



ELSEVIER

Available at
www.ComputerScienceWeb.com
POWERED BY SCIENCE @ DIRECT®

Pattern Recognition
Letters

Pattern Recognition Letters 24 (2003) 2175–2183

www.elsevier.com/locate/patrec

Automatic gait recognition by symmetry analysis

James B. Hayfron-Acquah ^{*}, Mark S. Nixon ^{*}, John N. Carter

*Image, Speech and Intelligent Systems Group, Department of Electronics and Computer Science, University of Southampton,
Southampton, S017 1BJ, United Kingdom*

Abstract

We describe a new method for automatic gait recognition based on analysing the symmetry of human motion using the Generalised Symmetry Operator. This approach is reinforced by the psychologists' view that gait is a symmetrical pattern of motion and results show that gait can indeed be recognised by symmetry analysis.

© 2003 Elsevier Science B.V. All rights reserved.

Keywords: Biometrics; Symmetry operator; Gait recognition

1. Introduction

We present a new method of extracting symmetry for automatic gait recognition based on analysing the symmetry of human motion. The method is reinforced from the psychologists' view that human gait is a symmetrical pattern of motion and also by other works that suggest symmetry is suitable for gait recognition. We use the Generalised Symmetry Operator which locates features according to their symmetrical properties rather than relying on the borders of a shape or on general appearance and hence does not require the shape of an object to be known in advance.

Current interests in biometrics include ear and face recognition, body tracking and hand gesture recognition, to mention just a few. Gait recogni-

tion is the most recent biometric to be added to this domain. As a biometric, gait involves recognising people by the way they walk regardless of their clothes or the background to them. One major advantage of gait over other biometrics (e.g. fingerprints) is the absence of need for contact with subjects. Further, gait can be difficult to disguise or conceal in application scenarios like bank robbery where robbers usually put on helmets, masks, spectacles and gloves making it virtually impossible to use any of the other known biometrics such as fingerprints or face. Currently, gait is also the only biometric at a distance and can be used when other biometrics are either obscured or at too low a resolution to be perceived. Though it could be argued that physical condition such as drunkenness, pregnancy and injuries involving joints can affect an individual's motion, these factors are similar in principle to factors affecting other biometrics.

There have been allied studies of gait, Nixon et al. (1999), notable among these are medical and

^{*} Corresponding authors.

E-mail addresses: jbha99r@ecs.soton.ac.uk (J.B. Hayfron-Acquah), msn@ecs.soton.ac.uk (M.S. Nixon), jnc@ecs.soton.ac.uk (J.N. Carter).

psychological studies, together with computer vision studies aimed to model human motion and to track people. Amongst these, psychologists, Cutting et al. (1978), suggest gait is a symmetrical pattern of motion and that humans perceive gait as unique.

There is already a number of approaches to gait recognition even though it is a fairly new research area. The spatio-temporal approach, presented by Niyogi and Adelson (1994) is probably the earliest. Here, the gait signature was derived from the spatio-temporal patterns of a walking person. The different patterns of the motions of the head and the legs in translation and time were extracted. The patterns were then processed to determine the motion of the bounding contours from which a five-stick model was fitted. The gait signature was then derived by normalising the fitted model in terms of velocity, that is by linear interpolation. Encouraging (85%) recognition rates were achieved.

Little and Boyd (1998) used optical flow to derive gait signatures by analysing the motion content (shape of motion) of a human walking and reported a recognition rate exceeding 90% on a small database of six subjects. Generic object-motion characterisation is also another approach where the gait signature is derived from a parametric eigenspace. Murase and Sakai (1996) applied the approach to a database of seven subjects with ten image sequences of each. The recognition rates were 88% and 100% for 8 and 16 eigenvectors, respectively. The approach was extended by Huang et al. (1999) using canonical analysis, a model free approach to reduce the dimensionality of the input data whilst optimising class separability. A recognition rate of 100% was obtained on a database of five subjects. Recently, Shutler et al. (2000) extended statistical gait recognition via temporal moments. This derived statistics with an intimate relationship to gait, with symmetry properties. On a small database of four subjects with four image sequences each, a recognition rate of 100% was obtained. The notion that gait is a biometric is supported by the most recent papers on automatic gait recognition, BenAbdelkader et al. (2001) and Johnson and Bobick (2001), especially the motion of self-similarity for recogni-

tion. The only model-based approach by Cunado et al. (1999) derived gait signatures from the frequency components of the variations in the inclination of the human thigh. As pendulums modelled the periodic motion of the thigh during walking, this again suggests that symmetry analysis is suited to gait recognition.

