

ACADEMIC
PRESSAvailable at
www.ComputerScienceWeb.com
POWERED BY SCIENCE @ DIRECT®

Computer Vision and Image Understanding xxx (2003) xxx–xxx

Computer Vision
and Image
Understanding

www.elsevier.com/locate/cviu

Automatic extraction and description of human gait models for recognition purposes

David Cunado, Mark S. Nixon,* and John N. Carter

Department of Electronics and Computer Science, University of Southampton, Southampton SO9 5NH, UK

Received 4 August 2000; accepted 21 November 2002

Abstract

Using gait as a biometric is of emerging interest. We describe a new model-based moving feature extraction analysis is presented that automatically extracts and describes human gait for recognition. The gait signature is extracted directly from the evidence gathering process. This is possible by using a Fourier series to describe the motion of the upper leg and apply temporal evidence gathering techniques to extract the moving model from a sequence of images. Simulation results highlight potential performance benefits in the presence of noise. Classification uses the k -nearest neighbour rule applied to the Fourier components of the motion of the upper leg. Experimental analysis demonstrates that an improved classification rate is given by the phase-weighted Fourier magnitude information over the use of the magnitude information alone. The improved classification capability of the phase-weighted magnitude information is verified using statistical analysis of the separation of clusters in the feature space. Furthermore, the technique is shown to be able to handle high levels of occlusion, which is of especial importance in gait as the human body is self-occluding. As such, a new technique has been developed to automatically extract and describe a moving articulated shape, the human leg, and shown its potential in gait as a biometric.

© 2003 Elsevier Science (USA). All rights reserved.

1. Introduction

Most people can recognise acquaintances by the way they walk, although it is not just their gait that identifies them—for example, their hair style or clothing is usually

* Corresponding author. Fax: +23-8059-4498.

E-mail address: msn@ecs.soton.ac.uk (M.S. Nixon).

recognisable. This research investigates the possibility of recognising people by way of a gait signature as obtained by computer vision. We first review the field of biometrics and present current approaches to gait recognition by computer vision.

There has been considerable study of gait in a number of fields, although there has been little cross-fertilisation of ideas between these fields: psychological gait cues have yet to find deployment elsewhere. Amongst this work, there are emergent techniques aimed at recognising people by their gait. None of these techniques use a known mechanical topology or medical studies, but concentrate more on heuristic and statistical metrics. It is however possible to develop a model-based gait extraction technique from which a metric directly applicable to the mechanics of walking can be generated. Before the performance advantages associated with such an approach are detailed, current research in gait and its allied fields shall be reviewed.

1.1. Biometrics

In today's society, a reliable means of identification is in great demand. Incidents like releasing a child from a day-care centre to a complete stranger or a welfare recipient claiming benefit under six different identities could have been avoided with improved means of identification. One study [29] has shown that one area where technology has enhanced the ability to identify people is biometrics.

A biometric is a measure taken from a living person and used as a method of identity verification or recognition. This measure can be based on some physiological characteristic, like a fingerprint, or some aspect of human behaviour, like handwriting. Currently, the main application for biometrics is in security, for access control. It is anticipated that the stock of potential applications will increase as the techniques mature. Potential future applications could include immigration control and surveillance.

The areas currently under research include automatic face recognition, eye (retina) identification, fingerprints, hand geometry, vein patterns, and voice patterns. Indeed, the use of retinal recognition is under investigation in the UK for biometric use in ATM transactions. Recent surveys [11,32] reveal the immense wealth of research in the field of biometrics, with many applications ranging from banking to security. Even with the vast range of applications and research in the field of biometrics, gait has appeared only recently [26].

In many applications of person identification, particularly those involving serious crime, many established biometrics are obscured. The face may be hidden or at low resolution; the palm is obscured; the ears cannot be seen. However, people need to walk so their gait is usually apparent. This motivates using gait as a biometric. The Oxford Dictionary definition of gait is '*Manner of walking, bearing or carriage as one walks*' suggesting that studies can concentrate on different facets of a person's walk. In fact, a unique advantage of gait is ability to operate at a distance, when other biometrics are of too low a resolution to be perceived. Apart from perceptibility, another attraction of using gait is that motion can be difficult to disguise. Consider for example a robbery situation. The robber will need to make access either quickly, to minimise likelihood of capture, or without being too obvious, in order not to

provoke attention. On escape, the robber will either exit at speed (if the crime is known to have been observed), or in (apparent) leisure. The motion in both cases is natural, for the subject will either not want to draw attention to themselves or will want to move quickly. In this case, using gait as a biometric would appear attractive.

Prior lack of research in gait was perhaps due to technological limitation: analysing gait involves processing a sequence of images and only recently have mainstream architectures and image acquisition been able to offer the requisite performance. Clearly, there are limits to the use of gait as a biometric. A clearer examination of the potency of some apparent limitations awaits development of a practical gait recognition system. However, it is not unlikely that footwear can affect gait, as can clothing. Equally, physical condition can affect gait such as pregnancy, affliction of the legs or foot, or even drunkenness. These factors are not new to biometrics: a face can have make-up or spectacles; ears can be obscured by hair. These can be of natural occurrence, but could also be used for purposes of deception. Deception could go further, a subject can wear a mask and in palm recognition it is of concern that hands can even be cut off. Human perception can be impeded by these factors, but the major question concerns whether a biometric system can still perceive the underlying characteristics of the biometric, regardless of the presence or absence of these factors. In the case of gait, this underlying characteristic can be the musculature which essentially limits the variation of motion. However, factors such as clothing, footwear, and physical condition must await investigation, since methods of recognising the basic nature of gait have yet to become established. Similar to the role of facial expressions or make-up in automatic face recognition, these factors would appear secondary to the principal drive of study of gait as a potential biometric.

Other than its availability, or its naturalness, recognising gait is attractive since it is non-invasive and requires no subject contact, in common with automatic face recognition and other biometrics. Clearly there is a view that gait can be used to recognise individuals. This view is actually not new: Shakespeare used a rich lexicon of adjectives to describe gait, including ‘princely,’ ‘lion’s,’ ‘heavy,’ ‘humble,’ ‘weary,’ ‘forced,’ ‘gentle,’ ‘swimming,’ and ‘majestic.’ Further, in *The Tempest* [Act 4 Scene 1], Ceres observes

‘High’st Queen of state, Great Juno comes; I know her by her gait.’

Even more, in *Troilus and Cressida* [Act 4 Scene 5], Ulysses states

‘Tis he, I ken the manner of his gait; He rises on the toe: that spirit of his in aspiration lifts him from the earth.’

The former is one of Shakespeare’s many observations concerning the ability to recognise people by their gait; the latter includes a concise description of Diomedes’ demeanour. Clearly, Shakespeare’s observations are interpreted through current English usage. However, the earliest dictionary definition of gait suggests a similar meaning, given in Johnson’s *Dictionary of the English Language* [13] as ‘The manner and air of walking’ and the first of the two quotes above is presented as an example of usage.

Accordingly, there appears not only to be a view that gait has potential as a biometric, but also that there is application potential too, as gait appears to have several practical advantages over other biometrics, and some application advantages in common with them. There have been allied studies, particularly those in medical studies for reasons of therapy. Also, there have been psychological studies concentrating on human capabilities. There have also been computer vision approaches aimed to model and to track human targets through a sequence of images, though not usually for recognition. These works are reviewed by [1] and more recently [20]. Tracking approaches are potentially less suitable for gait as a biometric since occlusion (temporal and spatial) can complicate implementation greatly.

1.2. Current approaches to automatic gait recognition

In what was perhaps the earliest approach to automatic recognition by gait, [27] derived a gait signature from the spatio-temporal pattern of a walking person. Here, in the XT dimensions (translation and time), the motions of the head and of the legs have different patterns. These patterns were processed to determine the body motion's bounding contours and then a five stick model was fitted. The gait signature was derived by normalising the fitted model for velocity and then by using linear interpolation to derive normalised gait vectors. This was then applied to a database of 26 sequences of five different subjects, taken at different times during the day. Depending on the values used for the weighting factors in a Euclidean distance metric, the classification rate varied from nearly 60% to just over 80%, a promising start indeed. Clearly, a model was not used within the study. As such, it would appear difficult to confirm that the metric was indeed purely related to a subject's gait.

Later, dense optical flow was used to derive a gait signature [15,16]. This model-free approach developed a description of instantaneous motion, the *shape of motion*, that varied with the type of moving object and type of motion. Applying this description to person recognition, discrimination was accomplished by periodic variations in the shape of their motion. From a sequence of $n + 1$ images, n dense optical flow images are generated. A set of m scalar characteristics (for instance, the centroid of all moving points) that describe the shape of motion are derived from each optical flow image. The values of these scalars are arranged to form m time series that represent the varying motion in the sequence. Each series shares the same fundamental frequency, or a simple multiple, but with varying phase. These phase measurements are the basis of the described gait signature. Phase measurements of different sequences are made comparable by subtracting a reference phase from the phase of each of the remaining scalars, and as such there is no need for temporal alignment. Each sequence is now characterised by $m - 1$ relative phases, forming a feature vector that is used to recognise individuals. Classification was accomplished by computing the mean phase vector for each subject and using it as an exemplar vector. Each sequence is classified by the Euclidean distance of its phase vector to each exemplar vector, the sequence being classified as the class of its nearest-neighbour exemplar vector. The number of features in the feature vector was varied to find the subset of features that would produce the best classification. Experimentation on a database

of six subjects found that with as little as three features a good recognition rate (90.5%) was achieved, plateauing at five features with a rate of 95.2%.

It would appear that the central goal of the work by [16] was to determine what content of motion aids recognition rather than to develop a biometric metric using this approach. Although demonstrating that a model-based approach is not necessary to produce good recognition results, [16] expected that *'discrimination would suffer when the database of individuals became large; simple motion cues could identify types of gait, but would no longer uniquely identify an individual.'*

Another approach [21] was aimed more at generic object–motion characterisation, using gait as an exemplar of their approach. The approach was similar in function to spatio-temporal image correlation, but used the parametric eigenspace approach to reduce computational requirement and to increase robustness. The approach first derived body silhouettes by subtracting adjacent images, with further processing to reduce noise. Then, the images were projected into eigenspace, a well established approach in automatic face recognition. Eigenvalue decomposition was then performed on the sequence of silhouettes where the order of the eigenvectors corresponds to frequency content. Recognition from a database of 10 sequences of seven subjects showed classification rates of 100% for 16 eigenvectors and 88% for eight, compared with 100% for the (computationally much more demanding) spatio-temporal image correlation approach. Further, the approach appeared robust to noise in the input images. Clearly, this is a statistical approach; the system can only learn to classify labelled data. Should the system have been trained erroneously, then the wrong identification will occur since a unique model of human walking has not been assumed, rather that humans can be recognised from walking by the successive distributions in a sequence of images.

Eigenspace transformation (EST) based on Principal Component Analysis (PCA), or the Karhunen-Loeve Transform, has been demonstrated to be a potent metric in automatic face recognition and gait analysis, but without using data analysis to increase classification capability. A new approach combines canonical space transformation based on Canonical Analysis (CA) or Linear Discriminant Analysis (LDA), with the eigenspace transformation, for gait analysis [9]. This gives a 'statistical' (area-based) approach to automatic gait recognition wherein the image sequence is described as a whole, and neither by a model- nor by a motion-based approach, but one which describes the motion content.

Face image representations based on PCA have been used successfully for various face recognition methods. However, PCA based on the global covariance matrix of the full set of image data is not sensitive to class structure in the data. In order to increase the discriminatory power of various facial features, [6] use LDA, also called CA, to optimise the class separability of different face classes and improve the classification performance. The features are obtained by maximising between-class and minimising within-class variations. Unfortunately, this approach has high computational cost. Moreover, the within-class covariance matrix obtained via CA alone may be singular. Combining EST with canonical space transformation (CST) reduces the data dimensionality and optimises the class separability of different gait sequences simultaneously.

On five sequences of five people from the Visual Computing Group, University of California, San Diego [9,15] achieved a classification rate of 85% by CST alone whereas 100% was achieved with combined EST and CST. Further, the clusters were better separated by CST, suggesting that the technique would appear suited to larger databases. A later extension [10] was based on using dense optical flow [16] and again achieved 100% on the same data, showing that the best performance was achieved by using dense optical flow which combined both vertical and horizontal motion components. However, these techniques remain statistical in basis, describing human motion by the statistics of a sequence of area distributions rather than any attempt to match the data to a model.

Recognition of different kinds of gait motion was investigated by [18] and [19]. The trajectories of body parts were extracted from grey-level image sequences, and feature vectors were then derived from this trajectory information. Optic flow was used to determine the trajectories of the body parts in [18], whilst [19] modelled several body parts and the background as mixture densities to extract the trajectories. The feature vectors for the gait analysis were formed from the displacements of the body parts in the x - and y -directions. Classification was performed using Hidden Markov models, one trained for each kind of gait type. In common with all the statistical (area-based) approaches discussed in this section, the features derived for the gait analysis in [18] and [19] are not based on a model of the human body or its motion. As such, the discriminatory powers of the classification will be limited to the quantity and quality of the training data available.

