

Automatic Gait Recognition via Fourier Descriptors of Deformable Objects

Stuart D. Mowbray and Mark S. Nixon

University of Southampton
Southampton, SO17 1BJ, UK
{sdm01r,msn}@ecs.soton.ac.uk

Abstract. We describe a new method for Automatic Gait Recognition based around the use of Fourier descriptors that model the periodic deformation of human gait. Fourier descriptors have been used successfully in the past to model the boundary of static or moving, rigid-bodied objects, but many objects actually deform in some way as they move. Here we use Fourier descriptors to model not only the object's boundary, but also the spatio-temporal deformations under which the object's boundary is subjected. We applied this new method to the Large Gait Database, compiled at the University of Southampton, and found that the Fourier descriptors obtained for each person appear to be unique and can be used for recognition. Successful recognition rates of over 85% were obtained from the Large Gait Database using only a small set of descriptors.

1 Introduction

The use of biometrics to uniquely identify a person has grown rapidly over the years. From the use of finger prints to the identification of humans through analysis of iris patterns, biometric-based techniques are fast becoming common place in a variety of application domains. Biometrics have been used most notably in the areas of security and surveillance, where reliable individual verification, without the risk of forgery, is essential. Usual methods of personal identification either involve lengthy procedures to gain accuracy (such as DNA testing), or are fast but unreliable (such as a PIN code system). Biometrics-based methods, however, provide the end user with a system which is both fast and (potentially) very reliable.

Recently much attention has been devoted to the use of human gait patterns as a biometric and to the analysis of human motion in general. Studying this motion, however, is a difficult task. Studies of human gait have been carried out in many disciplines, such as psychology, biomechanics, and various medical fields. However, these studies have usually required human interaction, such as manual labelling of the data or the use of markers. The aim of computer vision-based analysis of human gait therefore, is to automatically recover and describe human gait accurately and with minimal human intervention. Several models have been proposed for the description of the human body and also for the description of human gait [1, 3]. These models can only be applied in computer

vision, however, if a correct distinction can be made between the human subject in question and their surroundings — unfortunately, this in itself is a not a trivial problem. Human gait has many advantages over other biometrics, but perhaps its most notable advantage is that it is non-invasive, offering a means to verify identity without a subject’s permission or cooperation.

Many descriptors of human gait are kinematic in nature, relying on geometric descriptions of the various body parts and mathematical modelling of the affine transformations which describe their movement. Many kinematic *features* exist, such as the angles of various body parts through time, their velocity, and their acceleration. Other studies have taken a more statistical approach to produce a unique gait descriptor. These have included the use of Principal Component Analysis (for dimensionality compression) combined with Cananonical Analysis (for classification) [5], the use of velocity moments [10], and more recently the use of Hidden Markov Models [6]. Current research is also being carried out into describing human motion by analyzing the bilateral symmetry inherent in human gait [4].

As yet, the only models of human gait used in computer vision have taken one of two forms. One approach has been to model the pendular movement of the human thigh using simple harmonic motion [2], while the other approach models the thigh and lower leg movement as coupled oscillators [11]. These models, however, represent only sections of the human body and do so by finding a best-fit of the model to the image data. The method we describe here demonstrates a new approach to modelling the full movement and deformation of the body, rather than just a specific body part.

2 Spatio-temporal Fourier Descriptors

Fourier descriptors have long been established as a method for representing a two-dimensional shape’s boundary [7, 8, 12]. A major advantage of Fourier descriptors is that, when representing a shape in the Fourier domain, one can readily access the its frequency components. This can be useful, as macroscopic features (the general features of the shape) are found in the lower frequencies, whereas microscopic features (the details) are found in the higher frequencies.

If a shape’s boundary forms a closed curve, then this curve can be considered to be periodic such that:

$$c(l + L) = c(l) \tag{1}$$

Where L is the total length of the curve. Due to this periodicity it is possible to represent the shapes boundary using a Fourier series, with the coefficients of the series, a_{xk} , b_{xk} , a_{yk} , and b_{yk} , being the Fourier descriptors of the shape:

$$c(l) = \frac{a_{x0}}{2} + \sum_{k=1}^{\lfloor \frac{L}{2} \rfloor} \{a_{xk} \cos\left(\frac{kl2\pi}{L}\right) + b_{xk} \sin\left(\frac{kl2\pi}{L}\right)\} + j \left(\frac{a_{y0}}{2} + \sum_{k=1}^{\lfloor \frac{L}{2} \rfloor} \{a_{yk} \cos\left(\frac{kl2\pi}{L}\right) + b_{yk} \sin\left(\frac{kl2\pi}{L}\right)\} \right) \tag{2}$$