Most of the approaches have used small databases. We show how symmetry can be used to recognise gait, again using these databases and using a much larger database. We also investigate the performance of symmetry with respect to noise, occlusion and missing frames.

2. Symmetry and its extraction

2.1. Symmetry

Symmetry is a fundamental (geometric) property suggesting it to be an important principle of perception, Reisfeld et al. (1995). An object is said to be symmetric if its shape remains unchanged by the application of symmetry operations. Boolean symmetry operations can only assess symmetry when the shape of the object is known in advance, rendering them inefficient in most cases. The discrete symmetry operator can estimate symmetry without the knowledge of the object's shape, unlike feature extraction operators that find a shape by relying on its border or on general appearance. The symmetry transform assigns a symmetry measure to each point in the image and is determined with respect to a given point-symmetry group. Reisfeld et al. (1995) reported that the performance of the symmetry transform is not affected by existence of several objects in the scene. This depends on values selected for controlling parameters.

2.2. Symmetry extraction

The discrete symmetry operator uses edge maps of images from the sequences of subject silhouettes to assign symmetry magnitude and orientation to image points, accumulated at the midpoint of each analysed pair of points. The total symmetry magnitude (also known as iso-

tropic symmetry), $M(p)$ of each point P is the sum of the contributions of image points that have P as their mid point, that is

$$M(P_k) = \sum_{(i,j) \in \Gamma(P_k)} C(P_i, P_j) \quad (1)$$

where P_i and P_j are pairs of points having P_k as their mid point and $C(P_i, P_j)$ as their symmetry contribution.

Fig. 1 shows two edge points P_i and P_j and their symmetry contribution $C(P_i, P_j)$. The symmetry relation or contribution, $C(P_i, P_j)$ between the two points P_i and P_j is

$$C(P_i, P_j) = D_{i,j} Ph_{i,j} I_i I_j \quad (2)$$

where $D_{i,j}$ and $Ph_{i,j}$ (see Eqs. (3) and (6) respectively) are the distance and the phase between the two points. I_i and I_j are the logarithmic intensities at the two points. The symmetry distance weighting function gives the distance between two different points P_i and P_j , and is calculated as

$$D_{i,j} = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\left(\frac{\|P_i - P_j\|}{2\sigma}\right)^2\right) \quad \forall i \neq j \quad (3)$$

where σ controls the scope of the function. The logarithm intensity function, I_i , of the edge magnitude M at point (x, y) is

$$I_i = \log(1 + M_i) \quad (4)$$

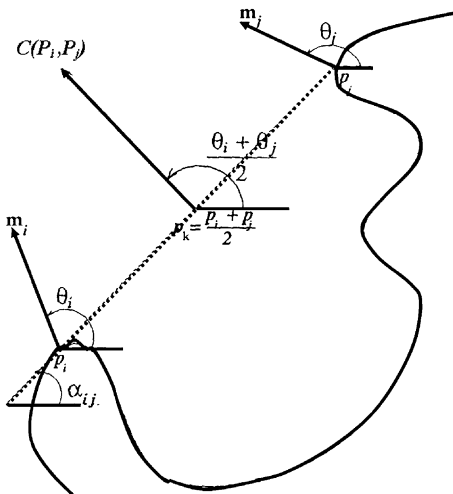


Fig. 1. The symmetry contribution of edge points P_i and P_j .

where

$$M_i = \sqrt{M_x^2(x, y) + M_y^2(x, y)} \quad (5)$$

where the M_x and M_y are derived by the application of (Sobel) edge templates. The phase weighting function between two points P_i and P_j is:

$$Ph_{i,j} = (1 - \cos(\theta_i + \theta_j - 2\alpha_{i,j})) \times (1 - \cos(\theta_i - \theta_j)) \quad \forall i \neq j \quad (6)$$

