Biometric Gait Recognition

Jeffrey E. Boyd[1] and James J. Little[2]

[1] Department of Computer Science
University of Calgary
boyd@cpsc.ucalgary.ca
[2] Department of Computer Science
University of British Columbia
little@cs.ubc.ca

Abstract. Psychological studies indicate that people have a small but statistically significant ability to recognize the gaits of individuals that they know. Recently, there has been much interest in machine vision systems that can duplicate and improve upon this human ability for application to biometric identification. While gait has several attractive properties as a biometric (it is unobtrusive and can be done with simple instrumentation), there are several confounding factors such as variations due to footwear, terrain, fatigue, injury, and passage of time. This paper gives an overview of the factors that affect both human and machine recognition of gaits, data used in gait and motion analysis, evaluation methods, existing gait and quasi gait recognition systems, and uses of gait analysis beyond biometric identification. We compare the reported recognition rates as a function of sample size for several published gait recognition systems.

1 Introduction

People often feel that they can identify a familiar person from afar simply by recognizing the way the person walks. This common experience, combined with recent interest biometrics, has lead to the development of gait recognition as a from of biometric identification.

As a biometric, gait has several attractive properties. Acquisition of images portraying an individual's gait can be done easily in public areas, with simple instrumentation, and does not require the cooperation or even awareness of the individual under observation. In fact, it seems that it is the possibility that a subject may not be aware of the surveillance and identification that raises public concerns about gait biometrics [1].

There are also several confounding properties of gait as a biometric. Unlike finger prints, we do not know the extent to which an individual's gait is unique. Furthermore, there are several factors, other than the individual, that cause variations in gait, including footwear, terrain, fatigue, and injury.

This paper gives an overview of the factors that affect both human and machine recognition of gaits, data used in gait and motion analysis, evaluation methods, existing gait and quasi gait recognition systems, and uses of gait analysis beyond biometric identification.

M. Tistarelli, J. Bigun, and E. Grosso (Eds.): Biometrics School 2003, LNCS 3161, pp. 19–42, 2005.

1.1 Gait and Gait Recognition

We define gait to be *the coordinated, cyclic combination of movements that result in human locomotion*. The movements are coordinated in the sense that they must occur with a specific temporal pattern for the gait to occur. The movements in a gait repeat as a walker cycles between steps with alternating feet. It is both the *coordinated* and *cyclic* nature of the motion that makes gait a unique phenomenon.

Examples of motion that are gaits include walking, running, jogging, and climbing stairs. Sitting down, picking up an object, and throwing and object are all coordinated motions, but they are not cyclic. Jumping jacks are coordinated and cyclic, but do not result in locomotion.

Therefore, we define gait recognition to be the recognition of some salient property, e.g., identity, style of walk, or pathology, based on the coordinated, cyclic motions that result in human locomotion. In the case of biometric gait recognition, the salient property is identity. We make the distinction between gait recognition and what we call quasi gait recognition in which a salient property is recognized based on features acquired while a subject is walking, but the features are not inherently part of the gait. For example, skeletal dimensions may be measured during gait and used to recognize an individual. However, skeletal dimensions may be measured other ways, and are therefore not a property of the gait.

1.2 Human Perception of Gait

The ability of humans to recognize gaits has long been of interest to psychologists. Johansson [2, 3] showed that humans can quickly (in less than one second) identify that a pattern of moving lights, called a moving light display (MLD), corresponds to a walking human. However, when presented with a static image from the MLD, humans are unable to recognize any structure at all. For example, without knowing that the dots in a single frame of the sequence shown in Fig. 1 are on the joints of a walking figure, it is difficult to recognize them as such. What we cannot show in a print medium is, that within a fraction of a second after the dots move, one can recognize them as being from a human gait.

Johansson's contributions are important because they provide an experimental method that allows one to view motion extracted from other contextual information. With the context removed, the importance of motion becomes obvious. Johansson also suggests a set of gestalt rules that humans use to connect the moving dots and infer structure.

Bertenthal and Pinto [4] identify the following three important properties in the human perception of gaits.

- *Frequency entrainment.* The various components of the gait must share a common frequency.

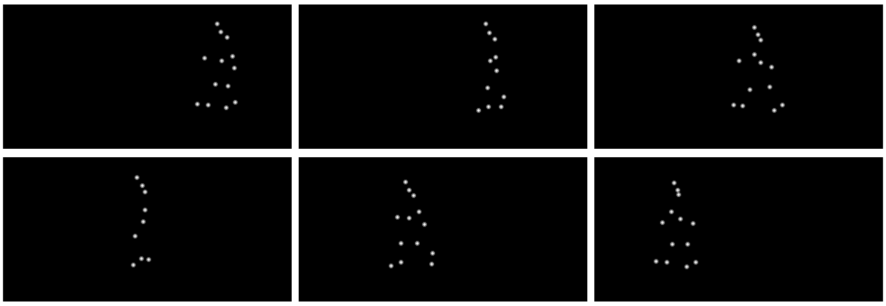


Fig. 1. Frames from a moving light display of a person walking. People can quickly identify that the motion is a gait from the moving sequence, but have difficulty with static frames.

- *Phase locking.* The phase relationships among the components of the gait remain approximately constant. The lock varies for different types of locomotion such as walking versus running.
- *Physical plausibility.* The motion must be physically plausible human motion.

As shown in Fig. 2, there are motions at different frequencies within a gait. However, the gait has a fundamental frequency that corresponds to the complete cycle. Other frequencies are multiples of the fundamental. This is frequency entrainment. It is not possible to walk with component motions at arbitrary frequencies.

When the motions are at entrained frequencies, the phase of the motions must be locked, i.e., the timing patterns of the motions are fixed. In a typical gait, the left arms swings in phase with the right leg and opposite in phase with the left leg, a pattern that is fixed throughout the gait. This is phase locking.

To understand physical plausibility, consider the motion of the star of an action movie such as Jackie Chan or Jet Li. On occasion, the actors will use wires to allow them to perform feats that would not be physically possible otherwise. However, even though the wires are not visible in the movie, viewers know that the wires are there because the motion is not physically plausible without them. Currently, physical plausibility is not employed in machine analysis of gait, other than by the use of exemplars which are real, and therefore physically plausible.

It appears that there is a special connection between human gaits and human perception. Cohen et al. [5] observed that while humans can easily recognize human motion, they have more difficulty recognizing animal motion. Cohen et al. explain this observation by suggesting that humans rely on the same mechanisms that they use to generate their own gait to perceive the gaits of others. If correct, this may indicate how to improve machine perception of gait.