Given that human gait demonstrates a periodic deformation, we can capture one gait cycle (one period) as a sequence of periodically deforming shapes, s , such that

$$s(t, l) = s(t + T, l) \tag{3}$$

Where T is the period of time that has elapsed in one complete gait cycle. The deformation of the boundary of the shape in this image sequence can therefore be represented by the two-dimensional discrete Fourier series

$$s(t, l) = \sum_{u=0}^{\lfloor \frac{T}{2} \rfloor} \sum_{v=0}^{\lfloor \frac{L}{2} \rfloor} S(u, v) e^{j2\pi(\frac{ux}{T} + \frac{vy}{L})} \tag{4}$$

Given here in complex form for brevity.

The Fourier coefficients of this series characterize the whole of the gait cycle. These coefficients can be calculated, by utilizing the separability property of the Fourier transform, as two separate one-dimensional complex Fourier transforms

$$S(u, v) = \sum_{t=0}^{T-1} \left[\sum_{l=0}^{L-1} s(n, l) e^{-j(\frac{2\pi}{L})(vl)} \right] e^{-j(\frac{2\pi}{T})(ut)} \tag{5}$$

3 Derivation of a Gait Signature

In order to derive a signature for a gait sequence each image in the sequence is first pre-processed. It should be noted at this point that for Fourier analysis to be applicable, the data needs to be regularly sampled with respect to the dependent variables. In terms of the gait data used in this study this requires regular sampling in both space and time. Additionally, an equal number of sample points should also be acquired spatially for each image in a specified gait sequence to aid in the calculation of equation (5).

The following gives an overview of the process used to generate each gait signature. Firstly, background extraction (via chroma-keying) is performed on each image, then each image is thresholded to produce a silhouette (see Figures 1(a), 1(b)). The boundaries from each subject are then extracted by following the outer contour of each image to produce a complex boundary signal. Further to this, linear interpolation is performed to extract an equal number (256) of regularly sampled data points for each frame.

To produce an actual gait signature the spatio-temporal Fourier descriptors of the complex boundary signals are calculated using a trigonometric, equivalent version of equation (5) and the resulting magnitude spectrum is used as the gait signature (see Figure 1(c)).

4 Feature Selection

The gait signatures produced using the method described above contain a large number of descriptors, although the majority of the information about the subject’s gait is contained in the low-order descriptors. In order to use these gait

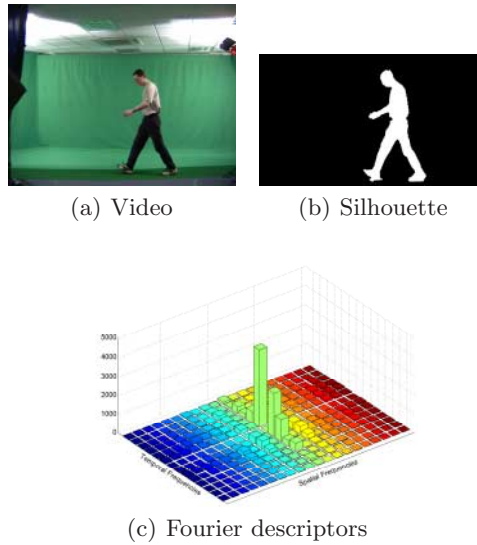


Fig. 1. Fourier descriptor generation

signatures for classification purposes it was necessary to reduce the number of descriptors that describe each subject. This was necessary for two reasons, firstly to extract only the descriptors that would be useful for classification, and secondly to reduce the dimensionality of the feature space – ensuring feasible classification speeds.

As mentioned above, the primary aim of feature selection in this case was to minimize intra-subject variance and maximize inter-subject variance, in order to increase the Correct Classification Rate (CCR). To achieve this we used a variation of the Bhattacharyya distance metric to measure inter-class separation with respect to the class covariances. The separation between the two classes a and b , for a given feature, is given by

$$S_{a,b} = [m_a - m_b] \left[\frac{\Sigma_a + \Sigma_b}{2} \right]^{-1} [m_a - m_b]^T \quad (6)$$

where m_a is the class mean and Σ_a is the covariance matrix of class a .