where

$$\alpha(i, j) = \text{atan}\left(\frac{y_i - y_j}{x_i - x_j}\right) \quad (7)$$

is the angle between the line joining the two points and the horizon. From Eq. (6), the phase weighting function has two factors. The first factor, $(1 - \cos(\theta_i + \theta_j - 2\alpha_{i,j}))$ makes it possible for maximum symmetry to be achieved when $(\theta_i - \alpha_{i,j}) + (\theta_j - \alpha_{i,j}) = \pi$. This is when the two gradients at the points P_i and P_j are oriented in the same direction towards each other. This situation corresponds to a dark object on a light background. The second factor, $(1 - \cos(\theta_i - \theta_j))$ is introduced because the case $(\theta_i - \alpha_{i,j}) = (\theta_j - \alpha_{i,j}) = \pi/2$ is included in $(\theta_i - \alpha_{i,j}) + (\theta_j - \alpha_{i,j}) = \pi$ when the first factor attains its maximum. This is when the edge directions are normal to the line joining the points P_i and P_j . This situation also corresponds to a light object on a dark background. The combination of factors makes it possible to achieve the same measure for different object reflectance and lighting conditions. In the distance weighting function, each value of σ implies a different scale thus making it suited to multi-resolution schemes. A large value of σ implies large-scale symmetry that gives distant points similar weighting to close points. Alternatively, a small value of σ implies local operation and local symmetry. Recently, Parsons and Nixon (1999) observed that increasing the value of σ increases the weighting given to the more distant points without decreasing the influence of the close points. A comparison with the Gaussian-like functional showed that the mean of the distribution locates the function on the mean value of the sample. A focus, μ , was therefore introduced into the distance weighting function to control the focusing capability of

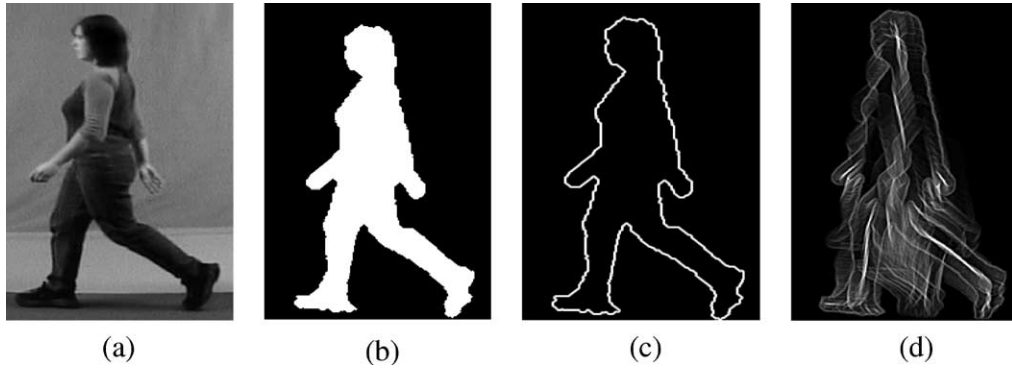


Fig. 2. Images from the new SOTON data: (a) original; (b) silhouette; (c) after Sobel and (d) symmetry map.

the function, hence further improving the scaling possibilities of the symmetry distance function. The resulting function is called the focus weighting function, FWF. This replaces Eq. (3) as follows:

$$\text{FWF}_{i,j} = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\left(\frac{\|P_i - P_j\| - \mu}{2\sigma}\right)^2\right) \quad \forall i \neq j \quad (8)$$

The addition of the focus into the distance weighting function moves the attention of the symmetry operator from points close together to a selected distance.

The symmetry contribution obtained is then plotted at the midpoint of the two points. The symmetry transform as discussed here detects reflectional symmetry, Reisfeld et al. (1995), that is invariant under 2D rotation, translation transformations and under change in scale, and as such has potential advantage in automatic gait recognition.

3. Symmetry and gait

3.1. Deriving a gait signature from silhouette information

The gait signature for a subject is derived from an image sequence. Each sequence of image frames consists of one gait cycle taken between successive heel strikes of the same foot, thus normalising for speed. This cycle has been considered, Murray (1967), “a total walking cycle” of gait. The fol-

lowing gives an overview of the steps involved in extracting symmetry from silhouette information. First, the image background is computed from the median of five image frames and subtracted from the original image (Fig. 2a) to obtain the silhouette (Fig. 2b). This was possible because the camera used to capture the image sequences was static and there is no translational motion. Moreover, the subjects were walking at a constant pace. The Sobel operator is then applied to the image in Fig. 2b to derive its edge-map, Fig. 2c. The edge-map is thresholded so as to set all points beneath a chosen threshold to zero, to reduce noise or remove edges with weak strength. These processes reduce the amount of computation in the symmetry calculation. The symmetry operator is then applied to give the symmetry map, Fig. 2d. For each image sequence, the gait signature, GS, is obtained by averaging all the symmetry maps, that is

$$\text{GS} = \left(\sum_{j=1}^N S_j \right) / N \quad (9)$$

where N is the number of symmetry maps in a sequence.