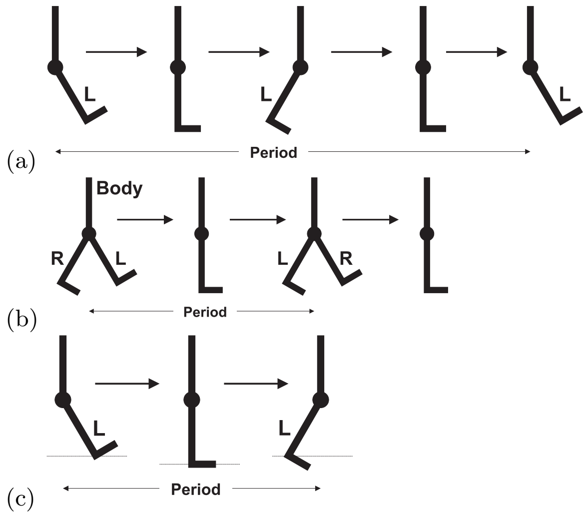


Fig. 2. Stylized body and legs showing sources of different frequencies in a synthesized gait: (a) the oscillation of a swinging limb repeats periodically, e.g., left foot fall to left foot fall, (b) the silhouette of a body repeats at twice that frequency, i.e., step to step, and (c) the pendulum motion of limbs has vertical motion at twice the frequency of the limbs horizontal motion.

1.3 Important Factors in Evaluation of Gait Analysis Systems

There are many and varied approaches to gait analysis. In order to interpret them in some common context, we suggest the following approach to understanding gait analysis systems.

1. Identify the oscillating signals that the system derives from the cyclic motion.
2. Determine how the oscillating signals establish frequency entrainment, phase locking, and physical plausibility.
3. Determine how the oscillating signals translate into features that can be used for recognition.

2 Potential for Gait as a Biometric

The use of gait as a biometric for human identification is still young when compared to methods that use voice, finger prints, or faces. Thus, it is not yet clear how useful gait is for biometrics. In this section we consider evidence from several sources, including known properties of the human body and human performance to gain insight.

2.1 Optimistic Viewpoint

Bhanu and Han [6] present an optimistic view of the potential for biometric gait recognition. Their analysis is built upon a gait recognition system that measures a subject's skeletal dimensions as he walks. Therefore, it is possible to estimate an upper bound on the performance of the system from known distributions of skeletal dimensions in a human population. They compute their estimate using a Monte Carlo simulation seeded with the population statistics and a set of assumptions about the accuracy of the skeletal dimension measurements. Plots showing the bounds they compute are in Fig. 8.

Since theirs is a quasi gait recognition system, it is reasonable to ask whether or not the bound might reasonably apply to gait recognition too. Do skeletal dimensions sufficiently constrain a gait for the purposes of recognition? The answer is unknown, but work in mechanical engineering can shed some light. McGeer [7, 8], and later Coleman and Ruina [9], Garcia et al. [10], and Collins et al. [11] have demonstrated passive mechanical walkers. These are mechanical machines that oscillate without external force to produce a gait as the machine *falls* down an incline. This implies that gait is a natural bi-product of the structure of the human body, and the mass and skeletal dimensions of the body are what determine the oscillations that produce the gait. Thus, to a large extent, Bhanu and Han are right to equate skeletal dimensions with gait. However, mass and other factors contribute to a human gait.

It is worth noting here that many gait analysis systems could benefit from the definition of a *standard* or *normal* gait. Passive mechanical walkers have the potential to define such a gait because they show the innate gait of the kinematic structure in the absence of muscular forces.

Bhanu and Han's results show one important feature of gait and other biometric systems. Regardless of the quality of biometric, the system performance in terms of recognition rate drops with increased population size. The best that one can hope for is that the rate at which performance drops is tolerable.

2.2 Human Performance

People often have the impression that they can recognize friends by their gaits. Although this ability has been confirmed by experiments using MLDs, human ability to recognize people from motion is limited.

For example, Barclay et al. [12], and Kozlowski and Cutting [13] showed that humans can recognize the gender of a walker from an MLD. However, for short exposures to the MLD (two seconds or less), humans were no better than random. It required longer exposures, on the order of four seconds, for humans to perform better than random. Even at that, the recognition rate was 66% when random was 50%.

Cutting and Kozlowski [14] also showed that people can recognize their friends from MLDs. Again, this result needs clarification. The experiment involved six students who knew each other well. Experimenters recorded MLDs for the six students. Then, at a later date, the original six, plus a seventh who was also a friend, tried to recognize their friends from the MLDs. The correct recognition rate was 38% which is significantly better than random (17%). Thus, the conclusion that people can recognize friends from motion is correct, but not well enough to be a reliable form of identification. It seems that people rely on other contextual clues more than they realize.

2.3 Confounding Factors

If passive mechanical walkers are a good indication, then the primary determinant of a gait is a person's skeletal dimensions and mass. Other factors play a role too, including:

- **terrain** (Laszlo et al. [15] illustrate variations in human gait due to terrain in computer graphic),
- **injury** (Murray et al. [16] and Murray [17] describe the effects of injury on gait),
- **footwear**, (von Tscharner [18] shows that muscle activation in walkers changes when people walk bare foot as opposed to wearing shoes),
- **muscle development**,
- **fatigue**,
- **training** (athletic training or military marching drills),
- **cultural artifacts** (e.g., mince, swagger, and strut), and
- **personal idiosyncrasies**.

Each of these factors may confound biometric gait recognition.

3 Data in Gait Recognition

In this section we give an overview of the types of data used in gait and motion analysis systems.

3.1 Background Subtraction

Background subtraction is a method for identifying moving objects against a static background. Although there are many variations on the theme, the basic idea is to

1. estimate the pixel properties of the static background,
2. subtract actual pixel values from the background estimates, and
3. assume that if the difference exceeds a given threshold that the pixel must be part of a moving object.

Normally one follows the last step by forming connected components, or blobs, of moving pixels that correspond to the moving objects. Factors that confound background subtraction include background motion, moving objects that are similar in appearance to the background, background variations over long periods of time, and objects in close proximity merging together. In general, the variations on the theme of background subtraction involve selecting pixel properties to compare, background models, and innovations to address any number of confounding factors. Examples include Hunter et al. [19], Horprasert et al. [20], Stauffer and Grimson [21], and Javed et al. [22].

Fig. 3 shows an example of background subtraction taken from the MoBo database [23].

(a) (b)

Fig. 3. Example of background subtraction from MoBo database [23]: (a) original image (deliberately blurred to conceal the subject's identity), and (b) segmented image.

3.2 Silhouettes

Background subtraction provides a set of pixels within the region of a moving object. Alternatively, one may only be interested in the outline of that region. We refer to this outline as a silhouette. An examples of gait analysis that uses silhouettes is in Baumberg and Hogg [24].