1.3. Aims of the research

This paper explores the possibility of extracting a gait biometric from a sequence of images of a walking subject without using markers. Sophisticated computer vision techniques are developed, aimed to extract a gait signature that can be used for person recognition. A model-based approach is sought to produce a biometric that has high fidelity to the original data. Medical studies are used to develop the model. This is a novel approach to gait recognition by computer vision, as current techniques have concentrated on heuristic and statistical gait metrics. Naturally, the application of a model will potentially alleviate the restrictions imposed on statistical (area-based) approaches, namely that the extracted metric can be directly attributed to human motion. As such, this would appear to offer the potential for discriminating not only gait, but also factors that affect gait.

A fundamental assumption within this work is that a model-based approach should offer suitable potential for automatic gait recognition. Current established approaches to automatic gait recognition are purely statistical: they derive a discriminatory measure that is based on the motion content in a sequence of images rather than a model based on the human body and its motion. The inherent advantages of a model-based approach are the potential ability to handle appearance transformations and practical effects, such as occlusion. Appearance transformations imply that an object's shape will be distorted by the camera's viewpoint. This can only be handled in area-based approaches by inclusion of marker points in each scene. It is likely

that a model-based method can handle such a distorted scene without marker points since it relies on the presence of human motion in the sequence, and as such can inherently model its time history/future.

Furthermore, it would appear impossible for an area-based technique to explicitly handle occlusion. The discriminatory measure derived in statistical techniques is formed directly from the motion information of pixels in the image sequence. In the case of an occluded object, where the motion information of the pixels has been removed in some parts of the sequence, an area-based approach could fail since there is no model with which to handle the missing information. A model-based technique can overcome the reduced information by globally integrating the information gathered locally in each image of the sequence. In this way, the importance of the information in each individual image is dramatically reduced, so much so that an object can be missing/occluded in several frames without affecting the extraction of the signature. The ability to handle occlusion successfully is fundamental to automatic gait recognition. In practical applications the person can be occluded by accessories such as briefcases, umbrellas, shopping bags or rucksacks. Also, the human body is self-occluding in many of its motion types.

Similarly, secondary effects such as footwear and apparel cannot be analysed by statistical techniques. The description of motion in the image sequence produced by the statistical (area-based) methods cannot be extrapolated to include these effects as the description has no basis in the mechanics of human motion. With a model-based approach, these secondary effects can potentially be accommodated in the feature extraction stage. For instance, a tired person will walk with less flexion of the legs, but nevertheless the image sequence will still contain a moving person and as such its model can still be extracted.

Initially, extant computer vision techniques were used to extract a gait biometric from a sequence of images [5]. Although a gait signature was successfully generated from the extracted features, there were several strategic problems. Essentially, the use of extant techniques did not consider the potential performance gains possible with sequence-based processing of moving features, being more concerned with a frame-by-frame approach to moving feature tracking. New modelling techniques were then developed to solve these problems. Using novel temporal evidence gathering methods, a gait model can be extracted from an image sequence and a signature based on this model generated totally automatically. We could have improved matters by using one of the new approaches for background subtraction to isolate moving subjects before analysis but instead we chose to phrase our new approach as a conventional approach to feature extraction, here targeted at extracting and describing moving shapes, from sequences of images.

Previous medical studies on gait are reviewed in Section 2. New techniques developed to overcome the inherent problems discovered in the preliminary study are detailed in Section 3. Section 4 outlines possible further work that would extend the techniques discussed in the previous section. The conclusions in Section 5 summarise the work presented in this paper and describe how well the aims of the research have been met.

2. Modelling human gait

2.1. Medical studies

The goal of most gait research has been to classify the components of gait. [23] and [22] produced standard movement patterns for normal people that were compared to the gait patterns for pathological patients [22]. No statistical or mathematical analysis was performed on the collected data.

Gait was considered by [22] as ‘a total walking cycle’—the action of walking can be thought of as a periodic signal, with an associated frequency spectrum. [2] examined the frequency content of kinematic gait data, and showed that normal gait had a maximum frequency component of 6 Hz. As such, kinematic gait data could be represented by a limited number of frequency components.

2.2. Gait description

The following terms are used to describe the gait cycle, as given in [23]. Fig. 1 illustrates the terms described. A *gait cycle* is the time interval between successive instances of initial foot-to-floor contact (‘heel strike’) for the same foot. Each leg has two distinct periods; a *stance* phase, when the foot is in contact with the floor, and a *swing* phase, when the foot is off the floor moving forward to the next step. The cycle begins with the heel strike of one foot, the left foot for example. This marks the start of the stance phase. The ankle flexes to bring the left foot flat on the floor (‘foot-flat’) and the body weight is transferred onto it. The right leg swings through in front of the left leg as the left heel lifts off the ground (‘heel-off’). As the body weight moves onto the right foot, the supporting left knee flexes. The remainder of the left foot, which is now behind, lifts off the ground ending the stance phase with toe-off.

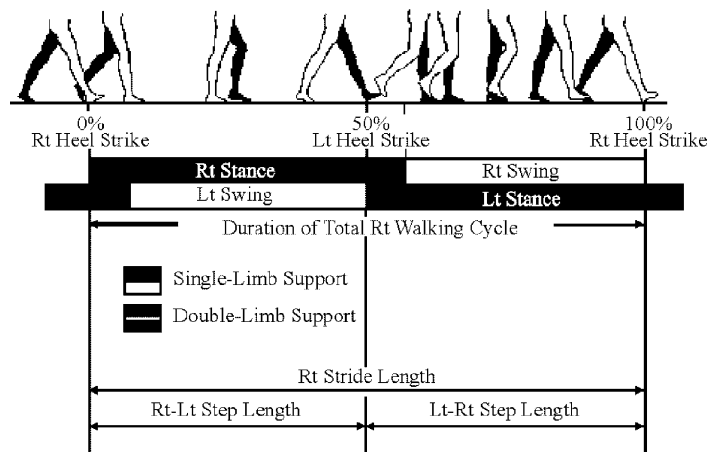


Fig. 1. Relationship between temporal components of the walking cycle and the step and stride lengths during the cycle.

The start of the swing phase is when the toes of the left foot leave the ground. The weight is transferred onto the right leg and the left leg swings forward to strike the ground in front of the right foot. The gait cycle ends with the heel strike of the left foot.

Stride length is the linear distance in the plane of progression between successive points of contact of the same foot. *Step length* is the distance between successive contact points of opposite feet. A *step* is the motion between successive heel strikes of opposite feet. A complete gait cycle is comprised of two steps.

2.3. Characteristics of human gait

From the work carried out by [23] and [22] it can be concluded that if all gait movements were considered, gait is unique. In all there appear to be 20 distinct gait components, some of which can only be measured from an overhead view of the subject.

Murray [22] found ‘the pelvic and thorax rotations to be highly variable from one subject to another.’ These patterns would be difficult to measure even from an overhead view of the subject, which would not be suited to application in many practical situations. [22] also suggested that these rotation patterns were not found to be consistent for a given individual in repeated trials. As such, they would not appear suited to an automated computer vision-based biometric system.

In [23] and [22], ankle rotation, pelvic tipping, and the spatial displacements of the trunk (vertical oscillation, lateral oscillation, and forward displacement) were shown to possess individual consistency in repeated trials. Naturally, given the resolution of most general purpose cameras, the ankle is difficult to extract consistently, let alone its rotation. Equally, the pelvis can easily be obscured by clothing, making a measurement of its inclination easily prone to confusion and error. The spatial displacements of the trunk are measured from the neck. Thus, the extraction of these displacements is hampered by a difficulty in finding the neck consistently and by the fact that the neck, like the pelvis, is easily obscured by clothing. As such, these components would be difficult to extract accurately from real images. Again, these would appear unsuited to an automated system.

Since many features established by medical studies appear unsuited to a computer vision-based system, the components for this investigation have been limited to the rotation patterns of the hip and knee. These patterns are possible to extract from real images and, from medical studies, appear to possess a high degree of individual consistency and inter-individual variability.

2.4. Rotation pattern of the hip

Fig. 2 shows the rotation angles for the hip and knee, as measured by [22]. The normal hip rotation pattern is characterised by one period of extension and one period of flexion in every gait cycle. Fig. 3 gives the average rotation pattern as presented by [22], generated from a database of 60 subjects. The upper and lower dotted lines indicate the standard deviation from the mean. In the first half of the

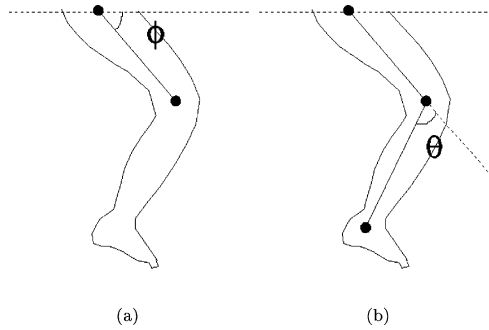


Fig. 2. Hip (a) and knee (b) rotation angles.

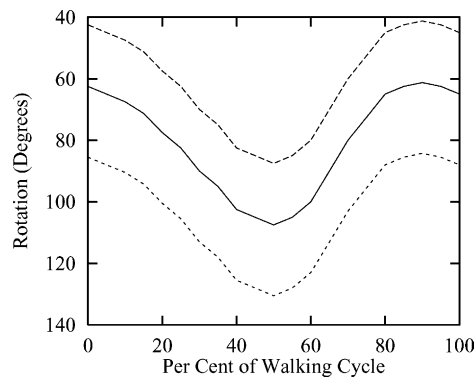


Fig. 3. Mean hip rotation pattern [22].

gait cycle, the hip is in continuous extension as the trunk moves forward over the supporting limb. In the second phase of the cycle, once the weight has been passed onto the other limb, the hip begins to flex in preparation for the swing phase. This flexing action accelerates the hip so as to direct the swinging limb forward for the next step. The angle of rotation is measured as the angle between the line joining the hip and knee, and the line passing through the hip point parallel to the ground.

2.5. Model of legs for gait motion

The potential of the periodic nature of gait for an analytic approach was first investigated by [14], who performed a feasibility study into using gait as a biometric. An analytic approach was used, describing the legs and the motion of walking as a model based on medical and perceptual studies. The human leg was modelled as two pendula joined in series (Fig. 4). The upper pendulum modelled the thigh and was suspended between the hip and the knee. The lower pendulum modelled the lower leg suspended from the knee to the ankle. This pendulum model is backed by [22]—‘for normal gait, the duration of successive temporal components and the length of successive steps are rhythmic.’

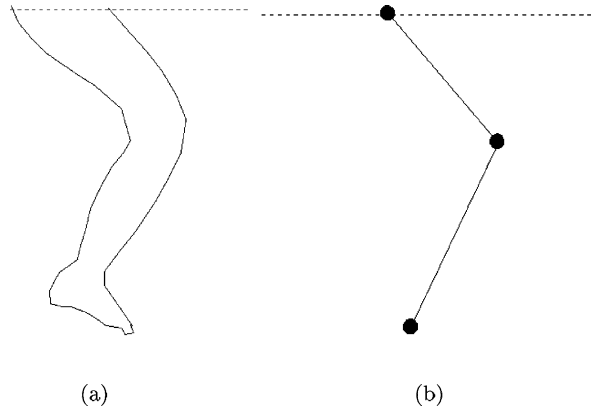


Fig. 4. (a) Leg outline. (b) Pendulum model of a leg.

Kuan [14] extracted the hip rotation pattern for three subjects from a sequence of images using computer vision techniques. The rotation patterns were curve fitted manually to in-fill for missing data points. Fourier analysis was performed on the rotation patterns, and the magnitude and phase spectra for each subject were examined. The magnitude plots showed some variation between subjects, whilst the phase plots exhibited greater variation between subjects. The greater inter-individual variation of the phase spectra makes the phase information an attractive measure for recognition. The manner in which the hip inclination changes is of as much interest as the actual angle itself, and as such both the magnitude and phase information were of use. Kuan concluded that this model-based approach looked promising, but as yet insufficient for an automated non-invasive technique.

The frequency information of hip rotation appears to be attractive for recognition as different walking speeds can be dealt with effectively. When fast walking was considered by [22] it was found that the increased ground speed was effected by an increase in stride length and reduction in gait period. Both the hip and knee exhibited greater extensions and flexions throughout the gait cycle in fast walking. The frequency scaling property of the FT predicts that the effect of a reduction in the period of the cycle is an expansion in frequency of the spectrum;

$$f(at) \stackrel{\mathcal{F}}{\leftrightarrow} \frac{1}{|a|} X\left(\frac{j\omega}{a}\right), \quad (1)$$

where $X(j\omega)$ is the FT of $f(t)$, and a is a real constant. Given the relationship shown above, the spectrum for fast walking might demonstrate the difference in speed by appearing to be a stretched version of the free walking spectrum, keeping a characteristic envelope for a given individual at different walking speeds—effectively, a gait signature. The scaling effect of the spectrum for different walking speeds can be eliminated by normalising the time axis for the hip rotation patterns to the length of gait cycle in each case; the time scale will cover from 0% to 100% of the walking cycle. In this way, direct comparisons between spectral data can be made, as the

Fourier information is now independent of the fundamental frequency of the gait cycle.

