G

Gabor Jets

Gabor jets are a set of filters that are used to extract the local frequency information from the face images. These filters are generally linear filter with impulse responses defined by a harmonic function and a Gaussian function. The Fourier transform of a Gabor filter's impulse response is the convolution of the Fourier transform of the harmonic function and the Fourier transform of the Gaussian function.

► Face Recognition, Component-Based

Gabor Transform

A complete representation of a signal or image in terms of coefficients on Gabor wavelets, such that the original data can be reconstructed exactly by combining together those wavelets using their computed coefficients. A complication is that the necessary coefficients cannot be obtained simply by operations of filtering or by the inner product projections of the data with the wavelets, since they do not constitute an orthogonal basis. More complex methods are required (biorthogonal bases; relaxation networks) to obtain the needed expansion coefficients from projection coefficients. Once obtained, a Gabor Transform is a powerful tool for signal or image encoding, analysis, and compression.

► Iris Encoding and Recognition using Gabor Wavelets

Gabor Wavelets

Complex exponentials (Fourier components) multiplied by Gaussian envelopes. Although they fail to satisfy some parts of the stricter mathematical definitions of wavelets, such as orthogonality and compact support, these elementary functions can constitute a powerful basis for signal or image encoding, representation, compression, and analysis. They are increasingly used today in computer vision and in pattern recognition, particularly in biometrics, where they are the basis of iris recognition and have also been used for several other biometric modalities. Among their advantages (besides forming a complete basis for signal or image encoding) are; their optimality under the Heisenberg Uncertainty Principle for simultaneous resolution in time/space and in frequency; their closed analytical form; their self-Fourier property and closure under convolution and multiplication; and their neurobiological basis in the receptive field profiles of neurons in the mammalian visual cortex. Their chief disadvantage is that they are not mutually orthogonal, and so the projection coefficients obtained by computing their inner product with an image are not the same as the expansion coefficients that would be needed to reconstruct the same image exactly from them.

► Face Recognition, Component Based
► Iris Encoding and Recognition using Gabor Wavelets
► Local Image Features
► Local Image Filters

Gait

The manner of a person's movement, specifically during walking is called gait. The human gait cycle consists of two main phases: during stance phase, the foot is on the ground, and during the swing phase, the leg is swinging forward in preparation for the next ground contact.

► Gait, Forensic Evidence of

Gait Analysis

► Gait, Forensic Evidence of

Gait Biometrics, Overview

Rama Chellappa, Ashok Veeraraghavan
Narayanan Ramanathan
University of Maryland, College Park, MD, USA

Synonym

Gait recognition

Definition

Gait is defined as the style or manner of walking. Studies in physchophysics suggest that people can identify familiar individuals using just their gait. This has led to a number of automated vision based algorithms that use gait as a biometric. Such a system usually consists of a video camera capturing images of a person walking within its field of view. Appropriate features such as joint angles or silhouettes are extracted from this video and are then used to compare with the stored gait signatures of known individuals. As with any other biometric system, the system can operate in both the identification and the verification mode. Gait as a biometric has several advantages compared to traditional biometrics such as fingerprint in that gait is non-intrusive, does not require cooperation from the individual, and can function at moderate distances from the subject.

Introduction

The study of human gait has gathered pace in recent years driven primarily by its potential as a biometric. Gait-based person authentication has several significant advantages compared to traditional biometrics such as fingerprint or iris. Firstly, gait based biometric systems do not require the individuals to be cooperative since the input of these systems is the video feed captured by passive cameras. Secondly, gait is a non-intrusive biometric – it does not require the individuals to wear any special equipment in order to be recognized. Thirdly, gait based biometric systems have an extended range compared to traditional biometrics – they can operate reliably even when the subjects are tens of meters away from the camera. Finally, such a system harnesses the potential of thousands of surveillance video cameras installed in public locations into a biometric authentication system.