3.3 Optical Flow

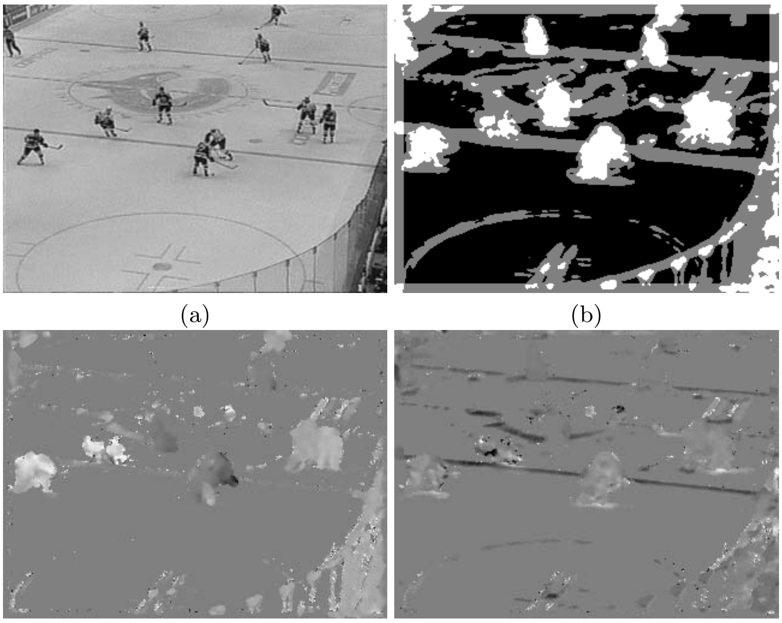
A motion field, is a projection of motion in a scene onto the image plane. Optical flow refers to the movement or flow of pixel brightness in an image sequence, and is a quantity that we can estimate from images sequences. Although the motion field and optical flow are not the same, we often use optical flow as an approximation to the motion field since most flow is caused by observed motion.

Barron, Jepson and Fleet [25] provide an excellent overview of several optical flow algorithms that compares their performance. They divide the algorithms into four categories: differential, region-matching, energy-based, and phase-based. We will consider only the first two categories since they are the most popular.

Differential flow algorithms find solutions to a differential equation, the optical flow constraint equation [26],

$$I_x u + I_y v + I_t = 0$$

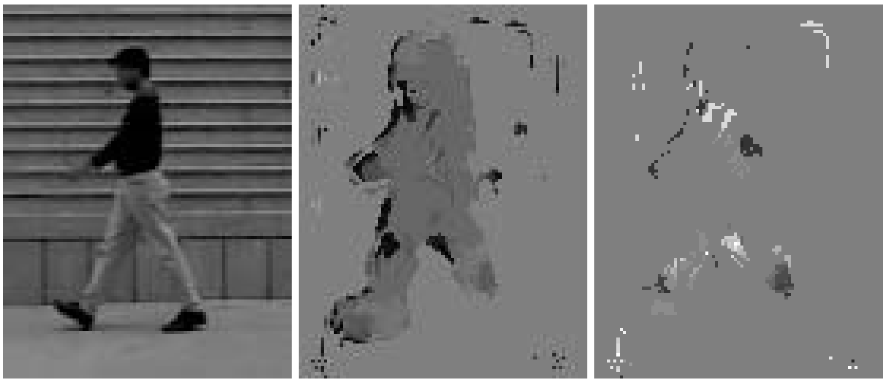
where I is the spatiotemporal (x, y, and t) image sequence, I_x, I_y, I_t are the partial derivatives of I with respect to space and time, and u and v are the x and y image velocities, i.e., the optical flow. Fig. 4 shows a sample frame of optical flow computed using the Lucas and Kanade [27] least-squares algorithm for differential flow.



(a) (b) (c) (d)

Fig. 4. Example of Lucas and Kanade [27] least squares optical flow: (a) original image from a sequence, (b) validity map, and (c) x- and (d) y-direction optical flow. In (b) black, gray and white mean no flow, gradient flow and least-squares flow respectively. In (c) and (d) gray is zero, black is negative (left/up), and white is positive (right/down).

Region-matching optical flow algorithms compute flow by comparing regions in consecutive images of a sequence. When regions match, the algorithms conclude that the region has moved and sets the flow accordingly. Fig. 5 shows an example of optical flow computed using the region-matching algorithm of Bulthoff et al. [28].



(a) (b) (c)

Fig. 5. Example of Bulthoff et al. [28] region-matching optical flow: (a) original image from a sequence, and (b) x- and (c) y-direction optical flow. In (b) and (c) gray is zero, black is negative (left/up), and white is positive (right/down).

3.4 Motion Energy and Motion History Images

Davis and Bobick [29] describe a motion energy image (MEI) and a motion history image (MHI), both derived from temporal image sequences. In the MEI, image pixels indicate whether or not there has been any motion at that pixel in previous frames. Note that an MEI cannot indicate in what order the pixels experienced the motion and therefore cannot encapsulate timing patterns in a motion. The MHI addresses this by indicating how recently motion occurred at each pixel. The brighter the region in an MHI, the more recent the motion. Fig. 6 shows images and the MHI from a sample sequence. Davis and Bobick [29] show that shapes in the MEI and MHI can be used to recognize various activities.

4 Evaluation of Gait Biometrics

4.1 Evaluation Methods

Typically, gait biometrics are tested in a recognition system like that shown in Fig. 7. The system extracts a set of descriptive features for an unknown test subject. It then compares the features to those of known subjects stored in a database. This model is adequate for evaluation of recognition and surveillance situations where there is no prior information provided about the identity of the subject.

Fig. 6. Example of a motion history image (MHI) [29]. The leftmost three images show the the motion sequence while the image on the right is the resulting MHI.

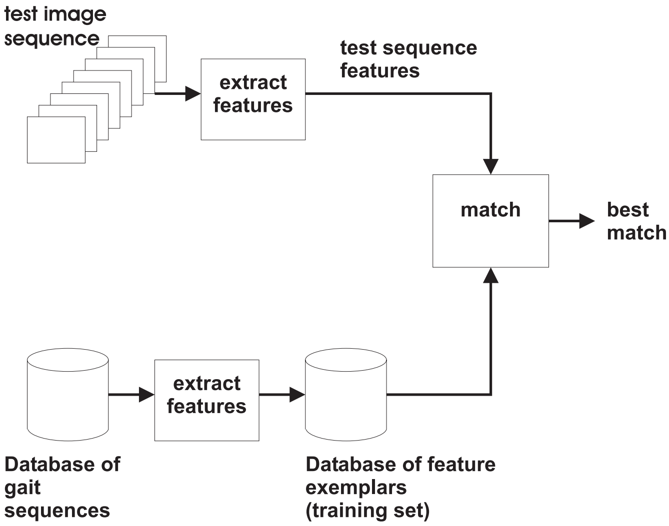


Fig. 7. Typical system for testing performance of gait recognition and other biometric systems.

Two broad approaches to evaluation have emerged. The first is to estimate the rate of correct recognition, while the second is to compare the variations in a population versus the variations in measurements. Neither method is entirely satisfactory, but they both provide insights into performance. We discuss both approaches in the remainder of this section.

4.2 Recognition Rate

Estimating the rate of correct recognition for a gait biometric has an intuitive appeal. It seems natural to think of system performance in terms of how often the system *gets it right*.

To arrive at such estimates, the procedure is to take a sample of the population of interest. One then divides the sample into two partitions, one for training the system (the database in Fig. 7), and one for testing. The estimated rate of correct recognition is the fraction of the test set that the system classifies correctly.

Such an estimate is extremely sensitive to context. Variations in any of the following factors will affect the resulting estimate.