3.2. Deriving a gait signature from optic flow information

The process involved in deriving a gait signature from optic flow information is little different from that of the silhouette information. A gait cycle of a walking subject is used. Having removed

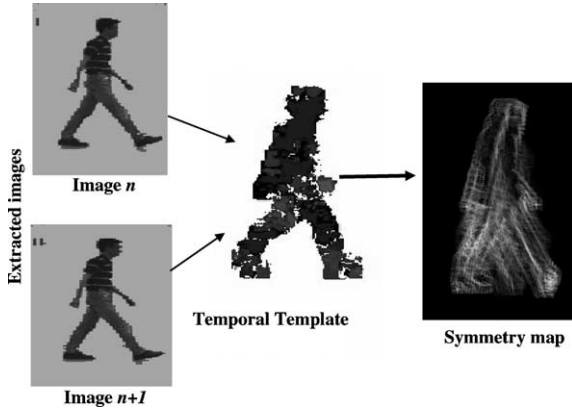


Fig. 3. Images involved in optic flow information.

the background from each original image to obtain the silhouette, two successive silhouettes are used at a time using the algorithm as proposed by Little and Boyd (1998) to extract the optic flow image which replaces the information supplied by the Sobel operator. The symmetry operator is then applied. Finally, all symmetry maps in the sequence are averaged to obtain the gait signature. Fig. 3 shows sample images of the process.

3.3. Gait recognition

Having obtained gait signatures for all image sequences in a database, we then apply the Fourier transform to obtain the description FD, that is

$$FD(u, v) = \sum_N \sum_N GS(x, y) e^{-j\frac{2\pi}{N}(ux+vy)} \quad (10)$$

where GS is a gait signature and the Fourier transform is implemented by the FFT. FD is then low-pass filtered to reduce sensitivity to high-frequency components. Different cut-off frequencies were used to determine the appropriate number of Fourier components. We select only the Fourier descriptions within a circle of radius, R , as

$$FD'(u, v) = \begin{cases} FD(u, v) & \text{if } (u^2 + v^2) \leq R^2 \\ 0 & \text{otherwise} \end{cases} \quad (11)$$

For purposes of classification or recognition, the similarity differences between the Fourier descriptions of the gait signatures are then calculated using Euclidean distance, that is

$$SD_{i,j} = \sum \|(|FD'_{i,x,y}| - |FD'_{j,x,y}|)\| \quad (12)$$

where the elements of $FD'i$ are the Fourier descriptions of the gait signatures of image sequence i remaining after low pass filtering. The magnitude spectra only were used here because they gave a better result than the inclusion of the phase information.

The new method was applied to three different databases of spatial templates. The original SOTON database has four subjects with four image sequences each and that of UCSD six subjects with seven image sequences of each. For both original SOTON and UCSD databases, we derived gait signatures for silhouette and optical flow information. These provide alternative versions of the input data for our technique to process. From similar work done by Huang et al. (1999), combining the silhouette and optical flow information can allow better recognition rates than achieved using the two modalities separately but our aim here is to establish whether human gait has distinct symmetrical properties that can be extracted for gait recognition. The values for σ and μ used were 27 and 90, respectively, unless otherwise stated. The k -nearest neighbour rule was then applied for classification, using $k = 1$ and 3, as summarised in Table 1. The correct classification rates (CCRs) were 100% for both $k = 1$ and 3 for the original SOTON database for both data types. For the UCSD database, the recognition rates for silhouette information were 97.6% and 92.9% for $k = 1$ and 3, respectively. A CCR of 92.9% was obtained for the optical flow information, for both $k = 1$ and 3.

