the silhouette preprocessing and the angular silhouette representation described in the previous sections. All transform coefficients of a silhouette are ordered in a single gait vector of dimension $K = 360/\Delta\theta$ where the angle step $\Delta\theta$ denotes the width of the angle intervals in which the transform is computed. Simulations indicate that the recognition performance is largely independent of $\Delta\theta$ in the range $3 - 15$ degrees. The transform vector corresponding to a silhouette is of the form:

$$\mathbf{A} = [\mathcal{A}(\theta_0)\ \mathcal{A}(\theta_1)\ \ldots\ \mathcal{A}(\theta_{K-1})]$$

The entire gait sequence is represented by a sequence of transform vectors of the above form. Each transformed vector was scaled so that its maximum value is M. The algorithm is invariant to the exact value of M as long as the same M is used for all transformed sequences. For the results in the present paper $M = 255$. The scaling operation makes the proposed representation scale-invariant, a property which shall be very useful in most practical cases since subjects may walk at different distances from the camera imposing a need to compensate the distance difference prior to the application of any gait recognition methodology. A typical transform representation for a gait sequence is shown in fig. 2.

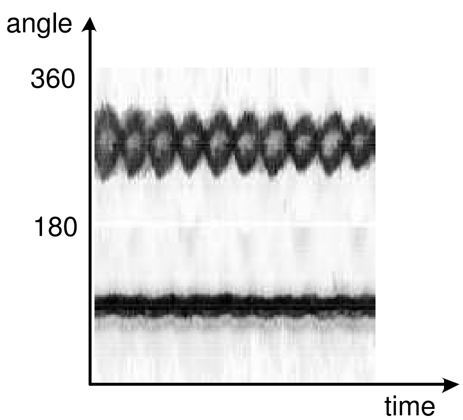


Fig. 2. Typical transform representation of a silhouette sequence (the grey scale was inverted for displaying convenience).

Our approach for identifying matching gait sequences is in the spirit of that in [1]. In [1], each test sequence was partitioned into several segments and the distance between each segment and the reference sequence was computed independently. The median of the resulting distances yielded the eventual distance between the test and the reference sequence. We take a similar approach in the present work by using segments of the sequence of feature vectors. Let $\mathcal{A}_{P_i}$ and $\mathcal{A}_{G_i}$ denote the angular transforms of the ith frame in the test segment and the reference sequence respectively. If N_p and N_G denote the respective number of frames, the distance metric between the test segment and the reference sequence is defined as

$$D = \min_{l} \sum_{i=1}^{N_p} \sqrt{\sum_{n=0}^{K-1} (\mathcal{A}_{P_i}(\theta_n) - \mathcal{A}_{G_{i+l}}(\theta_n))^2} \qquad (3)$$

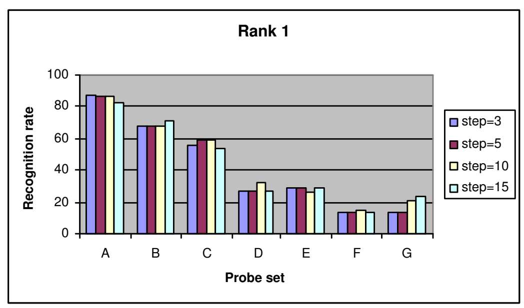
with $l = 1, \ldots, N_G$. Equation (3) implies that the distance is defined for an appropriate l, i.e. when the reference and the test sequences are aligned in phase. After all such distances are calculated between test segments and reference sequences of transform vectors, the *mean* of the distances is taken as the final distance.

For a given gait sequence, its distance from all the sequences in a reference database is calculated in order to determine a matching gait sequence. Since a smaller D means a closer match, its matching sequence in a set of sequences is identified as the sequence with which the above distance is minimum.