This study uses the model presented in [14] as the leg motion is periodic and each part of the leg (upper and lower) appears to have pendulum-like motion. Fourier theory allows periodic signals to be represented as a fundamental and harmonics—the gait motion of the lower limbs can be described in such a way. Accordingly, computer vision techniques were required to extract the measurements from video sequences in an automatic manner to allow the investigation of remote gait measurement as a potential biometric. This involved a re-examination of the basic processes by which the measurements were derived, and new development thereof.

3. Gait signature by evidence gathering

The preliminary study described in [5] demonstrated positive results in the use of gait as a biometric measure. A gait signature was extracted, but not automatically, using computer vision techniques and produced a high correct classification rate on a small database of subjects. However, the techniques used in [5] had inherent problems which would be likely to affect more general use of the technique, especially on a larger database. These are discussed in this section. Novel vision techniques have been developed specifically to overcome these problems and their motivation, as well as their implementation, are described. Further experimentation shows how these novel techniques can extract and describe gait, and results are presented showing how they can be used to recognise people by their gait. Furthermore, the new technique extracts the gait signature automatically from the image sequence, without human intervention, one of the major aims of this work.

3.1. Previous work

The study performed in [5] described a novel model-based approach to gait recognition, using the notion of gait as a periodic signal to create a gait signature. The lower limbs were modelled as two inter-connected pendula and gait was considered as the motion of these pendula.

Using image processing techniques, lines representing legs in a sequence of images were extracted using the standard Hough transform (SHT). The inclination of the line representing the leg in each frame was collated to create the hip rotation pattern for the subject. In-filling for missing data points was done by least-squares analysis of the collected data points to an eighth-order polynomial. This could, in hindsight, have been achieved by Fourier interpolation. Fourier analysis was performed on the extracted hip rotation pattern using the Discrete Fourier transform to find magnitude and phase information. The magnitude data and the phase-weighted magnitude data were classified using the k -nearest neighbour rule. The phase-weighted magnitude data was found to give a better correct classification rate than just the magnitude data.

This model-based approach uniquely gave a signature which could be directly related to the original image sequence. However, images were analysed singly, without reference to the whole sequence. As such, the technique would be unable to handle occlusion, except by interpolation such as by least squares.

Although the extracted hip rotation patterns in [5] concurred with those presented in medical research [22,23], the idea of gait as a periodic function was not reflected in the use of a polynomial to model the motion of the thigh. Any periodic signal, with period T , can be represented by a Fourier series (FS). The motion of the thigh is better represented as an FS rather than by polynomial fitting to the extracted data. Also, greater noise immunity can be achieved when extracting temporal features in a sequence of images by including the entire sequence in the evidence gathering process. [24] described a Velocity Hough transform (VHT) technique that enables the concurrent determination of structural and motion parameters of moving parametric shapes in an image sequence. Essentially, the VHT includes motion within the parametric model. The polar representation of a circle radius r with co-ordinates x_0, y_0 in the first frame and moving with horizontal and vertical velocity, v_x and v_y , respectively, has x and y co-ordinates at time t as

$$x(t) = x_0 + r \cos(\theta) + tv_x, \quad (2)$$

$$y(t) = y_0 + r \sin(\theta) + tv_y, \quad (3)$$

where θ is an index to points on the circle's perimeter. Votes are accumulated in a 5-dimensional accumulator (x_0, y_0, v_x, v_y, r) from edge images of each image in the sequence. The accumulator maximum corresponds to the best estimates of each of the parameters describing the moving circle. The VHT was demonstrated to have superior properties in low to moderate noise compared with a separately applied HT.

By combining VHT techniques with the FS representation of the hip rotation, a feature-based human gait model can be extracted from a sequence of images. This feature-based model has a high fidelity to the data, with a clear analytic justification. Evidence gathering using the VHT offers greater immunity to noise and occlusion, and produces a maximum likelihood estimate of the model parameters. By modelling the hip rotation as an FS, the gait signature described in [5] can be extracted directly, without intervention.

3.2. Modelling gait

The VHT is ideally suited for gait analysis as it considers both structural and motion information simultaneously. As such, features with a specified structure and demonstrating a particular type of motion can be extracted from a sequence of images. Other work by [25] has extended the VHT to find moving articulated objects, providing an excellent feature extraction primer for statistical approaches to gait recognition, since the technique isolates the moving area of interest. The following sections describe a gait model that can provide the biometric signature described in [5] directly from the evidence gathering process. For a given feature, gathering evidence using VHT techniques requires a structural motion model.

3.2.1. Motion model of the thigh

The structural model of the thigh presents the edge point values of the model at any time instant. Fig. 5 shows the model of the thigh used for temporal evidence gathering. The motion of the model of the thigh is described by the displacement functions $c_x(t)$ and $c_y(t)$, which control the movement of the hip, and the inclination $\phi(t)$ which controls the thigh rotation. [23] state that ‘during walking the pelvis remains relatively level.’ Therefore, the vertical velocity $v_y(t)$ is described by the average vertical velocity V_y only,

$$v_y(t) = V_y, \quad (4)$$

where V_y reflects the constant slope of the walking surface. The vertical displacement $c_y(t)$ of the hip can be modelled as

$$c_y(t) = V_y t. \quad (5)$$

The horizontal motion of the pelvis has two main influences: the average ground speed and the transverse rotation (about the y -axis) of the hip. The transverse rotation of the hip is characterised by the graph in Fig. 6, which was obtained from [22]. Even though the mean is not zero, reflecting bias on the measurement process, the entire cycle can be observed to be sinusoidal in nature. Accordingly, using a simple cosine to represent the influence of the transverse rotation, the horizontal velocity $v_x(t)$ can be described by

$$v_x(t) = V_x + A \cos(\omega_0 t + \psi), \quad (6)$$

where V_x is the average velocity, A is the amplitude, and ψ is the phase of oscillation, respectively, and ω_0 is the angular velocity of the gait cycle. Using standard trigonometric relations, $A \cos(\omega_0 t + \psi)$ can be represented in a more simple form as:

$$A \cos(\omega_0 t + \psi) = A \cos \omega_0 t \cos \psi - A \sin \omega_0 t \sin \psi \quad (7)$$

$$= \alpha \cos \omega_0 t - \beta \sin \omega_0 t, \quad (8)$$

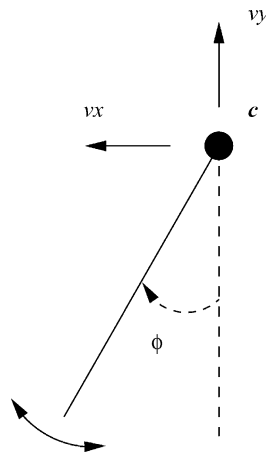


Fig. 5. Model of thigh for temporal evidence gathering.

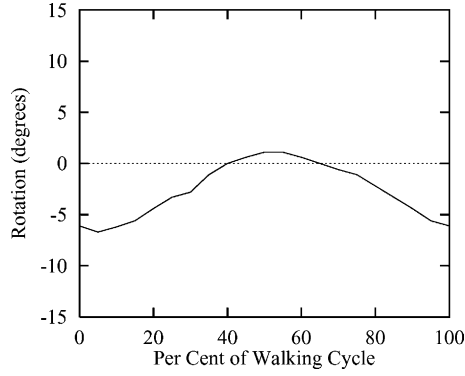


Fig. 6. Transverse pelvic rotation [22].

where

$$\alpha = A \cos \psi \tag{9}$$

and

$$\beta = A \sin \psi. \tag{10}$$

This actually simplifies later parameterisation in evidence gathering as the quantisation of the phase parameter $\psi \in [-\pi, \pi]$ is now spread over two magnitude parameters. These magnitude parameters are more easily quantised as their effect on the displacement $c_x(t)$ is measured in units of pixels. Integrating Eq. (8) gives the horizontal displacement of the hip, the constant of integration set to guarantee the initial position c_{x0} ($c_x(0) = 0$) at $t = 0$. The function $c_x(t)$ is given by

$$c_x(t) = -\frac{\beta}{\omega_0} + \left(V_x t + \frac{\alpha}{\omega_0} \sin \omega_0 t + \frac{\beta}{\omega_0} \cos \omega_0 t \right). \tag{11}$$

3.2.2. Structural model of the thigh

The thigh is described by a point \mathbf{c} that represents the hip joint, or pelvis, and a line passing through \mathbf{c} at an angle ϕ . The pelvis has horizontal and vertical velocities v_x and v_y , respectively. These velocities produce the displacement functions $c_x(t)$ and $c_y(t)$ from the initial coordinates of the hip (c_{x0} , c_{y0}), as described in Section 3.2.1. The inclination of the line, ϕ , is governed by the hip rotation model. The coordinates of points on the moving line are given by

$$r_x = c_{x0} + c_x(t) - \lambda \sin \phi(t), \tag{12}$$

$$r_y = c_{y0} + c_y(t) + \lambda \cos \phi(t), \tag{13}$$

where c_{x0} and c_{y0} are the initial coordinate values of the hip and λ can take any real value to produce any point on the line. As in the VHT, Eqs. (12) and (13) are used to map edge points in the image space into an accumulator space whose axes are the parameters of interest, namely c_{x0} , c_{y0} , V_x , V_y , α , β , a_0 , b_k , and c_k . Edge points from the

entire sequence are used, the mapping equations reflecting the time instance of the frame in which the edge points are in. Related edge points in different frames cast votes for the same parameter values, the ones relating them together.

3.2.3. Creating a gait signature

For a constant walking speed, the hip rotation $\phi(t)$ is a periodic function with period T . An FS can represent any periodic signal with fundamental frequency $\omega_0 = 2\pi/T$. For a real periodic signal, the FS representation can have the form

$$x(t) = a_0 + 2 \sum_{k=1}^{\infty} \Re\{a_k e^{j\omega_0 k t}\}. \quad (14)$$

The function $\phi(t)$ in Eqs. (12) and (13) can be modelled by the FS in Eq. (14). As such, the inclination of the thigh is represented by a series of harmonics, as consistent with observations from earlier medical studies [2] and the earlier model-based approaches to gait description. Also, Fourier series expansions of experimental data of actual human locomotion have been used as a basis in modelling human figure behaviour with emotions for animation purposes [30]. The method developed in [30] defined a functional model based on the Fourier series expansion of joint data from actual human behaviour. An extensive variation of movements could be generated from this functional model. For instance, using the functional models for a normal walk and a ‘tired’ walk, a parameter $s \in [0, 1]$ is used to interpolate between the two models to produce a walk with different degrees of ‘tiredness.’ Expressing exaggerated behaviours is achieved through extrapolation ($|s| > 1$) of the two models. Similarly to [2], [30] found that the number of harmonics required to model the joint motions were usually three, and at most seven—‘human locomotions are characterised with a small number of Fourier coefficients.’

Using the rectangular form of $a_k = b_k + jc_k$ in Eq. (14) gives an expression for $\phi(t)$;

$$\phi(t) = a_0 + 2 \sum_{k=1}^N [b_k \cos k\omega_0 t - c_k \sin k\omega_0 t], \quad (15)$$

where N is the number of harmonics. This equation reflects the periodic nature of the hip rotation.

In [5] a gait signature was constructed from the frequency components of the hip rotation. As this is a periodic signal, its Fourier Transform (FT) is a train of impulses at the harmonics of the fundamental frequency. The amplitude of these impulses is given directly from the coefficients a_k of the FS representation:

$$X(j\omega) = \sum_{k=-\infty}^{\infty} 2\pi a_k \delta(\omega - k\omega_0). \quad (16)$$

The FT of any function is a complex number that can be split into magnitude and phase components using its real and imaginary parts. The FS coefficients a_k are the only complex numbers in Eq. (16). Using the form $a_k = b_k + jc_k$, $X(j\omega)$ can be written in terms of b_k and c_k ;

$$X(j\omega) = \sum_{k=-\infty}^{\infty} 2\pi\delta(\omega - k\omega_0)[b_k + jc_k]. \quad (17)$$

From Eq. (17), the magnitude and phase components of the hip rotation pattern can be found from the FS coefficients b_k and c_k ;

$$|X(j\omega)| = 2\pi \sum_{k=-\infty}^{\infty} \delta(\omega - k\omega_0) \sqrt{b_k^2 + c_k^2}, \quad (18)$$

$$\arg(X(j\omega)) = \sum_{k=-\infty}^{\infty} \delta(\omega - k\omega_0) \tan^{-1} \left(\frac{c_k}{b_k} \right). \quad (19)$$

The feature extraction process generates values for the FS coefficients a_0 , b_k , and c_k . As such, using Eqs. (18) and (19), the gait signature can be directly extracted from the evidence gathering method: there is now no need to apply separate Fourier analysis. Accordingly, the frequency-based signature is derived directly from the whole image sequence using a model-based approach, avoiding the problems inherent in the earlier formulation.

3.3. Implementing a VHT to extract a gait model

3.3.1. Reduced gait model extraction by VHT and SHT

A VHT was implemented to extract the gait model described in Section 3.2. Each edge point in each frame of a sequence produced a set of possible initial coordinates (c_{x0}, c_{y0}) for a given range of values of the remaining parameters. Eqs. (12) and (13) produced this set of initial coordinates. The functions $c_x(t)$, $c_y(t)$, and $\phi(t)$ (Eqs. (11), (5) and (15), respectively) in Eqs. (12) and (13) were formed using the values of the remaining parameters, given the time t of the frame of the current edge point. The cells in the accumulator space that corresponded to each parameter combination for the given edge point were incremented. The axes of the accumulator space were the parameters of the model.

