2004 IEEE 6th Workshop on Multimedia Signal Processing

Gait Recognition Using Dynamic Time Warping

Nikolaos V. Boulgouris, Konstantinos N. Plataniotis, and Dimitrios Hatzinakos
The Edward S. Rogers Sr. Department of Electrical and Computer Engineering
University of Toronto
Toronto, Ontario, Canada
Email: nikos@comm.toronto.edu, kostas@dsp.toronto.edu, dimitris@comm.toronto.edu

Abstract— **We propose a methodology for gait recognition based on dynamic time warping. The gait sequences are initially partitioned into gait cycles and then the test cycles are compared to reference cycles using dynamic time warping. The final distance between a test and a reference sequence is determined using a nonlinear rule. Experimental results are reported showing an improvement in recognition performance in comparison to the baseline algorithm on the "Gait Challenge" database.**

I. Introduction

THE development of innovative systems for the efficient identification of individuals has attracted much interest during the past few years. A variety of new biometrics have been proposed, intended to complement the biometrics that are deployed in current security systems. Gait recognition [1], [2], [3] is a new technology which aims at identifying people from the way they walk.

An important issue in gait recognition is related to the identification of individuals from walking sequences in which the walking persons walk at different speeds. In such cases time normalization is necessary before or during the recognition process. In general, time warping can be performed implicitly, i.e. by the resizing along the time-axis of patterns that depict the evolution of a feature through time, or explicitly by finding the correspondence between frames/features in sequences of different durations. Dynamic time warping [4] is the most common technique that can be used for this purpose.

One of the first attempts to use gait for recognition purposes was presented in [1]. Initially, the existence of a walking person in a scene was detected by analyzing the spatiotemporal patterns of walking persons. Subsequently, patterns corresponding to angle signals were used for identifying a walking subject among a database of subjects. Time warping was used for recognition of gait patterns at different walking speeds.

A baseline algorithm was presented in [5] aiming to serve as a reference for the comparison of different gait recognition methods. The baseline algorithm directly compares gait sequences without attempting any time normalization. For this reason, it is not appropriate for cases in which the reference and the test subjects are walking at different speeds.

In [6], each gait sequence was mapped to similarity plots which were derived by the self-similarities between each pair of images in the sequence. The feature vectors that were used for classification consisted of units of self-similarity of size one period. The feature vectors were scaled to a constant size in order to compensate for period differences. The scaled feature vectors were used for gait recognition using standard statistical pattern recognition techniques.

The use of shape for gait recognition was investigated in [7]. The gait sequences were clustered into groups of similar frames. Test and reference cluster centers were compared and classification was performed using nearest neighbors. Although this approach does not exploit the dynamic information that exists in a gait sequence, and therefore does not require time warping, it yielded good results.

In [8], silhouettes were represented using a vector describing the width of the outer contour of the silhouette. The sequences of silhouettes were partitioned into gait cycles and dynamic time warping was applied, using several gait cycles, in order to match reference and test sequences.

In this work, we apply dynamic time warping in order to calculate distances between all combinations of test and reference gait cycles in a pair of test and reference sequences. Based on these distances, we use a nonlinear rule for the determination of the final distance between a test and a reference sequence. The proposed methodology is experimentally evaluated and is shown to attain very promising performance.

The organization of the paper is as follows: section II describes the partitioning of gait sequences into gait cycles. The computation of the distances between test and reference sequences based on dynamic time warping is described in section III. Experimental evaluation is presented in section IV and finally, conclusions are drawn in section V.

II. Partitioning into Cycles

We assume that input to our system are *sequences of silhouettes* that were extracted from video sequences depicting walking persons. Before applying our gait recognition technique, we align the silhouettes so that their center coincides with the center of the frame. In order to exploit the periodicity of walking in a gait sequence, we partition the gait sequence into cycles by locating the frame indices at which the sum of the foreground pixels is minimized. Apparently, the above process will partition the sequences in intervals between consecutive mid-stances [9]. Such intervals actually correspond to *half cycles* which hereafter will be termed *cycles* for simplicity.