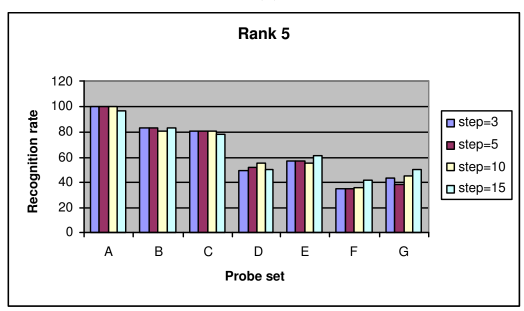
5. EXPERIMENTAL RESULTS

For the experimental evaluation for the present scheme we tested the present methodology on USF's Gait Challenge database which contains human gait sequences captured under different conditions. The Gallery (reference) set of gait sequences was used as the system database and the Probe (test) sets A-G were considered to contain sequences of unknown subjects who should be recognized by comparison of their gait sequences to the sequences in the Gallery set. The capturing conditions for the sequences in each of the Probe sets A-G may differ from the Gallery sequences in walking surface (cement/grass), shoe type (type A/B), and view-angle (left/right). The Gallery set contains gait sequences of individuals walking on Grass, wearing type-A shoes, and was captured using the Right camera. The exact conditions for the recording of the Probe sequences are summarized in brackets in Table 1 where C,G,A,B,L,R, stand for Cement, Grass, shoe type A, shoe type B, Left view, and Right view respectively.

For the performance evaluation, we report Cumulative Match Scores (as in [10]) at rank 1 and rank 5. Rank 1 results report the percentage of the subjects in a probe set that were identified exactly. Rank 5 results report the percentage of probe subjects whose true match in the Gallery set was in the top 5 matches.



(a)



(b)

Fig. 3. The effect of the variation of $\Delta\theta$ on the performance of the proposed gait recognition methodology. (a) rank-1, (b) rank-5.

Initially, we examined the effect of $\Delta\theta$ variation on the performance of the algorithm. Cumulative match scores derived for several $\Delta\theta$ values are reported at rank-1 and rank-5 in Figs 3(a) and (b) respectively. As seen, the recognition performance is not critically dependent on $\Delta\theta$ variations. However, we observed that using a larger angle step is usually beneficial for the higher-rank scores at the cost of a lower rank-1 score and vice-versa. It seems that the less detailed transform resulted by a larger angle step is more appropriate for approximate (higher-rank) recognition but less appropriate for exact (rank-1) recognition. In this work, we use $\Delta\theta = 5$ for all our experiments.

Cumulative Match Scores (CMS) at ranks 1 and 5 are reported in Table 1 in comparison to the methods in [1] and [6]. The complete CMS curves are presented in fig. 4. As seen, the proposed system, based on Preprocessing and our Angular analysis for Gait Recognition (PAGR), is able to capture sufficient gait information so that the results are comparable and in most cases better than the results reported in [1]. Even though our approach operates in the transform domain, i.e. using a reduced representation of the original silhouettes (the method in [1] uses raw silhouettes), the resulting scheme produces good results in comparison to the method in [1]. The efficiency of the proposed system is based on the silhouette preprocessing and the angular representation. It should be noted that in an angular representation a silhouette is represented using 24 (for $\Delta\theta = 15$) to 120 coefficients (for $\Delta\theta = 3$). This makes our system much faster than the system in [1] and is suitable for deployment in practical cases in which real-time recognition is important. We also compared our method with the method in [6]. As seen in Table 1, the approach taken in the present paper, which is significantly simpler since it does not require detection of contour pixels and does not perform PCA, outperforms in most cases the method in [6].