- **Randomization of sample:** For the estimate to have any relevance outside the experiment, the sample must be a randomly selected from the population of interest. Such sampling is time-consuming and expensive. Consequently, most estimates produced in research are based on a biased sample that reflects mostly graduate. Campbell and Stanley [30] give one of the most thorough treatments of experimental design and the need for randomization.
- **Randomization of partitions:** It is essential that the training and test partitions be selected at random. Failure to do this can introduce a bias into the estimate. Cohen [31] gives excellent descriptions of methods for cross validation that avoid such biases.
- **Sampling conditions:** It is time-consuming to acquire samples over extended periods of time, and over a variety of imaging conditions. Thus, current samples are biased toward conditions in a single session using a single imaging apparatus. When researchers have reported results for samples that span weeks to months, e.g., Tanawongsuwan and Bobick [32], recognition rates drop drastically when compared to samples acquired in a single session.
- **Sample size:** Recognition rates drop with increases in sample size. For example, see the trends in the plots in Bhanu and Han [6] and Ben-Abdelkader et al. [33]. Intuitively, this occurs because the larger the sample, the more opportunities there are to make a mistake. In terms of the features used for recognition, as the sample size increases, the feature space becomes crowded, thus providing less resolution between individuals.

In spite of their intuitive appeal, recognition rates must be considered only within the context in which they are produced. Failure to consider any of the above factors in comparing recognition rates will almost certainly lead to false conclusions.

4.3 Analysis of Variance

While there is no way to avoid the issues of sample randomization, partition randomization, and sampling conditions, there are methods for dealing with variations in sample size. Consider the f statistic,

$$f = \frac{MS_{between}}{MS_{within}},$$

where $MS_{between}$ and MS_{within} are the mean-square errors between classes (between individuals) and within classes (for a single individual) due to the accumulation of all factors that cause a gait and its measured features to vary. When f is large, individuals are spread widely throughout the feature space with respect to the variations for an individual. When $f = 1$, then individuals are indistinguishable. A large f does not eliminate the trend toward lower recognition rates with sample size, but it does reduce the rate at which recognition deteriorates.

The f statistic is the foundation of analysis of variance (ANOVA) [34]. ANOVA is a method of hypothesis testing that uses the known distribution of f under the condition that classes/individuals are indistinguishable, also referred to as the null hypothesis. If a sample produces a value of f that is large enough, one rejects the null hypothesis and concludes that there is significant variation between classes. Note that sample size is a parameter of the known distributions of f, so f may be interpreted for samples of different size. Bobick and Johnson [35] describe *expected confusion*, $E[A]$, a number that is directly related to f ($E[A] = 1/\sqrt{f}$), and its role in predicting performance for varying sample size.

While f address issues of sample size, it is not clear how to compare f for different feature spaces, especially when data can be linear, as in a persons height, or directional, as in the phase of a signal. Directional ANOVA exists [36], but is it correct to compare the values of f directly. Furthermore, the distribution of f can depend on the dimensionality of the feature space. Currently, f appears to be a useful way to compare results acquired with different sample sizes, but it needs further development.

5 Existing Gait Recognition Systems

In this section, we describe and compare a selection of biometric gait recognition systems. As the previous section suggested, it is difficult to compare different systems directly when each is tested with a different sample. To address this issue here, in Fig. 8 we plot the recognition rate versus sample size for the methods that report recognition rates. Note that this does not adequately address all the issues of sampling, but serves only to provide an approximate picture of the *state-of-the-art* in gait recognition.

In the following subsections, we categorize the methods by their source of oscillations: shape, joint trajectory, self similarity, and pixel.

5.1 Shape Oscillations

Fig. 9 shows the *shape-of-motion* system developed by Little and Boyd [37]. The system uses optical flow to identify a moving figure in a sequence of images. It then describes the shape of the moving figure with a set of scalars derived from Cartesian moments. For example, the descriptors include the x and y coordinates of the object centroid, the x and y coordinates of the object centroid

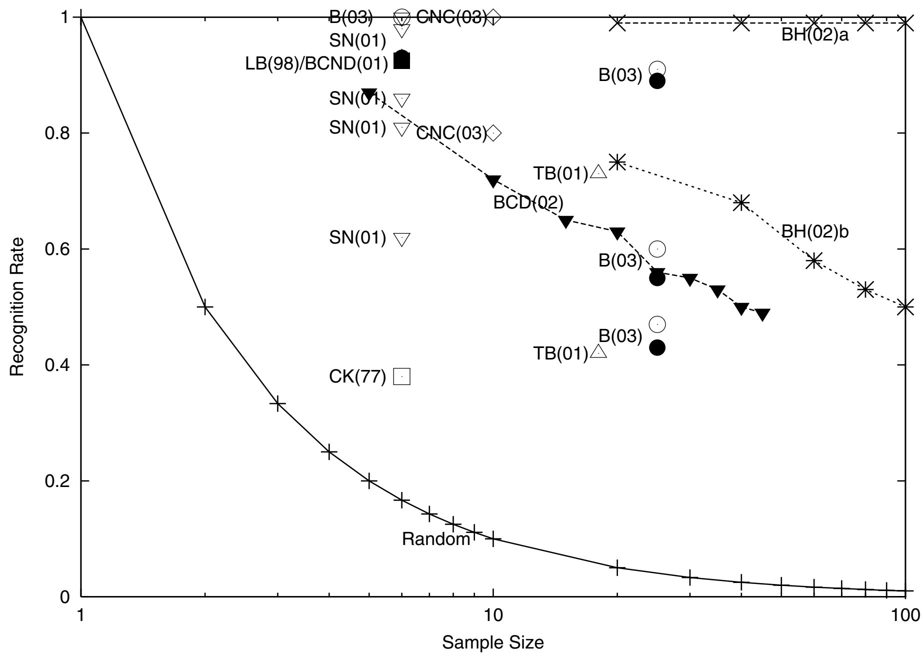


Fig. 8. Performance comparison of biometric gait recognition systems showing recognition rate versus sample size. The curve labeled **Random** indicates the expected recognition rate for random guesses. **CK(77)** refers to Cutting and Kozlowski [14], **BH(02)a** and **BH(02)b** refer to Bhanu and Han [6] 5mm and 40mm resolution respectively, **LB(98)** refers to Little and Boyd [37], **BCND(01)** refers to Ben-Abdelkader et al. [38], **SN(01)** refers to Shutler and Nixon [39], **TB(01)** refers to Tanawongsuwan and Bobick [32], **CNC(03)** refers to Cunado et al. [40], **B(03)** refers to Boyd [41], and **BCD(02)** refers to Ben-Abdelkader et al. [33].

weighted by the magnitude of the optical flow, and the aspect ratio of the distribution of pixels. When taken over the duration of the sequence, each scalar forms a time series. The shape-of-motion system extracts the oscillations from each series, then finds the frequency and phase of the oscillations, thus performing frequency entrainment and phase locking. The result is a set of m phases, one per scalar. The system takes one phase as a reference, then subtracts the reference to produce a feature vector of $m-1$ phases. In their evaluation, Little and Boyd achieved a recognition rate of approximately 92% for a sample size of six.