Since the function $f_s(t)$ of the sum of the foreground pixels with respect to time is noisy (see fig. 1(a)), we identify the

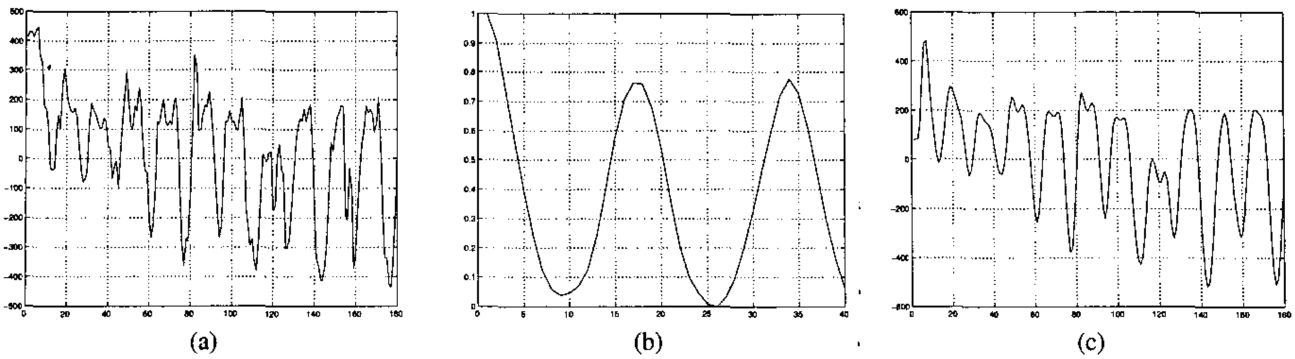

Fig. 1. (a) Original signal, (b) Autocorrelation, (c) Smoothed signal.

cycle length T_G by calculating the autocorrelation $R_w[m]$ (see fig. 1(b)) of the signal $w(t)$ which is derived from $f_S(t)$ by subtracting its mean value and dividing by its range. The autocorrelation of $w(t)$ is defined as

$$R_w[m] = E\{w(t)w(t+m)\} \tag{1}$$

where $E\{\cdot\}$ denotes expectation. The cycle length T_G is easily calculated as the smallest m, other than $m = 0$, at which there is a local maximum of $R[m]$. As seen in fig. 1(b), the determination of T_G from the autocorrelation function is unambiguous despite the fact that the original sequence was noisy.

The signal corresponding to the foreground sum can be smoothed using traditional smoothing techniques [10]. We formulate a system of Wiener-Hopf equations, in which the unknown optimal filter coefficients are expressed in terms of the autocorrelation of the signal, and solve it using numerical methods. Subsequently, we filter the foreground sum signal using the resulting filter. The filtered version is shown in fig. 1(c). A similar, in some sense, approach was taken in [11] where optimal prediction coefficients were used as the parameters required for the maximum entropy spectrum of a time-series corresponding to the vertical displacement of the silhouette centroid. Using this approach, the authors in [11] identified the fundamental frequency of the gait.

The methodology for partitioning into cycles relies on the detection of local minima. Since occasionally, due to the bad quality of silhouettes, a minimum may be signaled in the middle of a cycle yielding inaccurate cycle boundaries, we introduce an additional step which merges consecutive cycles whose length is small in comparison to the cycle length T_G of the gait sequence.

III. Distance Calculation Based on Dynamic Time Warping

For the description of the recognition process, we adopt the terminology in [12], according to which the reference gait sequences will be termed *gallery* sequences, whereas the test gait sequences will be termed *probe* sequences.

After the partitioning of gait sequences into cycles, using the technique of the previous section, dynamic time warping is performed between all cycles in a pair of test and gallery sequences. Dynamic time warping [4] is a nonlinear time normalization technique based on dynamic programming. Given two cycles of different duration, a warping function can be calculated which maps the time axis of the probe cycle to the time axis of the gallery cycle. The warping function is determined using a trellis consisting of nodes that represent the Euclidean distances between frames (silhouettes) in the probe and gallery cycles. Essentially, the warping function is defined by the path which links the beginning and the ending nodes of the trellis and exhibits the minimum cumulative distance. In this work, we assume warping functions with Type-I local constraints [13], according to which each node in the trellis is accessible from the three nodes shown in fig. 2(a). An example of a warping path is depicted in fig. 2(b).

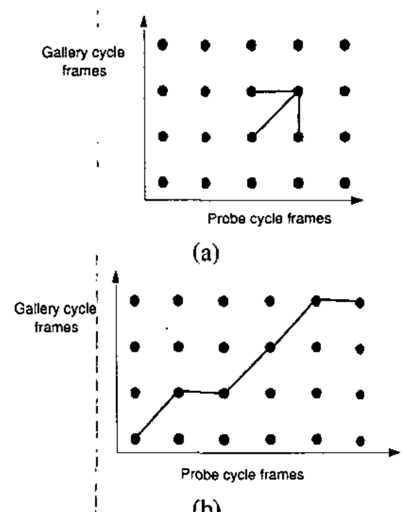

Fig. 2. (a) Type-I local constraints, (b) Warping path.