Angeloni et al. [2] showed that the magnitude of the frequency information of the hip rotation pattern had no significant terms above 5 Hz. As such, a full description of the hip rotation can be achieved with a fifth-order FS. The full gait model would have the 11 parameters for the hip rotation model, six for the pelvic model and a final parameter for the gait cycle period T . Implementing a VHT to extract this model would require an 18-dimensional accumulator space. With current technology, this would be impractical to create and search, even for a small number of possible values for each parameter. Table 1 lists the parameters along with the equations in which they are used. The hip has average velocities V_x and V_y in the x and y directions, respectively. The horizontal motion of the hip has a sinusoidal influence whose magnitude and phase is controlled by parameters α and β (Section 3.2.1). The hip rotation is modelled by an FS with period T and harmonic values of a_0 , b_k , and c_k , where $k = 0, 1, \dots, N$ and $N \in [1, 5]$ is the order of the FS.

Table 1
Parameters of gait model and the equations they are used in

Parameters	Equation
V_y	$c_y(t) = V_y t$
V_x, α, β	$c_x(t) = -\frac{\beta}{\omega_0} + \left(V_x t + \frac{\alpha}{\omega_0} \sin \omega_0 t + \frac{\beta}{\omega_0} \cos \omega_0 t \right)$
c_{x0}	$r_x = c_{x0} + c_x(t) - \lambda \sin \phi(t)$
c_{y0}	$r_y = c_{y0} + c_y(t) + \lambda \cos \phi(t)$
a_0, b_k, c_k	$\phi(t) = a_0 + 2 \sum_{k=1}^N [b_k \cos k\omega_0 t - c_k \sin k\omega_0 t]$
T	$\omega_0 = \frac{2\pi}{T}$

For the purposes of testing, a reduced gait model was used; a first-order FS was used for the hip rotation model, reducing the model to 10 parameters. [7] described a method to reduce the storage space required by conventional Hough transforms that is directly applicable to the VHT. The accumulator space can be reduced to two 2-dimensional spaces, one representing the $c_{x0}c_{y0}$ accumulator and the other for data storage. The range of values for the remaining ‘search’ parameters were combined exhaustively. For a given combination of these parameters, each frame in the sequence was then analysed in turn. Edge points from each frame incremented cells in the 2D accumulator indexed by values for the initial coordinates (c_{x0}, c_{y0}) for the current combination of the remaining parameters. For any combination of the search parameters an edge point, at a given time instance, is mapped onto a line in the $c_{x0}c_{y0}$ -plane. Bresenham’s line drawing algorithm [4] was used to cast the votes in the accumulator space. Once all the frames had been processed, the accumulator space was searched for a peak. A second 2D accumulator was used to store the peak value and the values of the parameters that related to it. The peak in the first space was compared with the peak in the second at the same coordinates (c_{x0}, c_{y0}) . If the new peak was greater, the parameter information and peak height in the second accumulator space was updated to that of the newer peak. The approach naturally selects the largest object, as usual in evidence gathering, though the accumulator space could be weighted so as to cause selection of a different feature. Note that with this technique, instances of the model that have the same initial coordinates (c_{x0}, c_{y0}) cannot be found, as one will overwrite the others (though this can be avoided in implementation). This method offered a reduction in storage space of the accumulator, from a 10-dimensional space to two 2-dimensional spaces (for the reduced model). There was negligible effect on the computation time since to create the first 2D accumulator space the 10-dimensional parameter space must be exhaustively searched.

In order to provide a basis for comparison, a technique based on extant work was used. As described in [24], the VHT demonstrates greater noise immunity compared with the traditional temporal feature extraction method, the Hough Transform Tracking (HTT) method. For a given feature, the HTT method applies the SHT

to each frame in a sequence to extract the static parameter values the feature at the time instances given by each frame. The temporal parameters are calculated from the static parameters by interpolation. The next section compares the performance in noise of the VHT and HTT when applied to extract the human gait model described earlier.

3.3.2. Performance in noise

A sequence of synthetic images was produced using the reduced gait model, and a VHT was implemented to extract the gait model parameters. An example is shown in Fig. 7(a) where the long line indicates the position of the simulated leg. The short line is for display only, depicting the position of the hip. Varying amounts of uniformly distributed, black and white noise were added to the sequence, and the model was extracted using both VHT and HTT techniques. The noise was added randomly to each image in the sequence such that each image had the same amount of noisy points, but in different positions. Fig. 7 shows examples of uniformly distributed, black and white noise added to an example image with levels of 14% and 42%. In the latter case it is impossible to determine the line in the static image by human vision.

The VHT to extract the reduced human gait model was implemented as described in Section 3.3.1. A restricted range of values for each parameter was used in order to maintain the computational time at an acceptable level. The HTT was performed by applying the SHT to each image in a given sequence, using established performance constraints [31]. The parameters ρ and ϕ of the most prominent line extracted from each frame, along with the time value t_i for that frame, were used to calculate the parameters of the human gait model. Least-squares analysis was used to fit this data to the model.

The extracted model parameters were compared against the actual values of the parameters used to produce the synthetic images, giving a Euclidean distance metric to quantify the performance of the extraction technique. Since all of the parameters

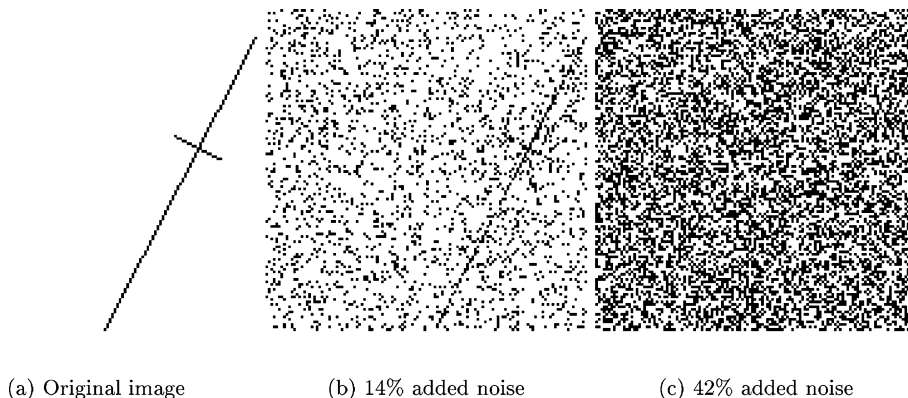


Fig. 7. Demonstrating uniformly distributed, black and white noise added to an image.

did not have the same quantisation in the VHT implementation, the Euclidean distance metric could not be calculated merely as the root of the sum of squared differences in each set of extracted parameters. For example, a difference of unity in the value of initial x coordinate c_{x0} would not be as significant as that same difference in the value of harmonic weight b_1 . As such, the difference for each parameter value was normalised by dividing it by its quantisation value as used in the VHT. Therefore, the Euclidean distance D between each model instance is given by

$$D = \sqrt{\sum_{i=0}^{P-1} \left(\frac{x_i - y_i}{q_i} \right)^2}, \quad (20)$$

where P is the number of parameters in the human gait model, x_i and y_i represent the i th parameter as extracted by the VHT and HTT, respectively, and q_i is the quantisation of the i th parameter. Table 2 lists the quantisation values for each parameter. Note that these values were determined through experimentation to give the most satisfactory results. In this way, the Euclidean distance D given by Eq. (20) represents the distance in accumulator space. For instance, a value of $D = 50$ for two given extracted models implies that their parameters were an average distance of 50 cells apart in the accumulator space. Using the quantisation values given in Table 2, this distance $D = 50$ produces a 50 pixel difference in values for the initial x -coordinate c_{x0} , or a difference of 0.5 radians (28.6°) in the harmonic weight b_1 .

Fig. 8 shows the performance of the VHT and the HTT techniques in extracting the model parameters from the synthetic images with noise levels of 2%, 5%, and then from 10% up to 42% in 2% steps. As Fig. 8 demonstrates, the VHT moving feature extraction method found the exact parameter values for the gait model even in noise levels of 42%, wherein the static image of Fig. 7(c) the line could not be seen. The performance of the HTT deteriorated at around the 30% level of noise. The extracted parameter values for the gait model were found to be progressively further away from the actual values as the noise level was increased above 30%. The extracted model parameters for the 34% case were found to be quite far from the actual values, compared to the distances for noise levels immediately above and below. This appears due to the random process in which the noise was added, leading to there being a line extracted in one of the images that was sufficiently out of place to effect the least squares analysis to produce more distant parameter values. Similarly, the comparatively small distance for the 38% case was due to the added noise affecting the line extraction to find lines not too distant from the original. The least squares analysis was able to produce parameter values that were not as erroneous as for

Table 2
Quantisation values used in the VHT and GA implementations

Parameters	Quantisation value q_i
$V_y, V_x, \alpha, \beta, c_{x0}, c_{y0}$	1 (pixel)
a_0, b_k, c_k	0.01 (radians)
T	0.04 (s)

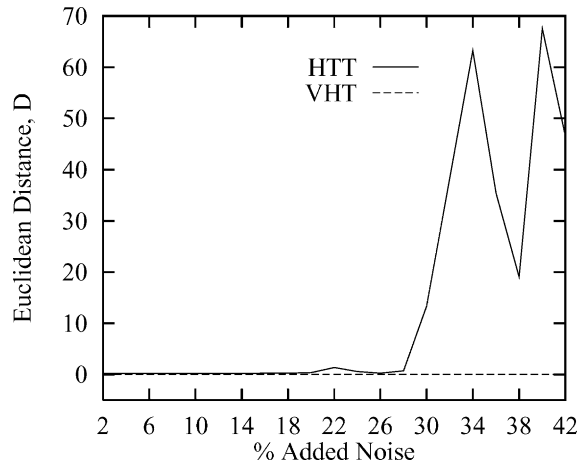


Fig. 8. Results of noise trials for VHT and HTT moving feature extraction techniques.

the surrounding levels of noise. This demonstrates the dependence of the HTT method on the structural information in every the frame of a sequence—one image with missing data will mislead the extraction of the model parameters. It can also be viewed to demonstrate practical advantage since the noise gives a varying background. Clearly, the VHT technique focuses better on the moving object and is less distracted by background information.

Like all Hough-based processes, the VHT suffers from a high computational complexity and storage cost. For problems of low dimensionality, Hough-based approaches to feature extraction are quick and precise, but it becomes rapidly unfeasible to create and search an accumulator as the number of parameters to estimate increases. Using the technique for fast contour recognition presented in [7], the storage and peak detection problem was reduced to that of two 2-dimensional accumulator spaces (see Section 3.3.1). The computational complexity remains as an exhaustive search of the parameter space was still required.

3.3.3. Implementation by genetic algorithm

Extracting model instances from images is effectively a problem in optimising the defining function of the model for a given set of edge points. A Genetic Algorithm (GA) [8] is an optimisation method that is shown to consistently outperform many other search methods in solving hard optimisation problems. For a satisfactory sized parameter space, the VHT implementation took the order of days to run on a P75 PC. The GA based VHT to extract the human gait model takes approximately 20 min on the same parameter space, offering a speed-up factor of approximately 100. As such, all further experimentation was performed using a GA based implementation.

In this GA implementation, each individual's chromosome was the binary coding of the parameters of the gait model. Each parameter was represented by n bits which gave an integer index to a position within a specified range. Naturally, the value of n

controlled the resolution for each parameter. The fitness was derived from the number of edge points matching those calculated for the (moving) template described by the current values within the chromosome. The fittest individuals were selected as those that had a greater probability that a spin of a biased roulette wheel would select them. Crossover was set to occur with a probability of 0.7 and mutation with a probability equaling the reciprocal of the maximum population (the effect of which was to complement the mutated bit). New generations were evolved until the stopping criterion was satisfied where the stopping criterion was either when the average fitness had evolved to 95% of the maximum fitness in the population, or when 200 generations had been reached. The final values provided by the GA were used to initiate a fine-grained VHT search.

3.4. Performance factor analysis

Due to the random nature of a GA, repeated trials of a GA on the same data set are not guaranteed to produce the same result. This problem is further compounded if the solution space is complex, containing many local peaks and valleys. Similarly, if the problem is unimodal with a large peak compared to the surrounding space, a GA will also have problems in finding consistent solutions. With the SHT for lines and circles, the complex solution space can be smoothed by correct quantisation of the parameters [31]. Even with this smoothing of the parameter space, it has not been possible to extract consistent results when using a GA to find geometric primitives in images [17]. In the work by [17], trials were repeated 100 times for a given image and the parameters of the extracted shape were calculated as an average of the repeated trials.

Due to the high dimensionality of the gait model, it is not possible to find optimal quantisation values for the various parameters with the methods used by [31]. Also, averaging the parameters determined in repeated trials [17] would not produce a viable solution. Due to the complexity of the solution space, the GA is likely to find one of the many false peaks that are due to anomalies present over the entire sequence, in at least one of the repeated trials. Accordingly, averaging a set of parameter values would be detrimental in finding an accurate solution.