Shutler and Nixon [39] extend the shape-of-motion concept to use Zernike *velocity moments* to compute shape descriptions over an entire sequence, rather than on a frame by frame basis. They test their system on the shape-of-motion [37]

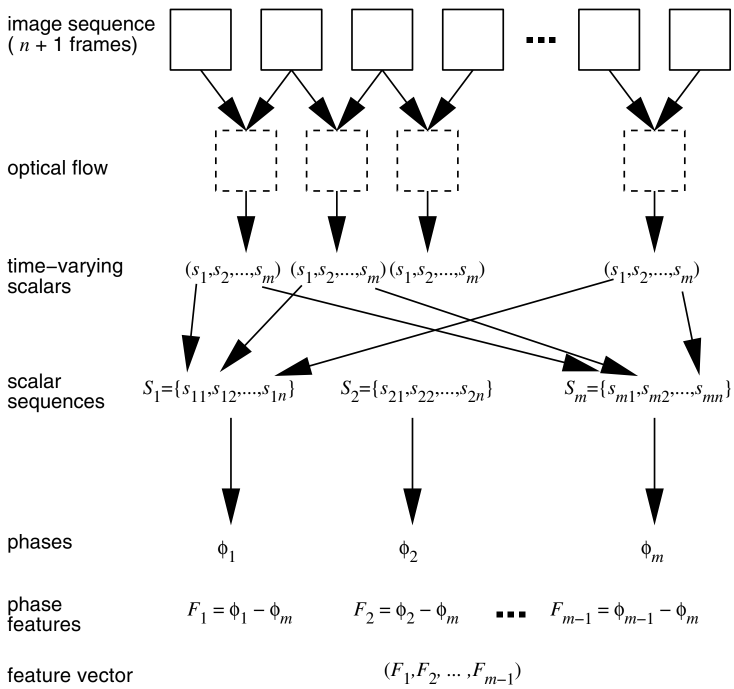
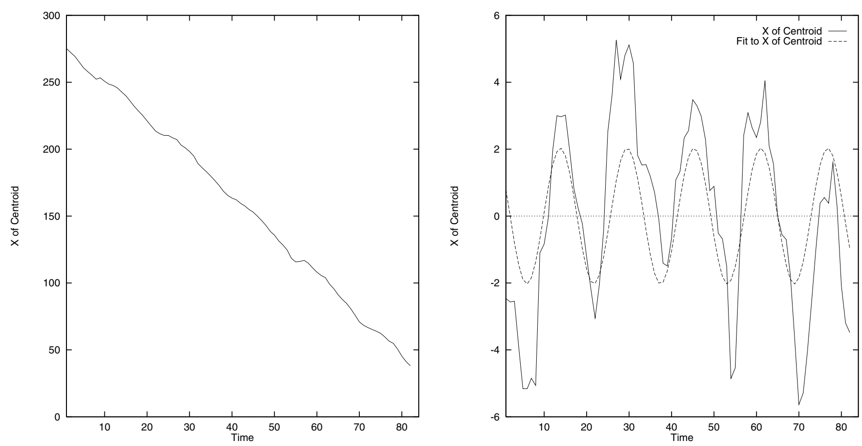


Fig. 9. The shape-of-motion gait recognition system [37].



(a) (b)

Fig. 10. Sample data from the shape-of-motion system: (a) a x-coordinate sequence, and (b) the sequence with the non-oscillatory component removed and a fitted sinusoid at the measured frequency and phase.

database, achieving recognition rates in the range of 62% to 100%, depending upon which velocity moments they include in their feature vector, for a sample size of six.

5.2 Joint Trajectory Patterns

Tanawongsuwan and Bobick (2001) [32] use joint angle trajectories measured using a magnetic-marker motion-capture system. As such, theirs is not a vision system and would not be practical for biometrics, but it does indicate the potential for joint angle trajectory features, if they were to be measured by some other means. They estimate the frequency of the gait and align the left and right, hip and knee joint trajectories to a common point in the gait cycle. They also resample the sequences to a common length. These steps effectively perform frequency entrainment and phase locking. The set of four trajectories combine to form one large feature vector used for recognition.

Tanawongsuwan and Bobick evaluated their system on a sample size of 18 and achieved a recognition rate of 73%. They further tested their system using an additional eight test sequences captured at a later date. When recognizing this latter sample using training data from the first sample, the recognition rate dropped to 42%. This demonstrates the deterioration in performance that occurs when samples span long periods of time.

Cunado et al. [40] extract a hip joint trajectory from a sequence of images. They acquire a trajectory for the hip closest to the camera only. They then use Fourier components of the trajectory as features for recognition. A test of their method on a database of size 10 yields recognition rates of 80% and 100% for Fourier features, and phase-weighted Fourier features respectively. Given the significance of phase locking in human perception of gaits, it is not surprising that the inclusion of phase information in the feature vector improves the recognition rate.

5.3 Temporal Patterns in Self-Similarity

As a person walks, the configuration of their body repeats periodically. For this reason, images in a gait sequence tend to be similar to other images in the sequence when separated in time by the period of the gait (the time between left foot strikes) and half the period (the time between left and right foot strikes). Fig. 11 illustrates this point.

Ben-Abdelkader et al. [38] exploit this *self similarity* to create a representation of gait sequences that is useful for gait recognition. From an image sequence, they construct a self-similarity image in which pixel intensities indicate the extent to which two images in the sequence are alike, i.e., pixel (i, j) in the self-similarity image indicates the similarity of the images at times t_i and t_j. With a cyclic motion such as a gait, the self-similarity image has a repeating texture. The frequency of the gait determines the rate at which the texture repeats (and thus is a form of frequency entrainment). Furthermore, variations in the timing of motions between individuals become details in the self-similarity

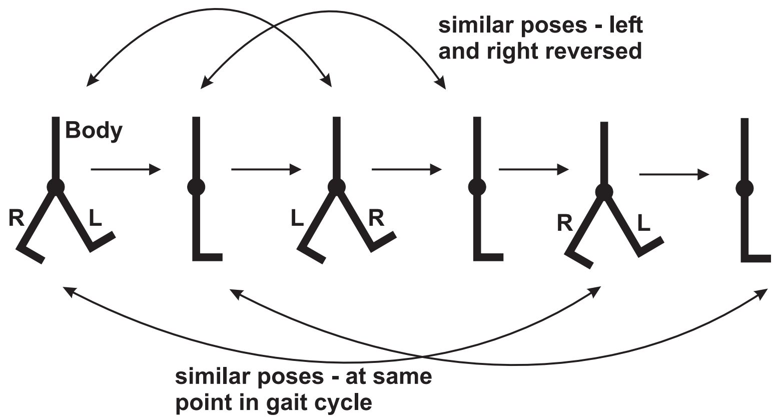


Fig. 11. Self similarity in gait sequences. Images separated by a full or half period of the gait tend to be alike.

image texture (and thus is a form of phase locking). Self-similarity images are large, so Ben-Abdelkader et al. use a principal component analysis on the space of similarity images to create a lower-dimensional eigenspace of images. The projections of self-similarity images onto this eigenspace become features for gait recognition.

Ben-Abdelkader et al. [38] test their system on the shape-of-motion database [37] and achieve a recognition rate of 93% with a sample size of six.

5.4 Pixel Oscillations

When a walker appears to be stationary in an image sequence, either as a result of tracking or walking on a treadmill, the cyclic motions of the gait result in intensity oscillations in pixels. The frequency of the gait and the timing of the component motions determine the frequency and phase of the pixel oscillations. Boyd [42] demonstrated that an array of phase-locked loops (PLL), one per pixel, can synchronize internal oscillators to the frequency and phase of pixel oscillations. This synchronization process inherently performs frequency entrainment and phase locking.