Operation of a Gait Based Biometric System

The sensor for a gait-based biometric system is a video camera capturing videos of human subjects walking within its field-of view. The raw sensor video is then processed to extract relevant features which can then be used for recognition. If the acquisition conditions are expected to be controlled and favorable, then the quality of the video will enable the extraction of features such as joint angles from the individual video frames. In more typical uncontrolled settings, the features extracted could either be background subtracted binary images, silhouettes, shapes or width vectors – all examples of features capturing the extent of the human body to differing amounts of detail. During the training phase, several such sequences of each individual in the gallery are collected and the appropriate features are then stored in the database. During the test phase, each test sequence is compared with the training sequences available in the database and the similarity is used to perform person authentication.

Challenges for Gait Based Biometrics Systems

The discriminative information in gait is present in both the shape of the individual and also in the manner of his/her gait. This means that gait based biometric systems must be able to model gait as a time series of features or as a dynamical model in order to perform accurate recognition. Static template based methods which have been used for most other biometric systems need to be adapted to a temporal sequence in order to achieve robust performance. In this regard, another challenge is time alignment of two sequences so that critical events during gait like "mid-stance", "toe-off " etc. are time aligned accurately so that recognition performance is not affected by inaccurate time alignment between postures that occur during gait. Since, gait based person identification often occurs without any particular viewpoint, view-invariance of the feature extracted from the video is another important challenge. This will ensure that recognition performance is robust to changes in the viewpoint of the camera. In scenarios with moderate amounts of acquisition control, one can set up multiple video cameras so as to ensure that the best possible viewpoint which happens to be the fronto-parallel gait is captured on atleast one of the cameras. Another challenge for automated gait-based biometrics is that of changing illumination conditions in the scene. In order to be robust to changing illumination conditions, background subtraction is typically performed on the raw videos before the video data is used in a recognition algorithm. Finally, another important challenge is the variability in the clothing, shoe type and the surface on which the individuals walk. Obviously, the clothing of the subject especially their type of footwear has significant impact on the gait features observed and it is important to bear this in mind while developing gait-based biometric systems.

Features for Gait Based Biometrics

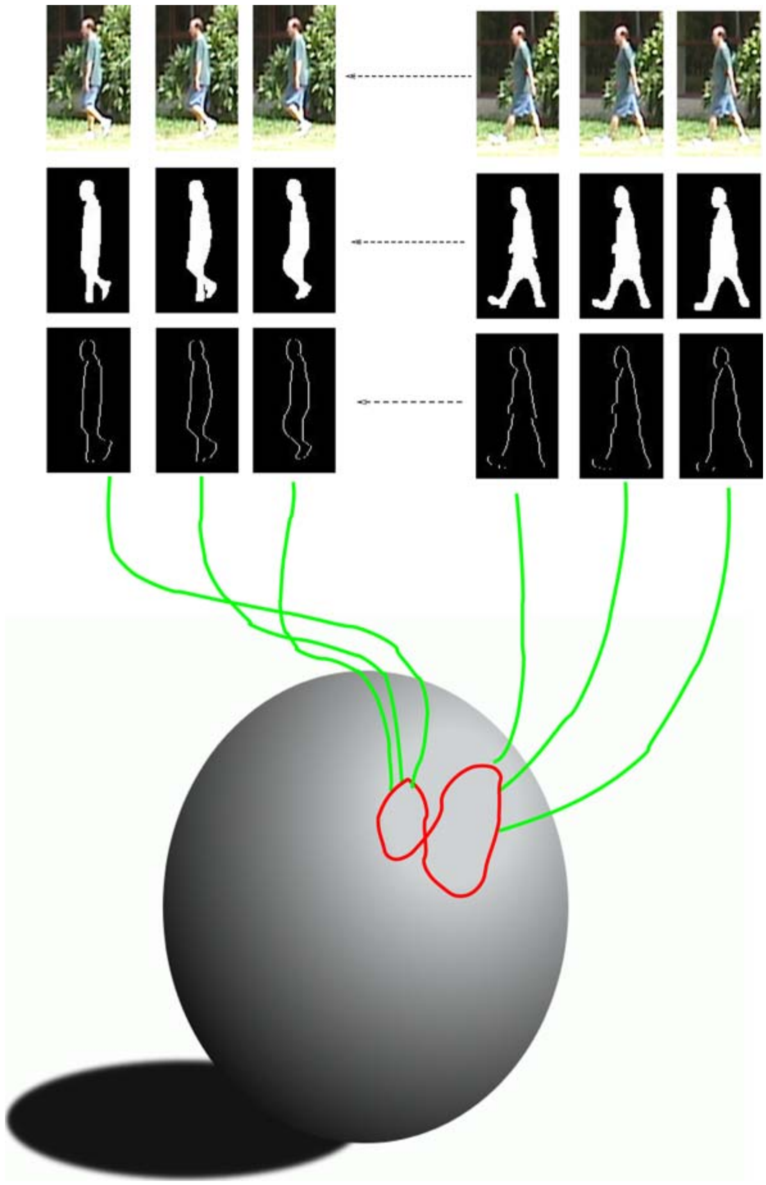
Silhouette: In most gait-based biometric systems the cameras can be assumed to be static during the short duration of time that they capture the gait of a single individual for verification. This allows simple background models to be built for each of these cameras. Background subtraction then identifies the set of all pixels in the image that belong to the moving individual. Figure 1 shows a sequence of color images captured by a video camera as a person walks through its field of view. Shown below are the binary background subtracted images in which all pixels belonging to the individual are white, while the background is black. This binary image is then scaled to a uniform size so that the feature extracted is independent of the distance of the camera from the subject. Several algorithms for gait based person identification use this binary silhouette as a feature [1–5].