After our earlier work, Hayfron-Acquah et al. (2001), we applied the same method to a much larger database of 28 subjects, the new SOTON database. This new database equals in size the largest (published) contemporaneous gait database though new and larger databases will emerge in the near future. Each subject has four image sequences giving in total 112 image sequences. With this database, only the silhouette information is used and the recognition rates obtained are also shown in Table 1. Clearly, using a much larger database still gave the same good recognition rates and that is very encouraging.

Table 1

Initial results obtained from three different databases

Database	# Subjects	# Sequences	Data type	CCR (%)	
				$k = 1$	$k = 3$
Original SOTON	4	16	Silhouette	100	100
			Optical flow	100	100
UCSD	6	42	Silhouette	97.6	92.9
			Optical flow	92.9	92.9
New SOTON	28	112	Silhouette	97.3	96.4

For the low pass filter, all possible values of radius were used to investigate the number of components that can be used (covering 0.1–100% of the Fourier data). Though the results of Table 1 were achieved for all radii greater than 3 (using the original SOTON database), selecting fewer Fourier components might affect the recognition rates on a larger database of subjects, and this needs to be investigated in future. Fig. 4 shows the general trend of the recognition rate against the different cut-off frequencies. The figure shows that selecting about 1.1% of the Fourier descriptions is enough to give the best recognition rates for both $k = 1$ and 3, but this is only on a small database.

3.4. Performance analysis of symmetry operator

Performance was evaluated with respect to missing spatial data, missing frames and noise

using the original SOTON database. Out of the 16 image sequences in the database, one (from subject 4) was used as the test subject with the remainder for training.

3.4.1. Missing frames

The evaluation, aimed to simulate time lapse, was done omitting a consecutive number of frames. For a range of percentages of omitted frames, Fig. 5a, the results showed no effect on the recognition rates for $k = 1$ or 3. This is due to the averaging associated with the symmetry operator. Fig. 5a shows the general trend of deviation of the best match of each subject to the test subject. The increase in the similarity difference measures as more frames are omitted appears to be due to the apparent increase in high frequency components when averaging with fewer frames.

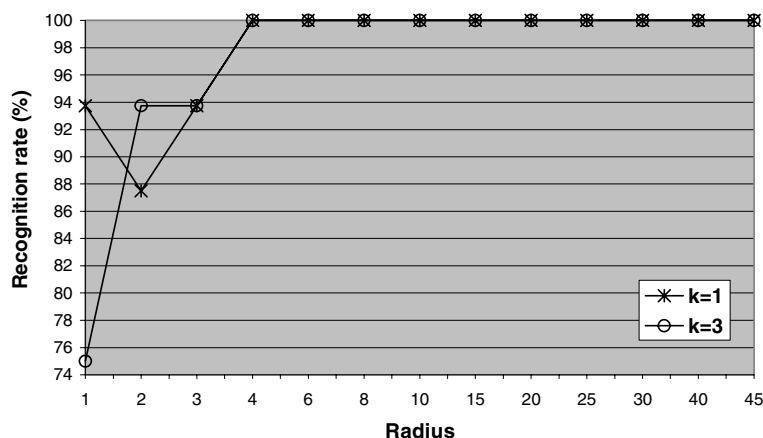


Fig. 4. Recognition rates against the different cut-off frequencies.