Boyd uses a phasor (Fig. 12), a complex number that represents a rotating vector, to represent the magnitude and phase of the oscillations at each pixel. Thus, once the PLL synchronization occurs, one can construct a complex image of phasors in which each pixel indicates the extent to which there are oscillations and the relative timing of the oscillations (Fig. 13). Procrustes shape analysis [43, 36] (Fig. 14) is a method for the statistical comparison of shapes represented as complex vectors. Thus, Procrustes shape analysis provides an ideal method to compare vectors of phasors that represent image oscillations.

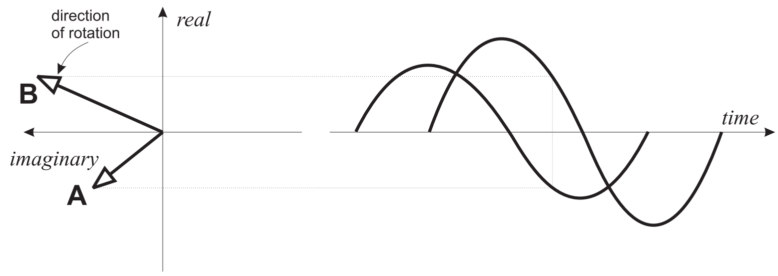


Fig. 12. A phasor, or phase vector, is a complex vector that rotates about the origin, generating a sinusoid when projected onto the real axis. The magnitude and direction of the vector gives the amplitude and phase of the sinusoid respectively. Timing is given by the relative phases. Here phasor **A** leads phasor **B**.

(a) (b) (c)

Fig. 13. Sample output of phase-locked loops: (a) superposition of frames from the input sequence, (b) magnitude of oscillations, and (c) phase of oscillations (note the phase wrap that results from the display of phase as a gray level).

Boyd [41] tested the phase-locked loops for the ability to recognize individual people using the shape-of-motion database [37] and the MoBo database [23]. With shape-of-motion data, recognition was perfect, 100% with sample size six. Using the MoBo database recognition rates were between 47% and 91% depending on whether or not sequences portraying the same style of gait were allowed to match. Boyd also observed that ignoring the phase information lowered the recognition rate in all cases.

In related work, Liu and Picard [44] Polana and Nelson [45] also look pixel level variations to analyze cyclic motions, however, they do not apply their analysis to biometric recognition. See Sec. 6 for more details.

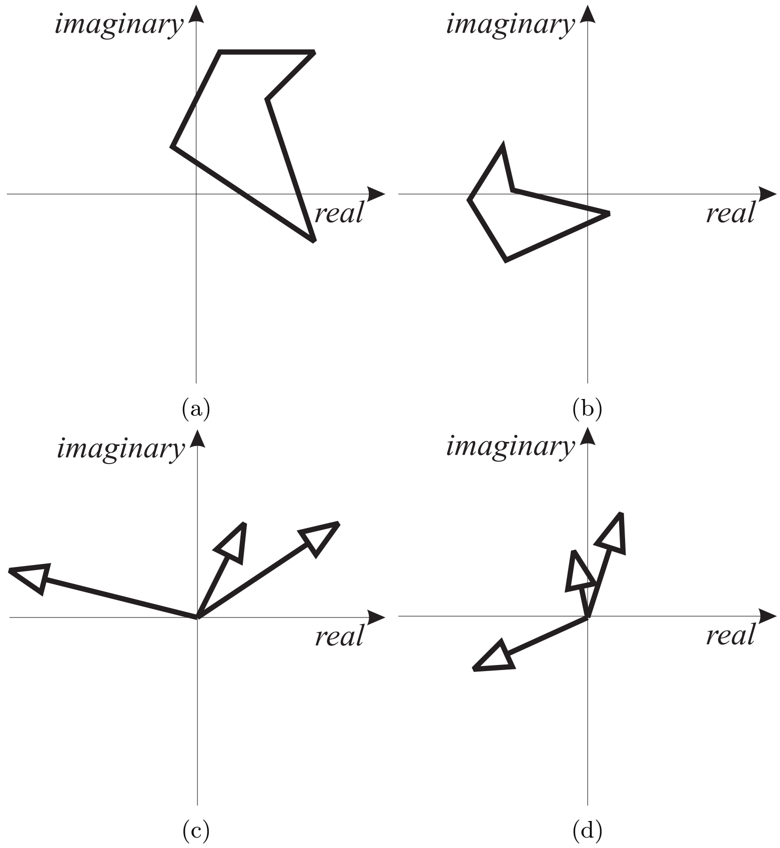


Fig. 14. Procrustes analysis applied to shape and phase configurations. In the conventional application, a shape is represented by a vector of complex vertices. The shape in (a), is the same as the shape (b) because each one is a translated, scaled, and rotated version of the other. A phasor configuration is also a vector of complex numbers. The configuration in (c) is the same as that in (d) because each one is a rotated and scaled version of the other. Rotation is always about the origin so translation can be ignored.

6 Other Systems

The methods described in this section are related to gait recognition, but are not mentioned in Sec. 5 because they are either quasi gait methods, not specific to gait, or do not do recognition. This is not to say that these methods are inferior, but that we merely choose to classify them differently.

6.1 Quasi Gait Recognition

Biometric recognition methods that do not rely on properties unique to a gait, but make measurements of a person during a gait, we refer to as quasi gait methods. One advantage to quasi gait approaches is that they may be less sensitive to variation in a gait. For example, a person's gait may vary for reasons discussed, but their skeletal dimensions will remain constant. Examples of quasi gait methods are discussed here.

Bobick and Johnson [35] measure a set of four parameters that describe a static pose extracted from a gait sequence. These parameters are height, torso length, leg length, and stride length, all of which can be estimated from a single image (see Fig. 15). Bobick and Johnson then use these parameters as feature vectors for recognition. The authors evaluate their method using *expected confusion*, a number related to the f statistic. For this reason we are unable to compare their results in the plot of Fig. 8.

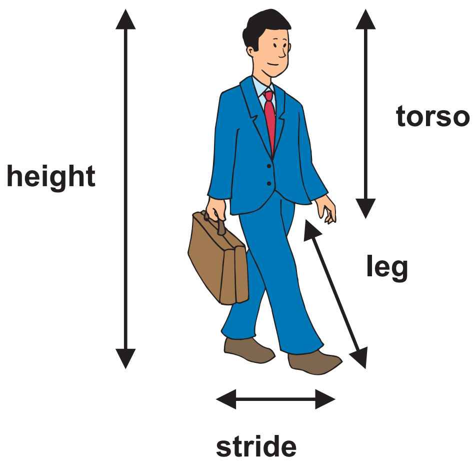


Fig. 15. Static gait features measured by Bobick and Johnson [35]: height, torso length, leg length, and stride length.

Ben-Abdelkader et al. [33] extract a subject's height, amplitude of height oscillations during gait, gait cadence, and stride length (see Fig. 16). They then use these values in a feature vector for recognition. Although the features include cadence, the method uses no timing information from the gait so we classify it as quasi gait recognition. Using the full feature set, they achieved a recognition rate of 49% with a sample size of 45, acquired over two days. They also look at subsamples to determine the rate at which performance deteriorated with sample size. The results are plotted in Fig. 8.