To gain a measure of a feature's ability to separate classes successfully a mean value of S was determined for each feature as

$$\bar{S} = \frac{1}{D^2} \sum_{a=1}^D \sum_{b=1}^D S_{a,b} \quad (7)$$

where D is the number of descriptors available. \bar{S} is then proportional to the class separability ability of the given feature.

Table 1. Results of k-nearest neighbour classification

Database	k	CCR(%)
SOTON	1	84.53
SOTON	3	86.20

5 Gait Recognition

To test the recognition capabilities of this new technique, spatio-temporal Fourier descriptors were calculated for each subject in the Large Gait Database at the University of Southampton [9]. This database consists of 115 subjects and 1062 sequences.

Classification was performed using a K-nearest neighbour classifier and cross-validated with the leave-one-out rule. This classifier assigns a test subject to be the same class as that of the modal class of the k^{th} nearest neighbouring subjects to it. The distance between classes is measured by the Euclidean distance, ED , given by

$$ED = \sqrt{\sum_{n=0}^N (\mathbf{x}_n - \mathbf{y}_n)^2} \quad (8)$$

where N is the dimensionality of the feature set, and \mathbf{x}_n and \mathbf{y}_n are the values of the n^{th} feature of the samples \mathbf{x} and \mathbf{y} respectively.

The classification results for $k = 1$ and $k = 3$ are shown in Table 1.

6 Performance Analysis

The performance and robustness of this new technique were evaluated with respect to missing data and varying resolutions, to simulate the effects of distance.

6.1 The Effect of Missing Frames

Missing frames or damaged data (if detectable) are simulated by omitting consecutive frames from the image sequences. As this is a Fourier-based technique, merely omitting frames would cause severe interference in the frequency-domain and therefore in the descriptors used for classification (the result would be the convolution of a sinc wave and the original descriptors in the case of contiguous missing frames). To overcome this Fourier interpolation is used to *predict* the missing or damaged data prior to classification. This is accomplished by sub-sampling the temporal data so that the data is still evenly spaced through time, but does not contain the missing data. The spectrum of this sub-sampled sequence is then zero-padded until its length is the same as the that of the original image sequence and interpolation is performed by inverse transforming this

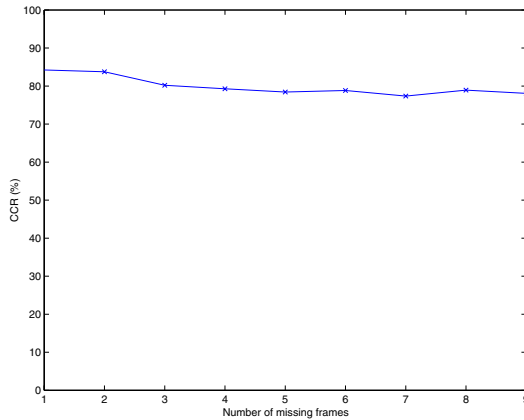


Fig. 2. Effects of interpolating missing temporal data

zero-padded spectrum back into the time domain. The resulting missing frames from the original sequence are then filled with these interpolated frames.

To test this technique several frames were omitted, then predicted through Fourier interpolation, for each sequence in the Southampton Large Gait Database. The results of classification using this technique are shown in Figure 2. As can be seen, quite a large proportion of temporal data can be omitted without a great loss of classification performance.

6.2 The Effect of Distance

The effect of distance is simulated by decreasing the spatial resolution of each image in an image sequence. The original images, which were at a resolution of 690×400 , were scaled so that their heights were 128, 64, and 32 respectively (see Figure 3). The results of classification at these resolutions, shown in table 2, show that the image resolution can be relatively small without a great loss in classification performance. The fact that such a loss of spatial resolution results in only a small drop in recognition rate can probably be accounted for by the facts that the majority of the information in a given Fourier descriptor is contained in the low-level descriptors and that for a spatial boundary of length N we can obtain up to $\frac{N}{2}$ descriptors. Therefore, as long as we have a sufficiently large value of N , an adequate number of descriptors can be obtained for recognition.