For each pair of probe and gallery sequences, the distances calculated as described above form a two-dimensional matrix, shown in Fig. 3, which contains all distances between probe and gallery cycles. The challenge here is to determine a distance D_r between the probe sequence and the gallery sequence which will accurately reflect the extent of dissimilarity between the two gait sequences. A prerequisite which we may intuitively set in the formulation of an algorithm for determining D_r is that the distance calculation process should be symmetric with respect to probe and gallery sequences, i.e. if the probe and gallery sequences were interchanged, the computed distance would be identical. Therefore, any process applied in a "probe-wise" manner should also be applied in a "gallery-wise" manner. For this reason, we treat each probe (gallery) cycle separately and evaluate its similarity with gallery (probe) cycles by taking the minimum distance of this probe (gallery) cycle to the gallery (probe) cycles with which it was compared. In practice, we find the minima in each line and each column of matrix $\mathcal{D}$ i.e.

$$d^p_{min,i} = \min_j\{d_{ij}\}, \text{for } \ i = 1, \dots, I$$

$$d^g_{min,j} = \min_i\{d_{ij}\}, \text{for } \ j = 1, \dots, J$$

where d_{ij} denotes the distance between the ith probe cycle and the jth gallery cycle. The above distances form two sets:

$$\mathcal{S}_p = \{d^p_{min,1}, \dots, d^p_{min,I}\} \text{ and } \mathcal{S}_g = \{d^g_{min,1}, \dots, d^g_{min,J}\}$$

We can take the median of the distances in the sets $\mathcal{S}_p$ and $\mathcal{S}_g$:

$$D^p_r = median\{d^p_{min,1}, \dots, d^p_{min,I}\}$$

and

$$D^g_r = median\{d^g_{min,1}, \dots, d^g_{min,J}\}$$

and average them in order to calculate the final distance D_r between the test and the reference sequence:

$$D_r = \frac{1}{2}(D^p_r + D^g_r) \tag{2}$$

This process is illustrated in fig. 3.

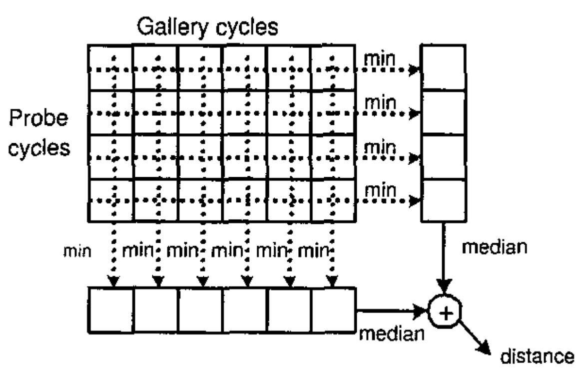

Fig. 3. Graphical representation of the proposed distance calculation process using a matrix comprising of distances among all cycles in a probe sequence against all cycles in a gallery sequence.

IV. Experimental Results

For the experimental evaluation of our scheme, we tested the present methodology on the Gait Challenge database[1] of the University of South Florida (USF) 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).

For the performance evaluation, we report in Table I Cumulative Match Scores at ranks 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.

Comparisons are with the USF method in [5]. As seen, the proposed method, termed DTW, outperforms in most cases the method in [5]. It appears that the time normalization achieved by the dynamic time warping process allows more accurate distance estimation and yields improved results.

TABLE I

Comparison to the baseline algorithm in [5]. The probability of identification P_I (in percent) at ranks 1 and 5 is reported.

	P_I (rank 1)		P_I (rank 5)	
Set	DTW	[5]	DTW	[5]
A	85	79	97	96
B	76	66	88	81
C	61	56	81	76
D	36	29	56	61
E	29	24	67	55
F	21	30	52	46
G	24	10	48	33

The proposed system was also evaluated in terms of verification performance. The most widely used method for this task is to present Rate Operating Characteristic curves (ROC). In an access control scenario, this means calculating the probability of positive recognition of an authorized subject versus the probability of granting access to an unauthorized subject. For calculating the above probabilities, we vary the distance that determines which subjects are granted access and which are not. For this calculation we use normalized distances (as in [14]). Using normalized distances is in general more suitable for verification purposes although it imposes additional computational complexity to the verification process.

The ROC's in their entirety are displayed in Fig. 5. As seen, the proposed system yields good performance although,

[1]We used the May 2001 database.