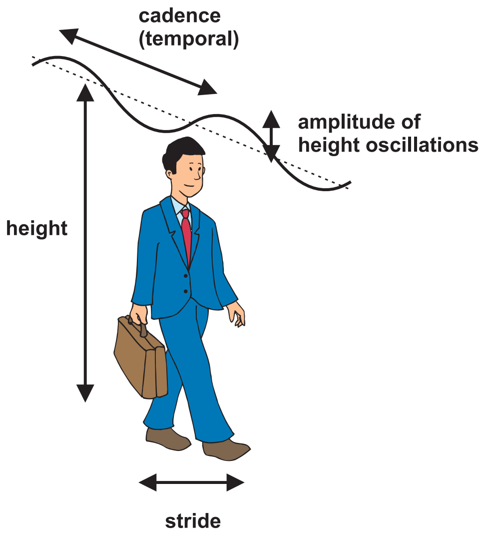


Fig. 16. Gait features measured by Ben-Abdelkader et al. [33]: mean height, amplitude of height oscillations, stride length and cadence.

It should be noted that these methods require some camera calibration and knowledge of the distance from camera to subject. This is done to obtain measurements in real-world units that can be measured with varying apparatus at different times.

6.2 Non-recognition Systems

Methods in this section either do not do gait recognition specifically, or do not do recognition at all.

Polana and Nelson [45] examine oscillations in the magnitude of the optical flow in a sequence containing periodic motion. They compute a coarse resolution (four by four) flow magnitude image at six points in the period of the motion. From this they form a 96-element vector that is used to recognize a broad range of periodic motions, but not individual gaits.

Liu and Picard [44] examine oscillations in pixel intensity for a gait sequence using fast Fourier transforms (FFT). Their analysis identifies the amplitude of the fundamental frequency of the gait. They did not use phase in their analysis, nor did they do recognition.

Baumberg and Hogg [46] describe a method that extracts the silhouette of a walking figure. They extend the concept by treating changes in shape with a vibration model [24]. They did not report testing their model for recognition.

From a sequence of images, Davis and Bobick [29] compute motion energy images (MEI) and motion-history images (MHI) that indicate where motion is occurring and how recently the motion occurred. They describe the shape of the moving regions with a set of Hu moments, which they in turn use to recognize patterns of motion, such as various aerobic exercises.

Several methods exist to match a kinematic model of a human to a sequence of video images, i.e., estimate a subject's pose. In general, these methods are not gait-specific, nor are they intended to do recognition. They may be viewed as methods for marker-less motion capture. Examples of these methods include work by Hunter et al. [47], Rowley and Rehg [48], Wachter and Nagel [49], Wren et al. [50], Bregler and Malik [51], and Morris and Rehg [52]. One problem with some model-based systems is that they are computationally intensive, which makes them either too slow or too expensive for use in a biometric system.

Bissacco et al. [53] extend results from acquisition of kinematic pose to recognition. They use Bregler's method [54] to extract joint angle trajectories from a motion sequence. They then compute an auto-regressive moving-average (ARMA) model of the joint movement which they in turn use as a feature vector for recognition. Their system can recognize different types of gaits such as running, walking, or walking a staircase. Although they did not test it for biometric gait recognition, this remains as a possibility.

7 Other Applications

Although the subject of this volume is biometrics, we feel it is worth noting some of the other applications that are related to biometric gait analysis.

One area of interest in gait analysis is gait-related pathology. Gait analysis can contribute in two important areas. The first is in diagnosis of gait-related disorders, and the second is in monitoring of treatment. Currently, the norm is to diagnose and monitor treatments using human observations. As in most applications of computer vision, we presume the machine can compensate for human deficiencies. In this case, we expect the machine to give consistent diagnoses and assessments of treatment that do not vary with the individual clinician, their training and experience, or their attentiveness at any particular moment.

Although improvements in human athletic performance are not likely to have an impact on quality of life for most people, athletics do have value as a source of entertainment. To that end, there is interest in evaluating human motion to predict athletic potential, or evaluate training.

Motion capture plays a vital role in the computer graphics and games industry. Currently, marker-based systems dominate industrial motion capture, but advances in human motion analysis are constantly improving marker-less systems. We expect that marker-less systems will eventually become the norm for motion capture.

8 Conclusions

Interest in gait-based biometrics has lead to a stream of recent results. Fig. 8, although not comprehensive, indicates what has been accomplished to date.

Clearly, the performance of gait recognition systems is below what is required for use in biometrics. When one considers that gait is best suited to recognition or surveillance scenarios where the databases are likely to be very large, one would expect high false alarm rates that will render a system useless. Furthermore, tests to date do not fully consider variation in gait measurements over long time spans, and under with different imaging conditions. Nevertheless, researchers are making progress and understanding more about gait with each new development. Areas that need further investigation include studies on variability with terrain, footwear, long time spans, and other confounding factors, in an effort to find gait features that vary only with the individual.

References

[1] McGrath, B.: The week in walks. The New Yorker **June 2** (2003) 35
[2] Johansson, G.: Visual perception of biological motion and a model for its analysis. Perception and Psychophysics **14** (1973) 201–211
[3] Johansson, G.: Visual motion perception. Scientific American (1975) 76–88
[4] Bertenthal, B.I., Pinto, J.: Complementary processes in the perception and production of human movements. In Smith, L.B., Thelen, E., eds.: A Dynamic Systems Approach to Development: Applications. MIT Press, Cambridge, MA (1993) 209–239
[5] Cohen, L., Shipley, T.F., Marshark, E., Taht, K., Aster, D.: Detecting animals in point-light displays. In: Twenty Second Annual Meeting of the Cognitive Science Society, Philadelphia, PA (2000) 70
[6] Bhanu, B., Han, J.: Bayesian-based performance prediction for gait recognition. In: IEEE Workshop on Motion and Video Computing, Orlando, Florida (2002) 145–150
[7] McGeer, T.: Passive dynamic walking. The International Journal of Robotics Research **9** (1990) 62–82
[8] McGeer, T.: Passive walking with knees. In: IEEE International conference on Robotics and Automation. (1990) 1640–1645
[9] Coleman, M.J., Ruina, A.: An uncontrolled toy that can walk but cannot stand still. Physcial Review Letters **80** (1998) 3658–3661
[10] Garcia, M., Chatterjee, A., Ruina, A., Coleman, M.: The simplest walking model: stability, complexity, and scaling. ASME Journal of Biomechanical Engineering **120** (1998) 281–288
[11] Collins, S.H., Wisse, M., Ruina, A.: A 3-d passive-dynamic walking robot with two legs and knees. International Journal of Robotics Research (2001) In Press.
[12] Barclay, C.D., Cutting, J.E., Kozlowski, L.T.: Temporal and spatial factors in gait perception that influence gender recognition. Perception and Psychophysics **23** (1978) 145–152
[13] Kozlowski, L.T., Cutting, J.E.: Recognizing the sex of a walker from a dynamic point-light display. Perception and Psychophysics **21** (1977) 575–580
[14] Cutting, J.E., Kozlowski, L.T.: Recognizing friends by their walk: gait perception without familiarity cues. Bulletin of the Psychonomic Society **9** (1977) 353–356
[15] Laszlo, J., van de Panne, M., Fiume, E.: Limit cycle control and its application to the animation of balancing and walking. In: SIGGRAPH 96. (1996) 155–162
[16] Murray, M.P., Bernard, A., Kory, R.C.: Walking patterns of normal men. The Journal of Bone and Joint Surgery **46A** (1964) 335–359