Also, increasing the ‘run time’ of the GA with the intention of finding even fitter solutions would not be adequate. The increased run-time could be achieved by extending the termination criteria—increasing the proportion of the highest fitness value on the population that the average fitness of the population must reach in order for convergence, and hence termination, to occur. The effect of this would be an increase in the number of solutions with the highest fitness value, rather than an increase in the highest fitness value. The genes of the fittest chromosome would be more thoroughly propagated throughout the population, and the decreasing probability of mutation p_m would ensure no deviation from these genes. In this way, if the GA finds a false peak it will remain there and that solution would spread throughout the population. As such, the method that has been employed has been to run 10 trials on each sequence, with the same set of parameters, and select the trial with the highest fitness as the solution for that sequence.

3.4.1. GA performance for occluded features

A characteristic of the VHT is that it can extract temporal features in scenes where the feature has been occluded. This is attractive in extracting human gait models as the human body is self-occluding in almost all its motions. As such, the GA based VHT for gait analysis, GAVHT, was tested for the presence or absence of this characteristic, at increasing levels of occlusion. The experimentation was performed on sequences of edge images of walking subjects. The Canny edge detector with hysteresis thresholding was used to produce the edge images from the greylevel sequences. The reduced human gait model (Section 3.3) was used in this performance experiment. The parameters of the model were quantised with the values given in Table 2. Also, the GA was implemented with a chromosome population of 500.

A major advantage of the VHT method of moving feature extraction over the traditional static, frame-by-frame moving feature extraction is its improved performance in sequences where the feature has been occluded in some frames. As such, an analysis of the GAVHT's performance in extracting the gait model for an occluded subject was performed to verify that this occlusion immunity characteristic of the VHT extraction method had been maintained in its translation to a GA. Fig. 10 shows the extracted model for the unoccluded sequence for subject CM using the GAVHT. The model used only a first-order FS to represent the hip rotation.

As Fig. 10 shows, the inclination of the hip is tracked throughout the cycle although without much precision. This was due to the low order of FS used to model the motion of the hip rotation, but for the purposes of this exercise it was sufficient.

The subject CM in Fig. 10 was occluded by simulating a column in the centre of the field of view of the camera, between the subject and the camera. The illusion was created by setting the pixels with the column to zero (black). Fig. 9 shows an example of this process, with the column width set to 90 pixels, representing almost two fifths of the image width.

A series of tests was then conducted on the sequence in Fig. 10. The same parameterisation and initialisation was used in all the tests, with the variable parameter between each sequence being the width of the occluding column. The width was varied from 10 pixels to 150 pixels in increments of 20 pixels. The model parameters extracted by the GAVHT for the occluded sequences were compared against those



Fig. 9. Example of occluded image.

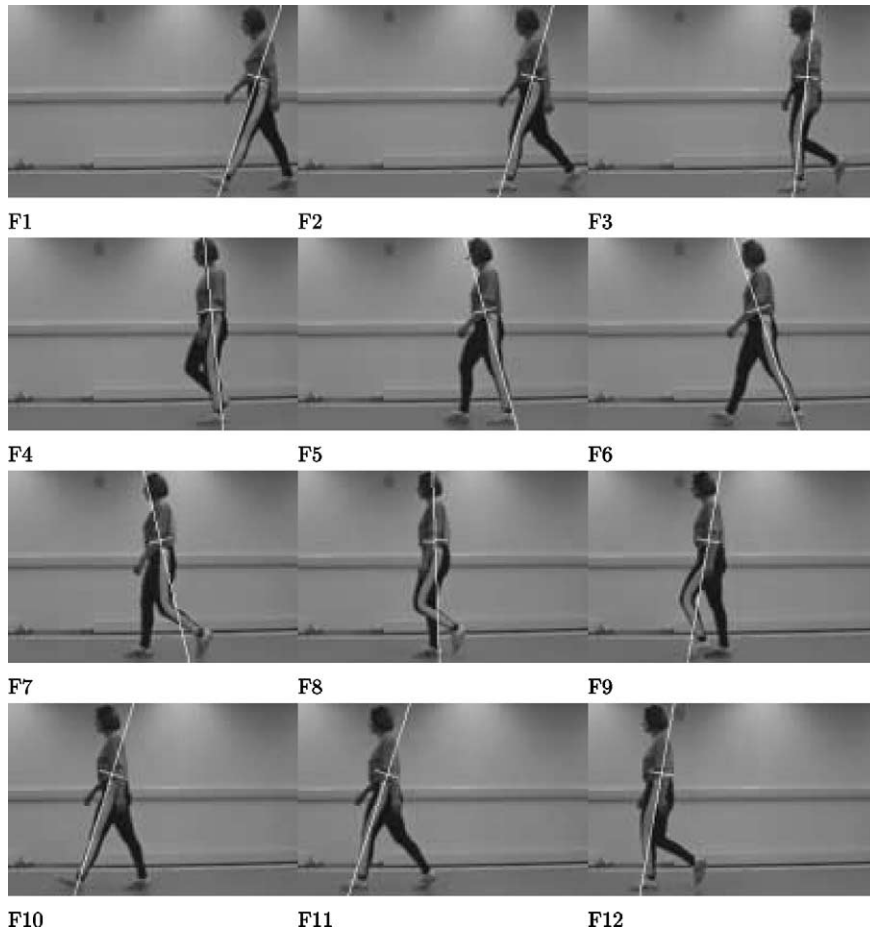


Fig. 10. Extracted reduced human gait model using GA implementation of VHT techniques for subject CM.

extracted from the unoccluded sequence. The performance of the GA for gait analysis in extracting the model from these occluded sequences was measured using the method described in Section 3.3.2. Fig. 11 shows the performance of the GAVHT for the occluded sequences. As the graph in Fig. 11 demonstrates, the performance of the GAVHT as the occluding column is widened remains comparable to its performance in the unoccluded sequence. The accuracy of the model parameter extraction deteriorates as the column width increases beyond 110 pixels, which represents almost half the image width. As such, the GA implementation of the VHT for gait analysis is shown to possess improved performance for occluded objects, which is characteristic of VHT techniques for extracting temporal features.

Fig. 12 shows the extracted model for an occluded sequence, with an occluding column width of 90 pixels, for subject CM. The subject was completely occluded

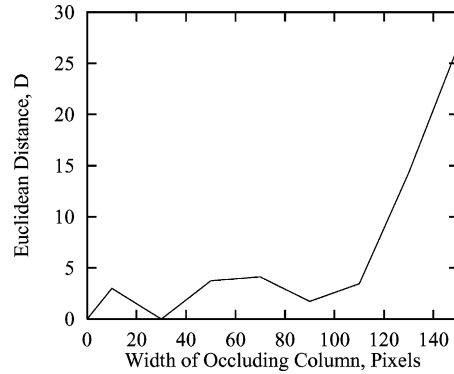


Fig. 11. Performance of GAVHT for occluded sequences with varying width of occluding column.

in four frames (F6, F7, F8, and F9) of the sequence, and partly occluded in two frames (F5 and F10): accordingly, the feature of interest was missing in 50% of the data available. Even with this amount of missing data, the subject's hip rotation was extracted with the parameter values only a single accumulator cell's distance apart from those extracted in Fig. 10, the unoccluded sequence. The initial x -coordinate of the hip was also the same as that found in Fig. 10, as well as the horizontal and vertical velocities, V_x and V_y respectively, the period T , and the parameter α . Small discrepancies between the two extracted models occurred in the values of the parameter β and the initial y -coordinate. Table 3 summarises the extracted parameter values.

3.5. Recognition results

A CCD array camera on a tripod without a shutter was used to collect data, and its output was recorded on a video recorder. The camera was situated with a plane normal to the subject's path in an environment with controlled illumination. Data collection was performed indoors, with lighting at a constant level. Subjects walked in front of a plain, static background. Each subject wore a special set of trousers that had a stripe down the middle of the outside of each leg. The new technique does not actually rely on the stripe as the front of the leg equally would be extracted. However, the presence of the stripe allows clearer assessment of extraction accuracy. In this way the camera-side leg could be distinguished visually from the other leg at all times. Fig. 13 shows an example image of a walking person used in this study.

Each subject was asked walk past the camera a total of ten times. From these ten sequences, the first and last three were discarded and only the middle four sequences were used for experimentation. In the first few sequences the subject would be *getting comfortable* with the experiment, and in the last few the subject would be anxious to finish the experiment. As such, the middle four sequences were considered to offer the most consistent walking cycles.

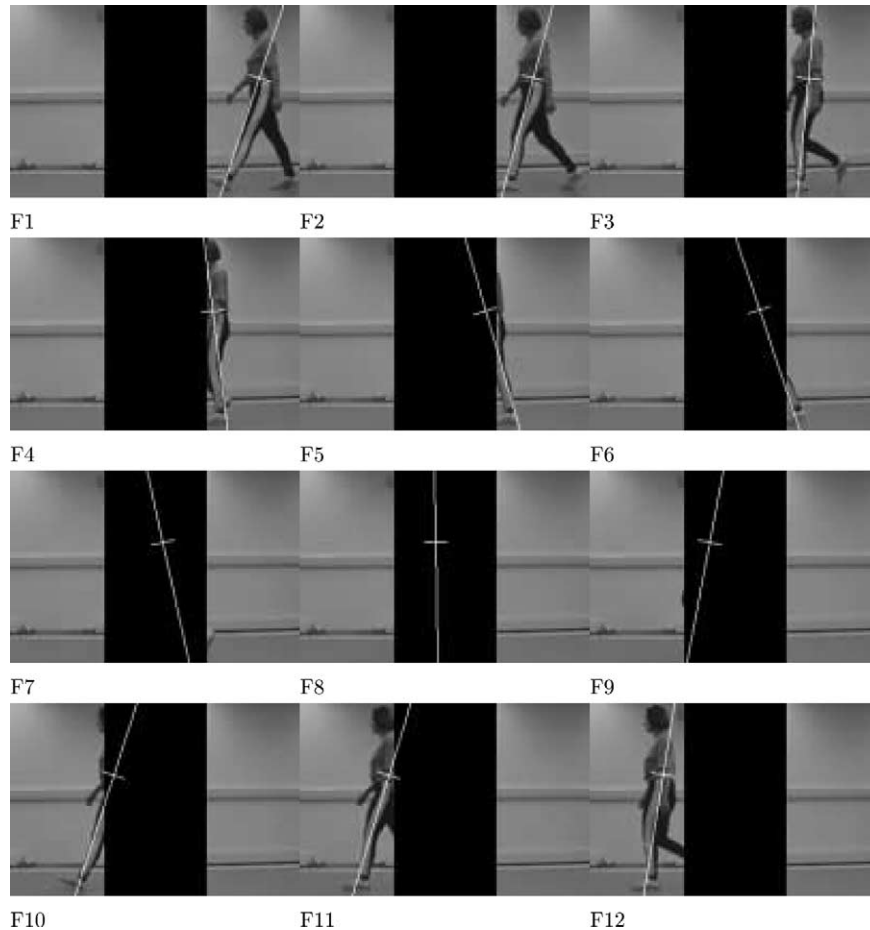


Fig. 12. Extracted thigh model for subject CM from an occluded sequence, with the occluding column width of 90 pixels.

Table 3

Extracted gait model parameters for unoccluded and occluded sequences of subject CM, from Figs. 10 and 12, respectively

Sequence	Hip model						Hip rotation model			T
	α	β	c_{x0}	c_{y0}	V_x	V_y	a_0	b_1	c_1	
Unoccluded	-5	3	219	63	118	1	0	0.16	-0.02	1.12
Occluded	-5	2	219	64	118	1	0	0.16	-0.01	1.12

Analysis used images between successive heel-strikes of the same foot, ensuring time alignment so that phase information could be compared and that subjects were not recognised by the speed of their walk. In these initial studies, to obtain



Fig. 13. Example image of a walking subject, subject 3.

meaningful information about the rotation pattern, it was essential that each subject travelled at a constant velocity. Room was given to allow the subjects to accelerate to a comfortable, constant walking speed before entering the field of view of the camera. The video sequences were digitised into the PGM image format. Resolution was reduced by halving the image size to 128×128 in order to speed processing.

Recognition analysis using the GAVHT, was initially performed on the walking sequences for ten subjects. Edge images of the sequence were produced by applying the Canny edge detector with hysteresis thresholding to the images. The GAVHT was used to extract a maximum likelihood estimate of the parameters for the thigh model (Section 3.2) in a given sequence. As an initial analysis of the performance of these novel techniques, the gait model extracted from the walking sequences described the hip rotation with a second-order FS. The full fifth-order FS was not used so as to reduce the complexity of the parameter search space. The parameters of the model were quantised with the values in Table 2, and the GA was again implemented with a chromosome population of 500.