7 Conclusions

The aim of this paper was to investigate the use of Fourier descriptors in capturing and classifying the spatio-temporal movement of human gait. We have therefore presented a new technique for extracting a Fourier descriptor from an

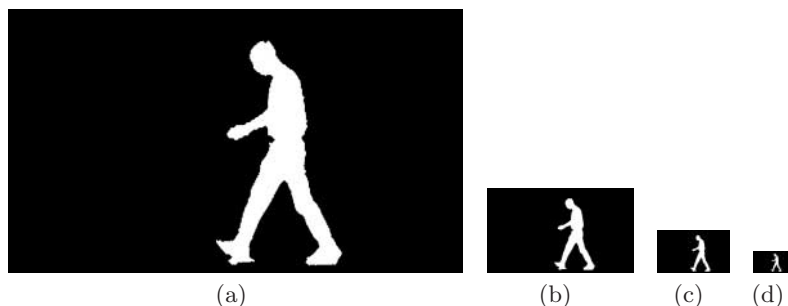


Fig. 3. Varying image resolutions: (a) 400; (b) 128; (c) 64; (d) 32

Table 2. Results of k-nearest neighbour classification ($k=3$) for varying resolutions

Image Height	CCR(%)
400	86.20
128	84.52
64	85.65
32	82.35

image sequence that is complete, compact, and captures the spatio-temporal motion associated with human gait. The fundamental motion of human gait is captured in a relatively low number of Fourier descriptors and as such this allows for the extraction of accurate Fourier descriptors from very low-resolution images sequences, necessary for identification of subjects at a distance. A secondary technique has been introduced which aids in the reconstruction of missing temporal data and allows for accurate Fourier descriptors to be calculated from sequences containing a reasonable amount of missing data. Recognition rates of over 85% were obtained from the Southampton Large gait database, which demonstrates that this new technique shows promise in automatic gait recognition.

Acknowledgments

We gratefully acknowledge partial support by the European Research Office of the US Army under Contract No. N68171-01-C-9002.

References

- [1] J. K. Aggarwal and Q. Cai. Human Motion Analysis: A Review. *Computer Vision and Image Understanding*, 73(3):428–440, March 1999. 566
- [2] D. Cunado, M. S. Nixon, and J. N. Carter. Automatic Gait Recognition via Model-Based Evidence Gathering. In *Proc. of AutoID '99: IEEE Workshop on Identification Advanced Technologies*, pages 27–30, 1999. 567
- [3] D. M. Gavrilu. The Visual Analysis of Human Movement: A Survey. *Computer Vision and Image Understanding*, 73(1):82–98, January 1999. 566
- [4] J. B. Hayfron-Acquah, M. S. Nixon, and J. N. Carter. Automatic Gait Recognition by Symmetry Analysis. In *Proc. of Audio-and-Video-Based Biometric Person Authentication*, pages 272–277, 2001. 567
- [5] P. S. Huang, C. J. Harris, and M. S. Nixon. Recognising humans by gait via parametric canonical space. *Journal of Artificial Intelligence in Engineering*, 13(4):359–366, 1999. 567
- [6] A. Kale, A. N. Rajagopalan, N. Cuntoor, and V. Kruger. Gait based recognition of humans using continuous HMMs. In *Proc. of the International Conference on Face and Gesture Recognition 2002*, 2002. 567
- [7] F. P. Kuhl and C. R. Giardina. Elliptic Fourier Feature of a Closed Contour. *Computer Graphics and Image Processing*, 18:236–258, 1982. 567
- [8] E. Persoon and K. Fu. Shape Discrimination Using Fourier Descriptors. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-7(3):170–179, 1977. 567
- [9] J. D. Shutler, M. G. Grant, M. S. Nixon, and J. N. Carter. On a large sequence-based human gait database. *Proc. Recent Advances in Soft Computing (RASC02)*, pages pp. 66–71, 2002. 570
- [10] J. D. Shutler and M. S. Nixon. Zernike Velocity Moments for Description and Recognition of Moving Shapes. In *Proc. of BMVC 2001*, 2001. 567
- [11] C. Y. Yam, M. S. Nixon, and J. N. Carter. Extended Model-Based Automatic Gait Recognition of Walking and Running. In *Proc. of 3rd Int. Conf. on Audio and Video-Based Biometric Person Authentication, AVBPA 2001*, 2001. 567
- [12] C. T. Zahn and R. Z. Roskies. Fourier Descriptors for Plane Closed Curves. *IEEE Transactions on Computers*, C-21(3):269–281, 1972. 567