[17] Murray, M.P.: Gait as a total pattern of movement. American Journal of Physical Medicine **16** (1967) 290–332

[18] von Tscharner, V., Goepfert, B.: Gender dependent emgs of runners resolved by time/frequency and principal pattern analysis. Journal of Electromyography and Kinesiology **13** (2003) 253–272

[19] Sudderth, E., Hunter, E., Kreutz-Delgado, K., Kelly, P.H., Jain, R.: Adaptive video segmentation: theory and real-time implementation. In: 1998 Image Understanding Workshop. Volume 1. (1998) 177–181

[20] Horprasert, T., Harwood, D., Davis, L.: A statistical approach for real time robust background subtraction. In: IEEE Frame Rate Workshop. (1999)

[21] Stauffer, C., Grimson, W.E.L.: Adaptive background mixture models for real-time tracking. In: Computer Vision and Pattern Recognition 99. Volume II. (1999) 246–252

[22] Javed, O., Shafique, K., Shah, M.: A hierarchical approach to robust background subtraction using color and gradient information. In: IEEE Workshop on Motion and Video Computing, Orlando, FL (2002) 22–27

[23] Gross, R., Shi, J.: The cmu motion of body (mobo) database. Technical Report CMU-RI-TR-01-18, Robotics Institute, Carnegie Mellon University (2001)

[24] Baumberg, A.M., Hogg, D.C.: Learning spatiotemporal models from training examples. In: British Machine Vision Conference, Birmingham (1995)

[25] Barron, J.L., Fleet, D.J., Beauchemin, S.S.: Performance of optical flow techniques. International Journal of Computer Vision **12** (1994) 43–77

[26] Horn, B.K.P.: Robot Vision. MIT Press, Cambridge, Massachusetts (1986)

[27] Lucas, B.D., Kanade, T.: An iterative image registration technique with an application to stereo vision. In: International Joint Conference on Artificial Intelligence. (1981) 674–679

[28] Bulthoff, H., Little, J.J., Poggio, T.: A parallel algorithm for real-time computation of optical flow. Nature **337** (1989) 549–553

[29] Davis, J.W., Bobick, A.F.: The representation and recognition of human movement using temporal templates. In: IEEE Computer Vision and Pattern Recognition. (1997) 928–934

[30] Campbell, D.T., Stanley, J.C.: Experimental and Quasi-Experimental Designs for Research. Houghton Mifflin, Boston (1963)

[31] Cohen, P.R.: Empirical Methods for Artificial Intelligence. MIT Press, Cambridge, Massachusetts (1995)

[32] Tanawongsuwan, R., Bobick, A.: Gait recognition from time-normalized joint-angle trajectories in the walking plane. In: Computer Vision and Pattern Recognition 2001. Volume II., Kauai, HI (2001) 726–731

[33] Ben-Abdelkader, C., R.Cutler, Davis, L.: Person identification using automatic height and stride estimation. In: 16th International Conference on Pattern Recognition, Quebec, Quebec (2002) 377–380

[34] Winer, B.J.: Statistical principles in experimental design. McGraw-Hill, New York (1971)

[35] Bobick, A., Johnson, A.: Gait recognitin using static activity-specific parameters. In: Computer Vision and Pattern Recognition 2001. Volume I., Kauai, HI (2001) 423–430

[36] Mardia, K.V., Jupp, P.E.: Directional statistics. Wiley, Chichester (2000)

[37] Little, J.J., Boyd, J.E.: Recognizing people by their gait: the shape of motion. Videre **1** (1998) 1–32

[38] Ben-Abdelkader, C., R.Cutler, Nanda, H., Davis, L.: Eigengait: motion-based recognition of people using image self-similarity. In: Audio- and Video-Based Biometric Person Authentication, Halmstad, Sweden (2001)
[39] Shutler, J.D., Nixon, M.S.: Zernike velocity moments for description and recognition of moving shapes. In: British Machine Vision Conference 2001, Manchester, UK (2001) Session 8: Modelling Behaviour
[40] Cunado, D., NIxon, M.S., Carter, J.N.: Automatic extraction and description of human gait models for recognition purposes. Computer Vision and Image Understanding **90** (2003) 1–41
[41] Boyd, J.E.: Synchronization of oscillations for machine perception of gait. In review (2003)
[42] Boyd, J.E.: Video phase-locked loops in gait recognition. In: International Conference on Computer Vision, Vancouver, BC (2001) 696–703
[43] Kent, J.T.: New directions in shape analysis. In Mardia, K.V., ed.: The art of statistical science: a tribute to g. s. watson. Wiley, Chichester (1992) 115–127
[44] Liu, F., Picard, R.W.: Finding periodicity in space and time. In: International Conference on Computer Vision. (1998)
[45] Polana, R., Nelson, R.: Detection and recognition of periodic, nonrigid motion. International Journal of Computer Vision **23** (1997) 261–282
[46] Baumberg, A.M., Hogg, D.C.: Learning flexible models from image sequences. Technical Report 93.36, University of Leeds School of Computer Studies (1993)
[47] Hunter, E.A., Kelly, P.H., Jain, R.C.: Estimation of articulated motion using kinematically constrained mixture densities. In: Nonrigid and Articulated Motion Workshop, San Juan, Peurto Rico (1997)
[48] Rowley, H.A., Rehg, J.M.: Analyzing articulated motion using expectation-maximization. In: Computer Vision and Pattern Recognition 97, San Juan, Peurto Rico (1997) 935–941
[49] Wachter, S., Nagel, H.H.: Tracking of persons in monocular image sequences. In: Nonrigid and Articulated Motion Workshop, San Juan, Peurto Rico (1997)
[50] Wren, C., Azarbayenjani, A., Darrell, T., Pentland, A.P.: Pfinder: real-time tracking of the human body. IEEE Transactions on Pattern Analysis and Machine Intelligence **19** (1997) 780–785
[51] Bregler, C., Malik, J.: Tracking people with twists and exponential maps. In: Computer Vision and Pattern Recognition 1998, Santa Barbara (1998)
[52] Morris, D.D., Rehg, J.M.: Singularity analysis for articulated object tracking. In: Computer Vision and Pattern Recognition 98, Santa Barbara, CA (1998) 289–296
[53] Bissacco, A., Chiuso, A., Ma, Y., Soatto, S.: Recognition of human gaits. In: Computer Vision and Pattern Recognition 2001. Volume II., Kauai, HI (2001) 52–57
[54] Bregler, C.: Learning and recognizing human dynamics in video sequences. In: Computer Vision and Pattern Recognition 1997, San Juan, Puerto Rico (1997) 568–574