3.5.1. Extracting the gait model

Fig. 14 presents the extracted thigh model superimposed over the original sequence of images for subject IM. The inclination of the thigh appears to have been found precisely in the majority of the frames. The model extraction showed less rotation than the leg in the first frame of the sequence and then aligned itself with the thigh in the second frame, only to over-shoot in the third. The accuracy was regained in the fourth frame. In frames eight and nine, the estimated motion of the thigh was not sufficiently fast to follow the true motion of the thigh in this part of the gait cycle. This inaccuracy was most likely due to the use of just two harmonics to model the thigh motion. As stated before, the motion of the thigh contains useful harmonic content up to the fifth harmonic. The model extraction in the last frame of the sequence shows the correct inclination, but the position of the thigh appears to trail

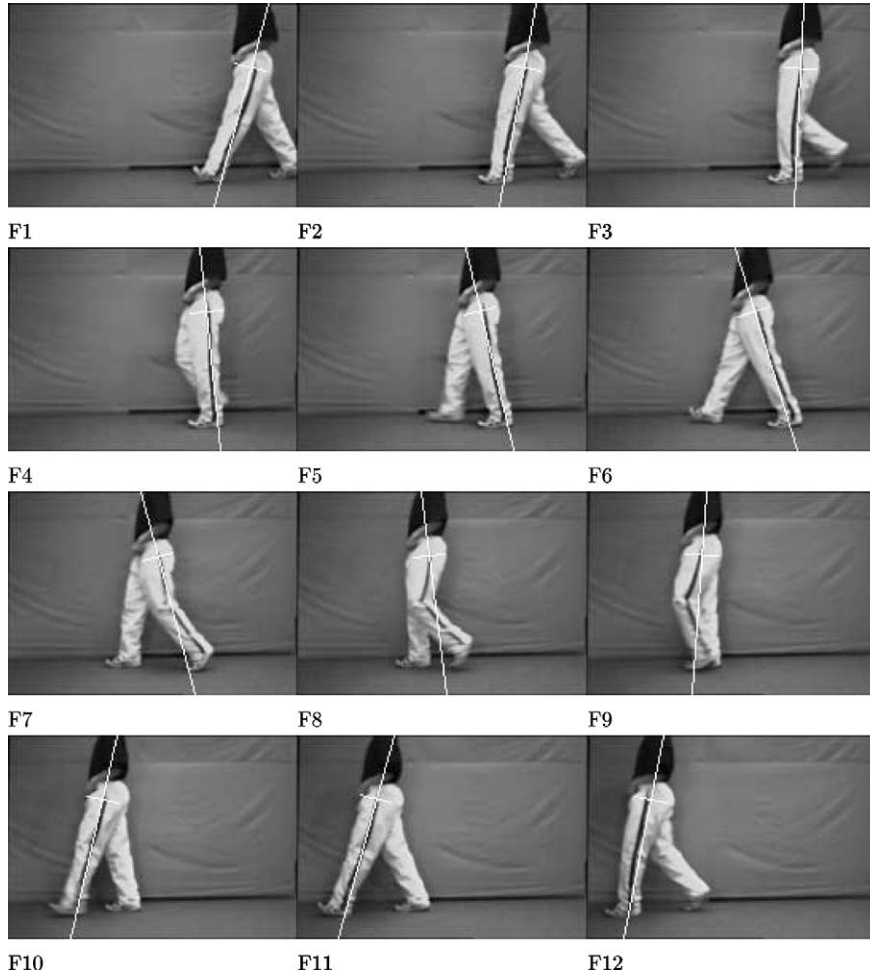


Fig. 14. Walking sequence with extracted thigh model for subject IM. Frames run from left to right, and top to bottom.

the hip. This was an inaccuracy in the estimation of the horizontal displacement of the hip (see Section 3.2.1) rather than an error due to the hip rotation model. Note that the estimated location of the hip was marked by a line normal to the long line representing the thigh.

Using the parameters of the gait model extracted from the walking sequence, the hip rotation pattern was given by Eq. (15) without the need for curve fitting. Fig. 15 shows the extracted hip rotation pattern for the gait cycle of subject IM shown in Fig. 14. The extracted rotation pattern also compared well with the generalised hip rotation pattern presented in medical research [22].

The GAVHT's ability to handle occlusion was again demonstrated in the extraction of the thigh model for the gait cycle of subject DC, shown in Fig. 16. The

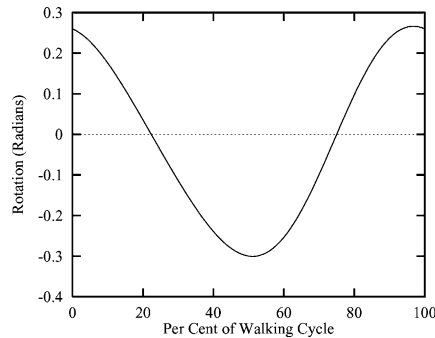


Fig. 15. Hip rotation pattern extracted with the thigh model for the sequence in Fig. 14.

extracted thigh model matches well the expected position of the moving thigh even though the subject's hand was occluding the top of the thigh in the some of the frames of the sequence. In the first frame of the sequence, the extracted inclination appeared to be in advance, but in the second frame the model extraction aligned itself precisely with the leg. A similar pattern occurred in the next two frames, where the model extraction appeared to trail the thigh in the third frame and aligned accurately in the fourth. Frames five, six, and eight showed a precise extraction of the thigh, whilst frame seven showed the model to have a greater rotation than the leg. In frames nine to eleven, the thigh inclination was extracted with precision, but again in frame twelve the extraction described a greater rotation than the thigh.

In Fig. 16, the estimated position of the hip was extracted with accuracy in the first two frames, and then again in the middle four frames. In the remaining frames the hip position appeared to trail the expected location of the hip, which was deemed to be intersection of the front of the stripe with the waist line.

Fig. 17 shows the hip rotation pattern for the sequence in Fig. 16 as described by the extracted parameters of the gait model.

3.5.2. Fourier analysis

The frequency information of the hip rotation was extracted directly from the evidence gathering process. Eqs. (18) and (19) were used to calculate the magnitude and phase of the frequency spectrum using the FS coefficients. For the gait cycle of subject DC in Fig. 16, a second-order FS was used to model the hip rotation. The magnitude and phase plots for this FS are shown in Fig. 18. Note that the Fourier transforms reflect around 0 Hz. From the magnitude plot, it can be seen that information above the second harmonic is meaningless as the magnitude drops to zero at and beyond the third harmonic. Accordingly, for k -nearest neighbour classification the metric used to calculate the Euclidean distance D between test feature vectors and training feature vectors was based on the components between 0 and +2 Hz.

Fig. 19 shows the magnitude and phase plots for the FS describing the hip rotation pattern for the gait cycle of subject IM in Fig. 14.

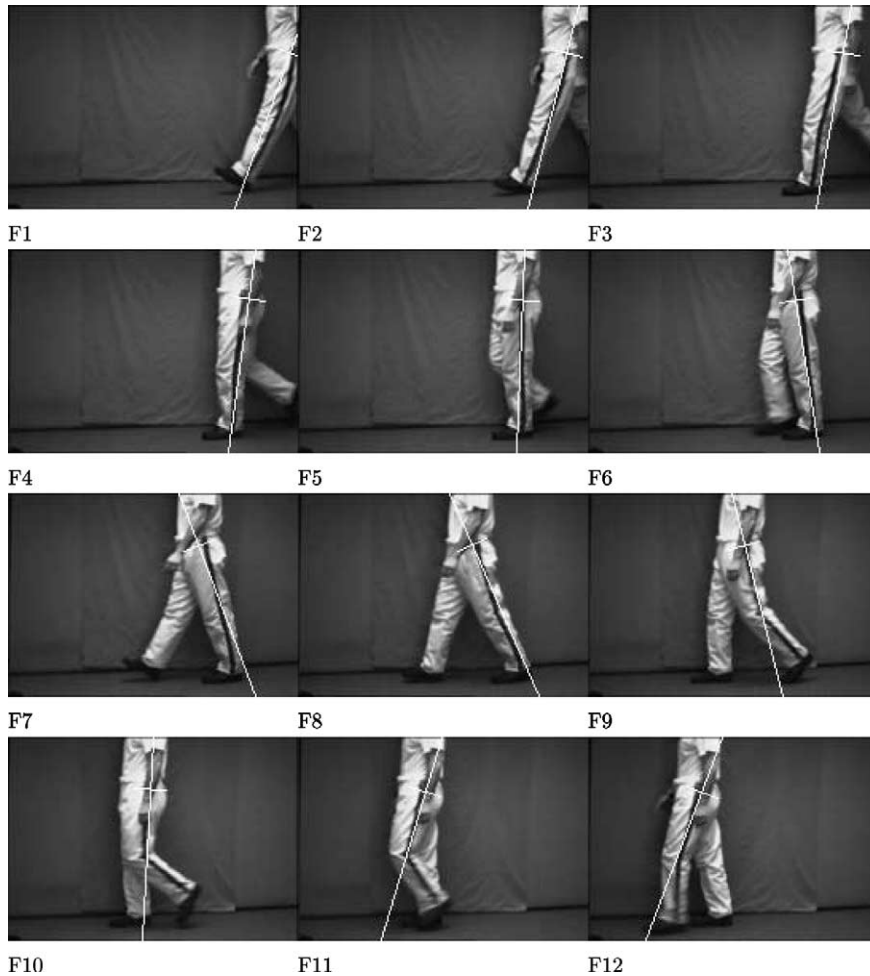


Fig. 16. Walking sequence with extracted thigh model for subject DC. Frames run from left to right, and top to bottom.

3.5.3. Classification results

For the classification analysis, two measures were compared; the Fourier magnitude and the phase-weighted Fourier magnitude. Walking sequences for ten subjects were used, each subject having four walking sequences; three training sequences and one test sequence. The measure for each test sequence was compared against those for the training sequences. The k -nearest neighbour rule was used to classify the differences in these measures for $k = 1$ and $k = 3$. Table 4 summarises the correct classification rates (CCR) for the two measures. Unlike the earlier study [5], the nearest neighbour classification led to the same classification performance as the 3-nearest neighbour rule.

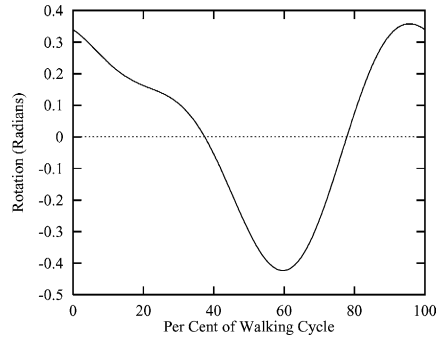


Fig. 17. Hip rotation pattern extracted with the thigh model for the sequence in Fig. 16.

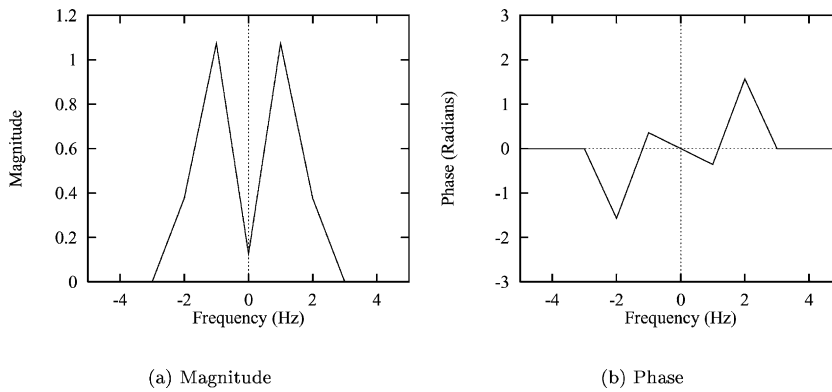


Fig. 18. Magnitude and phase plots for hip rotation as described by FS coefficients for gait cycle of subject DC in Fig. 16.

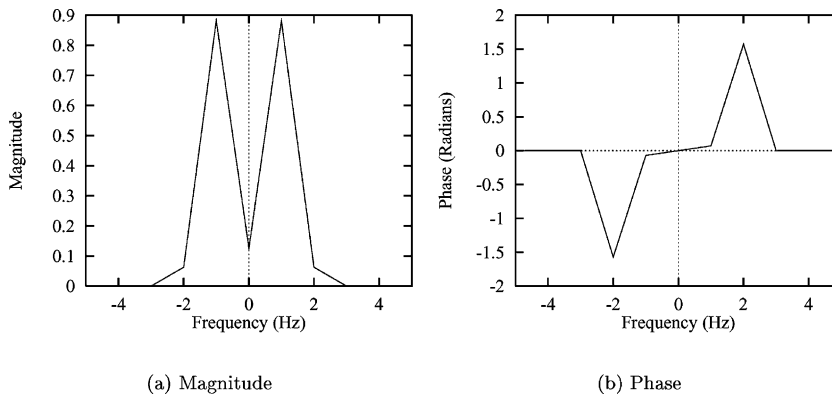


Fig. 19. Magnitude and phase plots for hip rotation as described by FS coefficients for gait cycle of subject IM in Fig. 14.