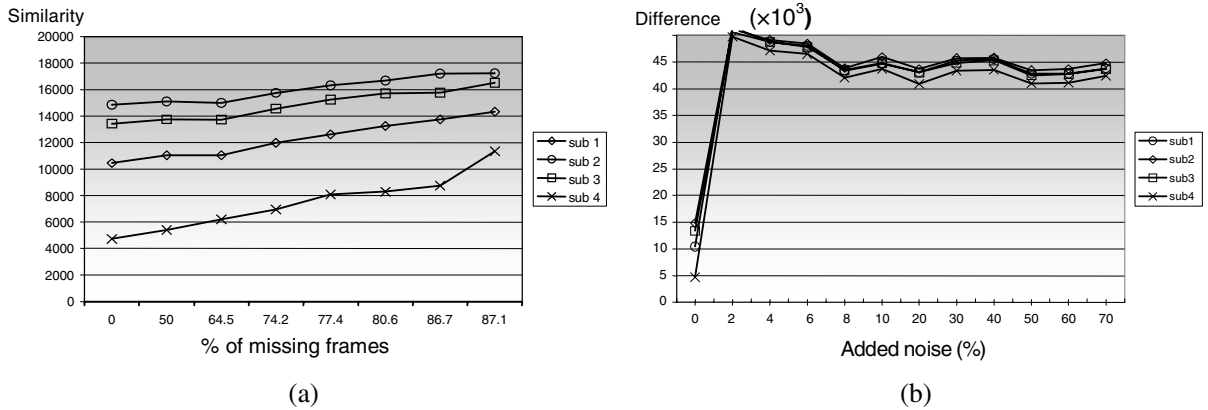


Fig. 5. Performance analysis: omitted frames and addition of noise: (a) effects of omitting image frames and (b) effects of adding noise.

3.4.2. Adding/omitting spatial data

The evaluation was done by masking with a rectangular bar of different widths: 5, 10 and 15 pixels in each image frame of the test subject and at the same position. The area masked was on average 13.2%, 26.3% and 39.5% of the image silhouettes, respectively. The bar either had the same colour as the image silhouette or as the background colour, as shown in Fig. 6, simulating omission and addition of spatial data, respectively.

In both cases, recognition rates of 100% were obtained for bar size of 5 pixels for both $k = 1$ and 3. For a bar width of 10 pixels, Fig. 6c failed but Fig. 6a gave the correct recognition for $k = 3$ but not for $k = 1$. Fig. 6c failed as the black bar changes effective body shape and recognition uses body shape and body dynamics. For bar sizes of 15 (i.e. Fig. 6b and d) and above, the test subject could not be recognised as subject is completely covered in most of the image frames. This suggests that

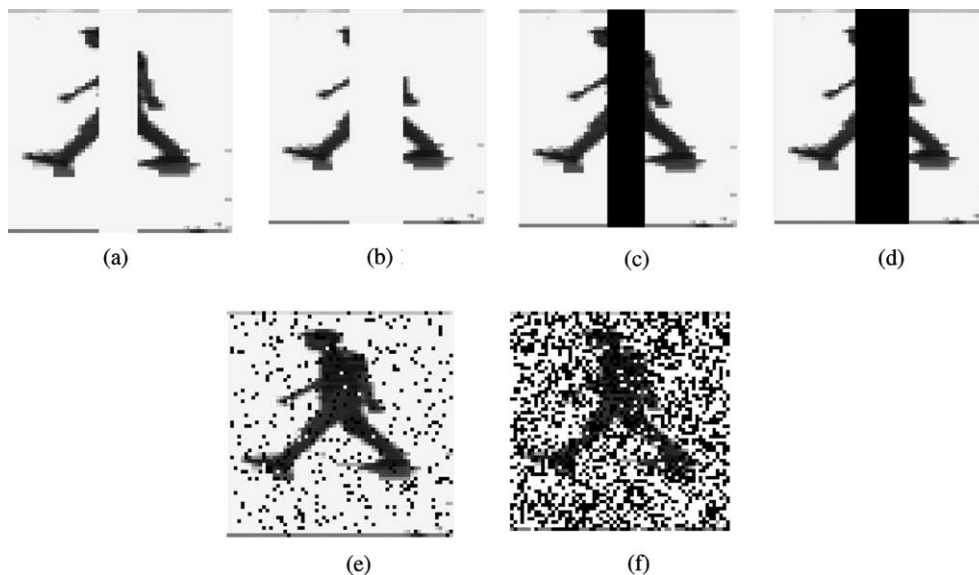


Fig. 6. Occluded and noisy data: (a) 10 pixels; (b) 15 pixels; (c) 10 pixels; (d) 15 pixels; (e) 10% and (f) 50%.

recognition is unlikely to be adversely affected when a subject walks behind a vertically placed object, such as a lamp post.