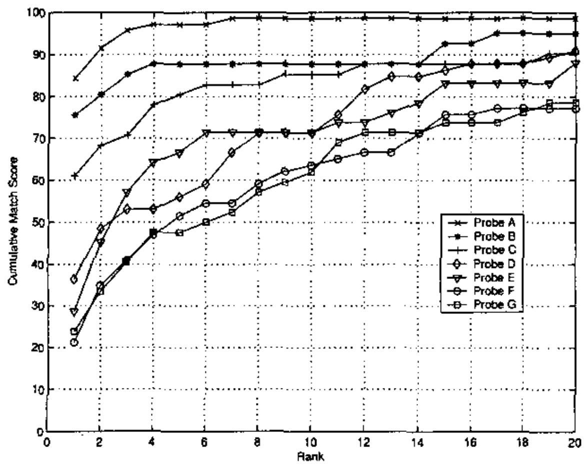

Fig. 4. Cumulative Match Scores for Probes A-G using the proposed method.

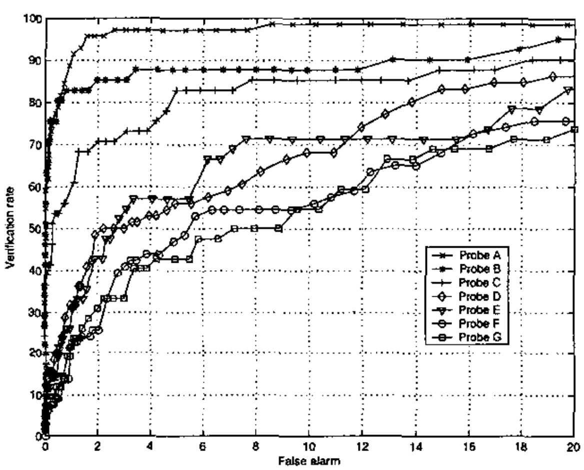

Fig. 5. Rate Operating Characteristics for Probe sets A-G using the proposed methods. Normalized distances were used.

at present, it appears that it might not be used on its own for the purpose of authentication without the use of additional biometrics or additional access control mechanisms. However, for interpreting these results, we should also take into consideration the fact that the database used for evaluating verification performance was captured outdoors and is very noisy. In a practical access control system, the capturing environment could be controlled, e.g. with uniform background, and multiple cameras could be used for generating high-quality gait sequences. In such a controlled environment, we expect that the verification performance would be significantly better allowing the deployment of gait-assisted verification in practical cases.

V. Conclusions

A novel method for gait recognition was presented based on dynamic time warping. The gait sequences were partitioned into gait cycles and each test cycle was matched to a reference cycle using dynamic time warping. The final distance between a test and a reference sequence was determined, from the collection of the distances between cycles, using a nonlinear rule. The experimental evaluation demonstrated improvements in recognition performance in comparison to the baseline algorithm on the "Gait Challenge" database.

Acknowledgment

This work was partially supported by Bell University Laboratories at the University of Toronto.

References

[1] S. A. Niyogi and E. H. Adelson, "Analyzing and recognizing walking figures in XYT," in *Proceedings of Computer Vision and Pattern Recognition*, June 1994.

[2] L. Lee and W. E. L. Grimson, "Gait Analysis for Recognition and Classification," in *Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition*, May 2002.

[3] D. Cunado, M. S. Nixon, and J. N. Carter, "Automatic Extraction and Description of Human Gait Models for Recognition Purposes," *Computer Vision and Image Understanding*, vol. 90, no. 1, pp. 1–14, 2003.

[4] H. Sakoe and S. Chiba, "Dynamic Programming Algorithm Optimization for Spoken Word Recognition," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 26, no. 1, pp. 43–49, February 1978.

[5] 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.

[6] C. BenAbdelkader, R. Cutler, and L. Davis, "Motion-based recognition of people in eigengait space," in *IEEE International Conference on Automatic Face and Gesture Recognition*, 2002.

[7] 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.

[8] A. Kale, N. Cuntoor, B. Yegnanarayana, A. N. Rajagopalan, and R. Chellappa, "Gait analysis for human identification," in *4th International conference on Audio Video based Person Authentication*, 2003.

[9] 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.

[10] A. Papoulis, *Probability Random Variables and Stochastic Processes*, 3rd Edition, McGraw-Hill, 1991.

[11] J. Little and J. Boyd, "Recognizing people by their gait: the shape of motion," *Videre, International Journal of Computer Vision*, vol. 14, no. 6, pp. 83–105, 1998.

[12] 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.

[13] L. Rabiner and B. Juang, *Fundamentals of Speech Recognition*, Prentice Hall, New Jersey, 1993.

[14] P. Grother, "Face recognition vendor test 2002. Supplemental Report. NISTIR 7083," *available from http://www.frvt.org*, Feb. 2004.