Table 4
Overall classification performance

No. of nearest neighbours	Magnitude CCR	Phase-weighted magnitude CCR
$k = 1$	80%	100%
$k = 3$	80%	100%

Classification analysis showed that the phase-weighted Fourier magnitude offered a better classification rate (100%) than just the Fourier magnitude (80%), verifying earlier work in [5]. This suggests that subjects are recognised not only by flexion, but also by the time when it occurs; both the phase and the magnitude of the oscillatory motion would intuitively appear to describe a particular pendulum better. Direct generation of the Fourier information from the evidence gathering process was possible using the FS coefficients, and as such no further transform processing was required. Using this evidence gathering technique, improved classification rates, of 100% for both $k = 1$ and $k = 3$, were achieved compared to those obtained using the computer vision techniques in [5], on the same data being 80% and 90% for $k = 1$ and $k = 3$, respectively.

3.5.4. Statistical analysis

A statistical analysis was performed on the experimental data to establish a basis for the improved classification performance of the phase-weighted Fourier magnitude metric over the Fourier magnitude metric. A statistical measure was required that described the distribution of subject, or class, clusters in the feature space. The separation \mathbf{S} between due to class means, normalised with respect to class covariances, was used. The separation $\mathbf{S}_{i,j}$ between subjects i and j is given by

$$\mathbf{S}_{i,j} = [\mathbf{m}_i - \mathbf{m}_j] \left[\frac{\Sigma_i + \Sigma_j}{2} \right]^{-1} [\mathbf{m}_i - \mathbf{m}_j]^T, \quad (21)$$

where \mathbf{m}_i is the mean and Σ_i is the covariance of class i . This measure is the first term of the Bhattacharyya-distance.

The mean signature \mathbf{m}_i for each class i is calculated by

$$\mathbf{m}_{i,k} = \frac{1}{M} \sum_{l=0}^{M-1} \mathbf{x}_{l,k}^i, \quad k = 0, 1, \dots, N, \quad (22)$$

where M is the number of experiments for class i , N is the number of harmonics (as used in Eq. (15)) and \mathbf{x}^i is an $M \times N$ data matrix of signatures for class i . The covariance matrix Σ_i is given by

$$\Sigma_i = \frac{1}{M} \sum_{l=0}^{M-1} (\mathbf{x}_l^i - \mathbf{m}_i)^T (\mathbf{x}_l^i - \mathbf{m}_i). \quad (23)$$

The data matrix \mathbf{x}^i for each subject was generated using the Fourier magnitude metric and the phase-weighted Fourier magnitude metric. Figs. 20(a) and (b) illustrate the separation \mathbf{S} as calculated from the Fourier magnitude metric and the

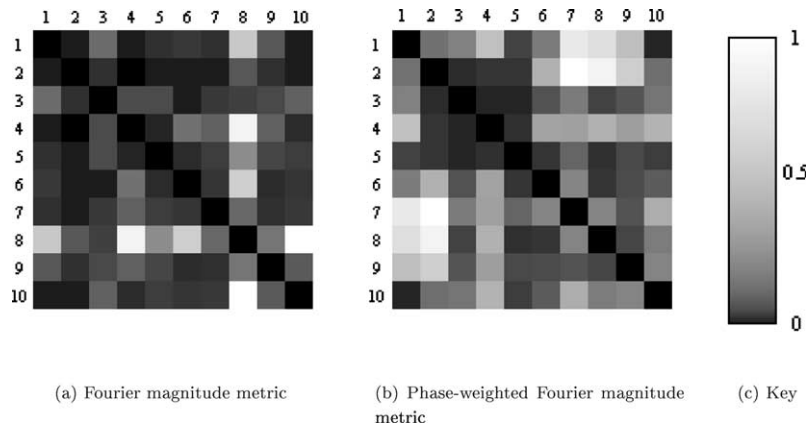


Fig. 20. Separation S of mean subject signatures, normalised with respect to subject covariances.

phase-weighted Fourier magnitude metric, respectively. The separation was normalised to be in the range $[0,1]$. As Fig. 20 shows, the phase-weighted Fourier magnitude metric demonstrates a greater separation between the mean signature of each subject in the database than the Fourier magnitude metric, as the points are much brighter in Fig. 20(a) than they are in Fig. 20(b).

Using the separation S , the average class separation was calculated for both metrics. Table 5 shows that the phase-weighted Fourier magnitude metric has a greater average distance between subjects than the Fourier magnitude metric. This suggests that a better CCR could be achieved using the phase-weighted Fourier magnitude metric over the Fourier magnitude metric, a suggestion that is supported by the classification results in Table 4. The average separation between subjects for a given metric indicates its relative performance to other metrics on a given database. It would be difficult to use this measure to estimate generalisation capability (the size of population that the metric could perform satisfactorily on) since new subject signatures could naturally lie outside the space defined by the current set of gait signatures. Clearly, since the average distance is larger for the phase-weighted data than it is for the magnitude data, the phase-weighted data would appear to be able to handle a larger population with equal performance (should the signatures fall within the space already defined) since the potential for confusion is much less than for magnitude data alone. However, the improved CCR of the phase-weighted Fourier

Table 5

Average separation between subjects for Fourier magnitude metric and phase-weighted Fourier magnitude metric

Metric	Average separation
Fourier magnitude	0.105
Phase-weighted Fourier magnitude	0.219

magnitude signature over the Fourier magnitude signature has been predicted through a statistical analysis of the experimental data.

4. Further work

The aims of the research initially outlined were fulfilled. In this process, several areas were identified for further study. Preliminary experimentation was performed into investigating the extension of the human gait model itself to encompass other limbs. Including the lower legs, or even the torso or arms, into the model has several implications. Firstly, there is the problem of how to handle the ever increasing dimensionality of the model. GAs offer one solution to the dimensionality problem. Also, the relative independence of each limb can be explored to maximise parallelisation of the model extraction. Secondly, given the extraction of multiple limbs there exists the problem of how this information can be combined to achieve the best discriminatory power in the classification process.

Also, the pre-processing of the images can be investigated for more reliable and accurate extraction of model parameters. Some pre-processing techniques and their potential implications are discussed.

Application of the technique described in Section 3 to a wider database is required to verify and improve the recognition performance. Also, the effects of footwear, running, weight-bearing, and loose apparel (and without stripes) on the classification rate requires study. In this research the data for each subject was obtained in the same session. This is not representative of a practical application and therefore subject data should be obtained on different occasions. This will allow the variability of signatures for a given individual to be assessed to establish possible class bounds. It would also appear worthwhile to assess the potential effect of background, though the simulation tests in noise (Section 3.3.2) have indicated good ability to handle background. Clearly, imagery with multiple moving people will require a new strategy to interpret the accumulator space.

The extraction of the human gait model from different camera perspectives of the walking subject must be addressed and evidence gathering techniques for these alternative view points developed. Work by [3] extracted 3D trajectories of moving objects in a scene. This work assumed all motion to be contained in a given (ground) plane, allowing the 3D trajectories to be tracked with a single camera. Application of this tracking technique to gait model extraction from varying camera perspectives requires investigation.

4.1. *Extending the human gait model*

The most basic extension to the human gait model is the addition of the lower leg. This entails the addition of a second moving line, with a periodically changing inclination, to the structural model. Similarly, the knee rotation must be included in the motion model. The position of the fulcrum (knee) for this second moving line (shin) is controlled by the motion and structure of the original line (thigh). To calculate the

position of the knee, the thigh requires a length tl . Similarly, the shin can be given a length kl , allowing even further extension to the inclusion of the foot. The extended human gait model is shown in Fig. 21. The motion model for the human gait model remains as described in Section 3.2.1, but with the addition of the knee rotation $\theta(t)$. As with the hip rotation $\phi(t)$, the knee rotation $\theta(t)$ can be modelled by a FS:

$$\theta(t) = a_0 + 2 \sum_{k=1}^L [b_k \cos k\omega_0 t - c_k \sin k\omega_0 t], \quad (24)$$

where L is the number of harmonics used to describe the knee rotation $\theta(t)$. The structural model for the human gait model requires a slight modification to accommodate the addition of the lower leg. An arbitrary point \mathbf{r} on the model is now described by

$$\mathbf{r} = \begin{cases} \mathbf{c}(t) + \lambda \mathbf{u}(t) & 0 \leq \lambda < tl, \\ \mathbf{k}(t) + \lambda \mathbf{p}(t) & tl \leq \lambda \leq tl + kl, \end{cases} \quad (25)$$

where tl and kl are the lengths of the thigh and shin, respectively, $\mathbf{c}(t)$ is the position vector of the pelvis, $\mathbf{u}(t)$ is the unit vector of the direction of the thigh, $\mathbf{k}(t)$ is the position vector of the knee, $\mathbf{p}(t)$ is the unit vector of the direction of the shin, and λ can take any real value between 0 and $tl + kl$. The values of $\mathbf{k}(t)$ and $\mathbf{p}(t)$ are given by

$$\mathbf{k}(t) = \mathbf{c}(t) + tl\mathbf{u}(t), \quad (26)$$

$$\mathbf{p}(t) = (-\sin[\phi(t) - \theta(t)], \cos[\phi(t) - \theta(t)]), \quad (27)$$

where $\phi(t)$ is the hip rotation and $\theta(t)$ is the knee rotation.

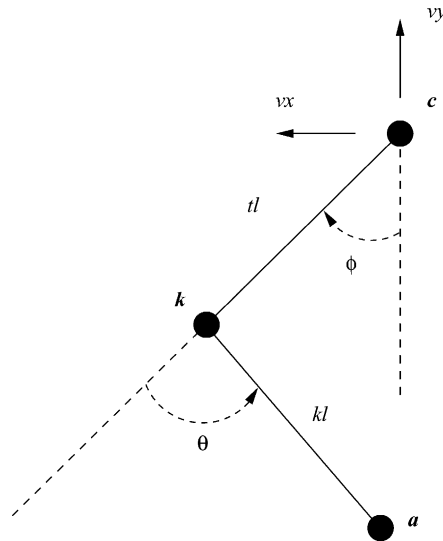


Fig. 21. Extended human gait model, including thigh and shin.

The GA for gait analysis was modified to extract the extended gait model described above. Fig. 22 shows this extracted model for subject RW. Both the rotation of the thigh and knee were modelled by second-order FS in the model extraction in Fig. 22. The inclination of the thigh appears to have been precisely extracted throughout the majority of the sequence. Frames four and ten show the thigh extraction to have a greater rotation than perceived for subject's thigh. The rotation of the knee appears to have been extracted with less accuracy. This is especially evident in frames four, five, and seven to nine. This could be due to its dependence on the precision of the thigh extraction. Also, from the medical studies described in chapter 2, the knee rotation characteristically has a rapid and large change in inclination over the latter part of the gait cycle, as it goes into flexion then extension. This behaviour will not have been faithfully reproduced by the second-order FS used to represent the knee rotation in this extraction. The greatest inaccuracies in the model extraction seem to have been in the motion of the pelvis. Slight inaccuracy in the extracted average velocity V_x resulted in the model extraction falling behind in the final frames,

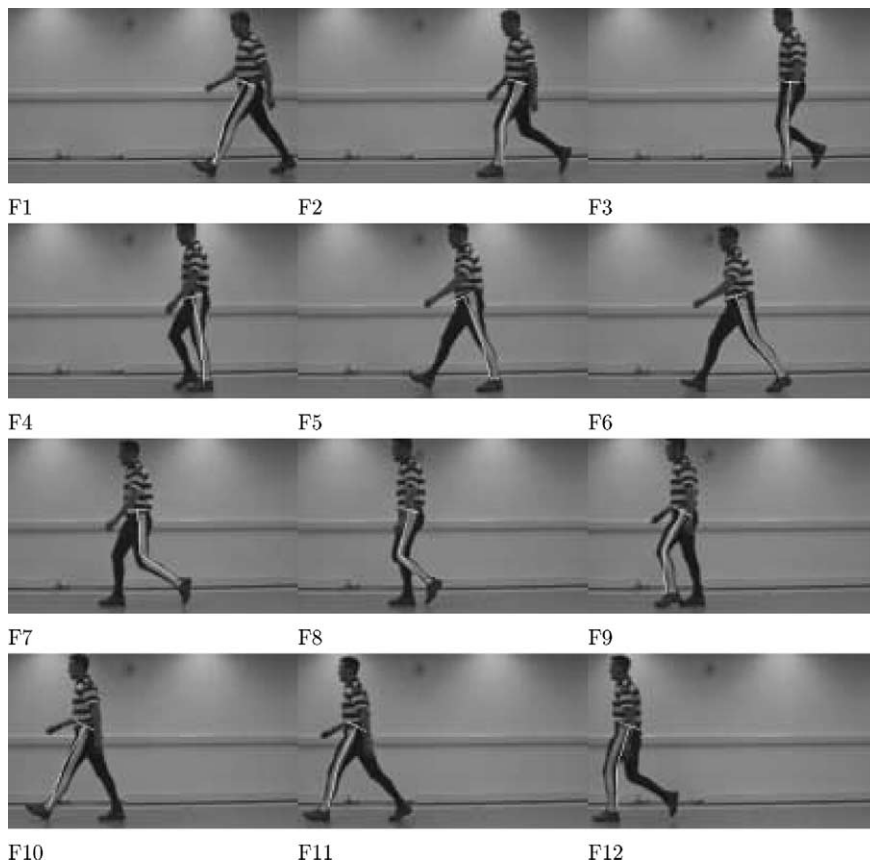


Fig. 22. Extracted human gait model for subject RW, model extended to include lower leg.

eleven and twelve. Improved extraction of the two rotation patterns with a higher order FS would improve on the determination of the pelvic motion parameters. However, this example was sufficient to demonstrate the potential for expansion of the model.