Probe Set	Rank 1			Rank 5		
	PAGR	[1]	[6]	PAGR	[1]	[6]
A (GAL) [71]	**86**	79	70	**100**	96	93
B (GBR) [41]	**68**	66	59	**83**	81	**83**
C (GBL) [41]	**59**	56	51	**81**	76	71
D (CAR) [70]	27	29	**34**	52	61	**64**
E (CBR) [44]	**29**	24	21	**57**	55	45
F (CAL) [70]	14	**30**	27	35	**46**	39
G (CBL) [44]	**14**	10	**14**	**38**	33	26

Table 1. Comparison to the baseline algorithm in [1] and the algorithm in [6]. The probability of identification (in percent) at ranks 1 and 5 is reported. The conditions under which each probe sequence was recorded are summarized in brackets. The number of subjects (gait sequences) in each probe set is shown in squared brackets.

6. CONCLUSIONS

A new method was presented for gait-assisted recognition based on binary silhouettes. The raw silhouette sequences were initially enhanced using preprocessing. A transform representation of silhouettes was subsequently applied which has many attractive features that make it very suitable for deployment in gait analysis and recognition applications. A system based on the new transform was experimentally evaluated using the Gait challenge database and was seen to outperform other gait recognition methods of similar or higher complexity.

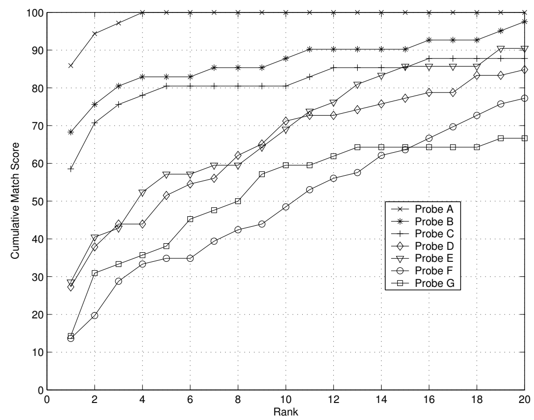


Fig. 4. Cumulative match scores using the proposed system.

7. REFERENCES

[1] P. J. Phillips, S. Sarkar, I. Robledo, P. Grother, and K. W. Bowyer, "The gait identification challenge problem: Data sets and baseline algorithm," in *International Conference on Pattern Recognition*, August 2002.

[2] A. Sunderesan, A. K. Roy Chowdhury, and R. Chellappa, "A Hidden Markov Model Based Framework for Recognition of Humans from Gait Sequences," in *Proc.ICIP 2003*, Barcelona, Spain, September 2003.

[3] Y. Liu, R. Collins, and Y. Tsin, "Gait Sequence Analysis using Frieze Patterns," in *European Conference on Computer Vision*, Copenhagen, May 2002.

[4] L. Lee, G. Dalley, and K. Tieu, "Learning Pedestrian Models for Silhouette Refinement," in *International Conference on Computer Vision*, 2003.

[5] J. P. Foster, M. S. Nixon, and A. Prugel-Bennett, "Automatic gait recognition using area-based metrics," *Pattern Recognition Letters*, vol. 24, pp. 2489–2497, 2003.

[6] L. Wang, T. Tan, H. Ning, and W. Hu, "Silhouette Analysis-based Gait Recognition for Human Identification," *IEEE Trans. Pattern Anal. and Mach. Intell.*, vol. 25, no. 12, pp. 1505–1518, Dec. 2003.

[7] F. Liu and R. Picard, "Finding periodicity in space and time," in *Proceedings of the Sixth International Conference on Computer Vision*, 1998.

[8] R. T. Collins, R. Gross, and J. Shi, "Silhouette-based Human Identification from Body Shape and Gait," in *Intl' Conference on Face and Gesture*, Oct 2002.

[9] D. Tolliver and R. T. Collins, "Gait Shape Estimation for Identification," in *4th Int'l Conf on Audio and Video-Based Biometric Person Authentication*, Guilford, UK, June 2003.

[10] P. J. Phillips, H. Moon, S. Rizvi, and P. Raus, "The feret evaluation methodology for face recognition algorithms," *IEEE Trans. Pattern Anal. and Mach. Intell.*, vol. 22, no. 10, pp. 1090–1104, 2000.