3.4.3. Noise

To investigate the effects of noise, we added synthetic noise to each image frame of a test subject and compared the resulting signature with those of the other subjects in the database. Fig. 6e and f shows samples of the noise levels used. The evaluation was carried out under two conditions. First by using the same values of σ and μ (Eq. (8)) as earlier, for a noise level of 5%, the recognition rates for both $k = 1$ and 3 were 100%. For 10% added noise, the test subject could still be recognised correctly for $k = 1$ but not for $k = 3$. This suggests that the recognition is not greatly affected by limited amounts of failure in background extraction, though those errors are less likely to be uncorrelated. With added noise levels of 20% and above, the test subject could not be recognised for $k = 1$ or 3. In order to reduce contribution of distant (noisy) points, the values of σ and μ were then made relatively small. Now the recognition rates (100%) were not affected for both $k = 1$ and 3 for added noise levels even exceeding 60%. Fig. 5b shows how in this case the best match of each subject deviated from that for the test subject, as more noise was added. Here, the high recognition rate is consistent with the same order being achieved despite addition of noise. As such, the symmetry operator can be arranged aiming to include noise tolerance, but the ramifications of this await further investigation.

4. Conclusions and further work

We have shown how symmetry can be used to recognise people by their gait, as supported by research in psychology. Human gait appears to have distinct symmetrical properties that can be extracted for recognition. The symmetry operator, essentially, forms an accumulator of points, which are measures of the symmetry between image points to give a symmetry map. By using the symmetry operator, the Discrete Fourier Transform and a basic nearest-neighbour approach, the

results have produced encouraging recognition rates on a small database, a performance that is comparable with other approaches applied to the same databases. Also, the recognition rates on a much larger database remained almost the same as those obtained from smaller databases. The symmetry operator has been shown to handle missing spatial data, missing image frames, and to some extent noise. As such, symmetry appears to be promising approach with performance advantages and we shall look forward to developing further the notion of moving object description by temporal symmetry.

Acknowledgements

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

References

- BenAbdelkader, C., Cutler, R., Nanda, H., Davis, L., 2001. EigenGait: Motion-based recognition using image self-similarity. In: 3rd Internat. Conf. on Audio-and-Video-Based Biometric Person Authentication, pp. 289–294.
- Cunado, D., Nixon, M., Carter, J., 1999. Automatic gait recognition via model-based evidence gathering. In: Proc. AutoID99: IEEE Workshop on Automated Identification Technologies Summit, pp. 27–30.
- Cutting, J., Proffitt, D., Kozlowski, L., 1978. A Biomechanical Invariant for Gait Perception. *J. Experim. Psychol. Human Percept. Perform.*, 357–372.
- Hayfron-Acquah, J., Nixon, M., Carter, J., 2001. Automatic gait recognition by symmetry analysis. In: 3rd Internat. Conf. on Audio-and-Video-Based Biometric Person Authentication, pp. 272–277.
- Huang, P., Harris, C., Nixon, M., 1999. Human gait recognition in canonical space using spatio-temporal templates. *IEE Proc. Vision, Image Signal Process.*, 93–100.
- Johnson, A., Bobick, A., 2001. A multi-view method for gait recognition using static body parameters. In: 3rd Internat. Conf. on Audio-and-Video-Based Person Authentication, pp. 301–311.
- Little, J., Boyd, J., 1998. Recognising people by their gait: The shape of motion. *Videre, Internat. J. Comput. Vis.* 2 (1), 83–105.
- Murase, H., Sakai, R., 1996. Moving object recognition in eigenspace representation: Gait analysis and lip reading. *Pattern Recognition Lett.* 17, 155–162.

- Murray, M., 1967. Gait as a total pattern of movement. *Amer. J. Phys. Med.* 46 (1), 290–332.
- Nixon, M., Carter, J., Cunado, D., Huang, P., Stevenage, S., 1999. Automatic gait recognition. In: Pankanti, S. (Ed.), *BIOMETRICS—Personal Identification in Networked Society*. Kluwer, Dordrecht.
- Niyogi, S., Adelson, E., 1994. Analyzing and recognizing walking figures in XYT. In: *Proc. Conf. on Comput. Vis. Pattern Recognition*, pp. 469–474.
- Parsons, C., Nixon, M., 1999. Introducing focus in the generalized symmetry operator. *IEEE Signal Proces. Lett.* 6 (3), 49–51.
- Reisfeld, D., Wolfson, H., Yeshurun, Y., 1995. Context-free attentional operators: The generalized symmetry transform. *Internat. J. Comput. Vis.* 14, 119–130.
- Shutler, J., Nixon, M., Harris, C., 2000. Statistical gait recognition via temporal moments. In: *Proc. 4th IEEE Southwest Symp. on Image Anal. Interpretation*, Texas. pp. 291–295.