By extending the human gait model to include the lower leg, the rotation pattern of the knee can be used as another possible feature in the classification process. [12] have shown that using orthogonal features can improve classification rates. In this work, it was shown that using several orthogonal feature measures to produce an extended feature vector gave improved recognition of faces. Taken individually, the measures did not correctly classify all test images. When each measure was given an equal weight in an extended feature vector, all test images were correctly classified. This extended feature vector could be applicable to gait recognition.

Indeed, [16] demonstrated that an increased number of features in the feature vector improved correct classification rates for gait recognition. Their approach extracted a feature vector based on a number of scalars characterising the shape of motion in a sequence of images. Using only a single phase feature, correct classification rate was 90.5%, increasing to a maximum of 95.2% with five features. Although experimentation was only performed on a limited database, the potential of increasing the number of features in the feature vector to improve classification rate in gait recognition is clearly demonstrated.

4.2. Pre-processing of image data

Relatively little pre-processing was performed on the images in the walking sequence; only the Canny edge operator was applied to produce edge images of the input data. Essentially, the human gait model describes a moving line whose inclination is constrained by a periodic signal and velocity governed by some initial conditions and characteristics. Further work could explore the effect on the model parameter extraction when the evidence gathering process is performed on line images rather than edge images. These line images can be produced by SHT for lines. Due to the nature of the voting method in the SHT, in complex noisy scenes containing many various sized lines the shortest lines are unlikely to be detected. As such, the use of the SHT for producing line images is limited.

A pyramidal approach to the Hough transform for lines, the Hierarchical Line Finder (HLF), was proposed by [28], wherein a Hough transform was performed at the base of the pyramid and the detected line segments were then reaccumulated at each level of the pyramid. The HLF suffered from a quantisation effect in the re-accumulation. Further to this approach, the Hierarchical Hough transform [33] (HHT) combines a hierarchical framework with a local-to-global line detection scheme to ensure that small, but meaningful, lines are not lost. Again the Hough transform is performed at the base of the pyramid and in the upper levels a merging process, based on distance discrimination between lines, is used to merge similar lines together. This avoids the quantisation effect of a re-accumulation at each level. Fig. 23 shows HHT applied to an image, and the extracted lines superimposed on the original image.

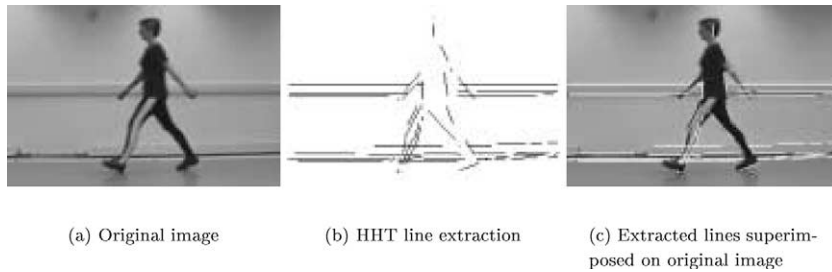


Fig. 23. Line extraction using the HHT for lines.

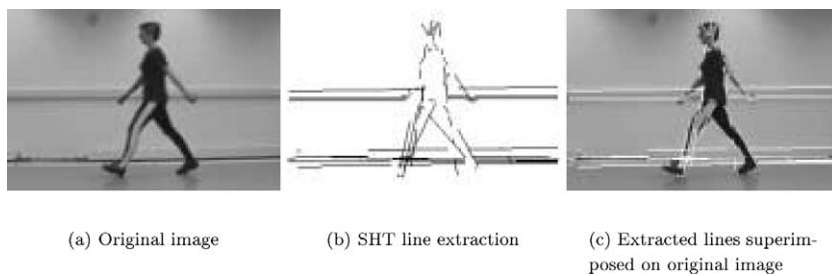


Fig. 24. Line extraction using the SHT for lines.

Comparing the line extraction of the HHT with the results of the SHT applied to the same image (Fig. 24), shows that the HHT line extraction produces a less noisy image. Examination of the trailing right foot of the subject demonstrates that the HHT extracted the contour of the sole of the foot, whilst this feature was missed by the SHT. The HHT could be used as a second stage of pre-processing to the GA for gait analysis. Accordingly, this promotes the investigation into a technique that incorporates HHT methodology for extracting temporal features—a Hierarchical Velocity Hough transform (HVHT), perhaps.

5. Conclusions

Previous work in [5] showed that a feature-based method could be used for gait recognition. It also showed that a gait signature based on the phase-weighted Fourier magnitude offered an improved correct classification rate over one based on just the Fourier magnitude or phase. Although these techniques achieved positive results, there were several strategic problems. The work described in this paper sets about solving those problems.

Greater immunity to moderate noise and feature occlusion when extracting temporal features in a sequence of images was achieved by using VHT evidence gathering techniques. All the frames in the image sequence were used in the

evidence gathering process, allowing the concurrent extraction of both structural and temporal parameters of the feature. An improved human gait model was described, having both a structural and temporal description of the upper leg. The hip rotation was modelled by a FS, paralleling earlier medical studies that described gait as a periodic signal. This FS description of the thigh motion allowed the generation of the gait signature directly from the evidence gathering process, via the FS coefficients.

A VHT for extracting the human gait model was implemented, and its performance was compared with the traditional Hough transform tracking (HTT) method for extracting temporal features from image sequences. Testing on a synthetic set of images, the VHT implementation was found to have better noise immunity than the HTT method.

Applying VHT methods to extract this new human gait model from a sequence of images required a high dimensional parameter space to be created and searched, which was unfeasible with current technology. A GA was implemented to perform this search, effectively using a multi-frame template matched across the entire sequence of images to extract model instances. The GA reduced the computational time to an acceptable level. The GA was applied to a sequence of images in which the subject had been (synthetically) occluded in some frames of the sequence. The extracted model parameters were shown to be comparable to those extracted from the unoccluded sequence. This demonstrated that the GA possessed the VHT's characteristic immunity to occlusion.

The experimental results for a database of ten subjects showed that the GA for gait analysis could extract parameters for the human gait model with a high fidelity to the original image data. Fourier analysis was not required to be performed on the hip rotation patterns of the subjects to obtain their Fourier magnitude and phase components. This information was generated directly from the FS describing the hip rotation. Gait signatures were formed using the magnitude information alone and the phase-weighted magnitude information. Classification was done using the k -nearest neighbour rule, and the results verified that an improved CCR (100%) was achieved when using the phase-weighted magnitude information rather than just the magnitude information (80%). A statistical analysis has shown how the phase-weighted Fourier data improves performance over the magnitude data, by resulting in much greater cluster separation as such confirming improved classification performance.

Using VHT techniques implemented as a GA for gait analysis, a gait signature was extracted from a walking sequence of images that was formed from parameters found directly in the evidence gathering process. The gait signature was generated using five parameters from the human gait model—the second-order FS representing the hip rotation. This biometric measure was found to give good discriminatory performance on a small database of subjects. The gait signature as generated by this novel temporal feature extraction technique gave a perfect CCR (100%) and it is not unlikely that future experimentation on a larger database will predicate further improvement on the current model to match this performance, perhaps along the lines suggested as further work.

References

- [1] J.K. Aggarwal, Q. Wai, Human motion analysis: a review, *Comput. Vision Image Understanding* 73 (3) (1999) 428–440.
- [2] C. Angeloni, P.O. Riley, D.E. Krebs, Frequency content of whole body gait kinematic data, *IEEE Trans. Rehabilitation Eng.* 2 (1) (1994) 40–46.
- [3] K.J. Bradshaw, I.D. Reid, D.M. Murray, The active recovery of 3D motion trajectories and their use in prediction, *IEEE Trans. Pattern Anal. Mach. Intell.* 19 (3) (1997) 219–233.
- [4] J.E. Bresenham, Algorithm for computer control of a digital plotter, *IBM Systems J.* 4 (1965) 25–30.
- [5] D. Cunado, M.S. Nixon, J.N. Carter, Using gait as a biometric, via phase-weighted magnitude spectra, in: J. Bigun, G. Chollet, G. Borgefors (Eds.), *Proceedings of 1st International Conference on Audio- and Video-Based Biometric Person Authentication AVBPA'97*, in: *Lecture Notes in Comput. Sci.*, vol. 1206, 1997, pp. 95–102.
- [6] K. Etemad, R. Challappa, Discriminant analysis for recognition of human face images, *J. Opt. Sci. Amer.* 14 (8) (1997) 1724–1733.
- [7] G. Gerig, F. Klein, Fast contour identification through efficient Hough transform and simplified interpretation strategy, in: *Proceedings of the 8th International Conference on Pattern Recognition*, 1986, pp. 498–500.
- [8] D. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley, Reading, MA, 1988.
- [9] P.S. Huang, C.J. Harris, M.S. Nixon, Human gait recognition in canonical space using temporal templates, *IEE Proc. Vision Image Signal Process.* 146 (2) (1999) 93–100.
- [10] P.S. Huang, C.J. Harris, M.S. Nixon, Recognizing humans by gait via parametric canonical space, *AI in Eng.* 13 (1999) 359–366.
- [11] A. Jain, R. Bolle, S. Pankanti (Eds.), *Biometrics: Personal Identification in Networked Society*, Kluwer Academic Publishers, Dordrecht, 1999.
- [12] X. Jia, M.S. Nixon, Extending the feature vector for automatic face recognition, *IEEE Trans. Pattern Anal. Mach. Intell.* 17 (2) (1993) 1167–1176.
- [13] S. Johnson, *Dictionary of the English Language*, Times Books, New York, 1755.
- [14] E.L. Kuan, Investigating gait as a biometric, Technical report, Department of Electronics and Computer Science, University of Southampton, 1995.
- [15] J.J. Little, J.E. Boyd, Describing motion for recognition, in: *Proceedings of the International Symposium on Computer Vision*, Coral Gables, FL, USA, 1995, pp. 235–240.
- [16] J.J. Little, J.E. Boyd, Recognising people by their gait: the shape of motion, *Veridic—Electronic J. Comput. Vision* 1 (2) (1998) 1–32.
- [17] E. Lutten, P. Martinez, A genetic algorithm with sharing for the detection of 2D geometric primitives in images, *Lecture Notes in Comput. Sci.* 1063 (1996) 287–303.
- [18] D. Meyer, Human gait classification based on hidden Markov models, in: H.-P. Seidel, B. Girod, H. Niemann (Eds.), *3D Image Analysis and Synthesis'97*, 1997, pp. 139–146.
- [19] D. Meyer, J. Pösl, H. Niemann, Gait classification with HMMs for trajectories of body parts extracted by mixture densities, in: *Proceedings of the British Machine Vision Conference*, Southampton, UK, 1998, pp. 459–468.
- [20] T.B. Moeslund, E. Granum, A Survey of Computer Vision-Based Human Motion Capture, *Comput. Vision Image Understanding* 8 (2001) 231–268.
- [21] H. Murase, R. Sakai, Moving object recognition in eigenspace representation: gait analysis and lip reading, *Pattern Recognition Lett.* 17 (1996) 155–162.
- [22] M.P. Murray, Gait as a total pattern of movement, *Amer. J. Phys. Med.* 46 (1) (1967) 290–332.
- [23] M.P. Murray, A.B. Drought, R.C. Kory, Walking patterns of normal men, *J. Bone Joint Surg.* 46-A (2) (1964) 335–360.
- [24] J.M. Nash, J.N. Carter, M.S. Nixon, Dynamic feature extraction via the Velocity Hough transform, *Pattern Recognition Lett.* 18 (1997) 1035–1047.
- [25] J.M. Nash, J.N. Carter, M.S. Nixon, Extraction of moving articulated-objects by evidence gathering, in: *Proceedings of the British Machine Vision Conference*, Southampton, UK, 1998, pp. 609–618.

- [26] M.S. Nixon, J.N. Carter, D. Cunado, P.S. Huang, S.V. Stevenage, Automatic gait recognition, see Ref. 11, pp. 231–250.
- [27] S.A. Niyogi, E.H. Adelson, Analyzing and recognizing walking figures in xyt, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1994, pp. 469–474.
- [28] J. Princen, J. Illingworth, J. Kittler, A hierarchical approach to line extraction based on the Hough transform, *Comput. Vision, Graphics Image Process.* 52 (1990) 57–77.
- [29] Spectrum, It had to be you, *IEEE Spectrum*, 1994.
- [30] M. Unuma, K. Anjya, R. Takeuchi, Fourier principles for emotion-based human figure animation. in: Proceedings ACM SIGGRAPH '95 (22nd ACM Conference on Computer Graphics), Los Angeles, USA, 1995, pp. 91–96.
- [31] T.M. van Veen, F.C.A. Groen, Discretization errors in the Hough transform, *Pattern Recognition* 14 (1981) 137–145.
- [32] Various, The Biometrics Report, SJB Services, 1995.
- [33] S.B. Yacoub, J.M. Jolin, Hierarchical line extraction, *IEE Proc. Vision, Image Signal Process.* 142 (1) (1995) 7–14.