Shape: "Shape is all the geometric information that remains when location, scale, and rotational effects are filtered out from the object"[6]. Kendall's statistical shape is a sparse descriptor of the shape that describes the shape configuration of k landmark points in an m-dimensional space as a $k \times m$ matrix containing the coordinates of the landmarks. Image space is 2-dimensional and therefore it is convenient to describe the shape vector as a k dimensional complex vector. First, a binarized silhouette denoting the extent of the object in an image is obtained. A shape feature is then extracted from this binarized silhouette. This feature vector must be invariant to translation and scaling since the object's identity should not depend on the distance of the object from the camera. So any feature vector that we obtain must be invariant to translation and scale. This yields the pre-shape of the object in each frame. Pre-shape is the geometric information that remains when location and scale effects are filtered out. Let the configuration of a set of k landmark points be given by a k-dimensional complex vector containing the positions of landmarks. Let us denote this configuration as X. Centered pre-shape is obtained by subtracting the mean from the configuration and then scaling to norm one. The centered pre-shape is given by

$$Z_c = \frac{CX}{\| CX \|}, \ \textit{where}\ C = I_k - \frac{1}{k} 1_k 1_k^T, \qquad (1)$$

where I_k is a k $\times$ k identity matrix and 1_k is a k dimensional vector of ones.

The advantage of using shape feature is that the differential geometric properties of the spherical manifold in which the shapes lie are very well understood and therefore, appropriate distance measures that can account for translational, rotational and scale invariances are well defined. For example, consider two complex configurations X and Y with corresponding preshapes α and β. The full Procrustes distance



Gait Biometrics, Overview. Figure 1 Graphical illustration of a video sequence obtained during a walking cycle and the corresponding features – silhoeuette and shape. Courtesy [7].

between the configurations X and Y is defined as the Euclidean distance between the full Procrustes fit of α and β and is chosen so as to minimize

$$d(Y, X) = \| \beta - \alpha s e^{j\theta} - (a + jb)1_k \|, \qquad (2)$$

where s is a scale, θ is the rotation and $(a+jb)$ is the translation. The full Procrustes distance is the minimum Full Procrustes fit i.e.,

$$d_F(Y, X) = \inf_{s,\theta,a,b} d(Y, X). \qquad (3)$$

The extracted shape sequence is shown in the bottom row of Figure 1 with a graphical illustration of the spherical manifold in which shapes lie. Shape is a very popular feature for gait-based biometrics and several state of the art algorithms perform gait matching as a matching of a sequence of shapes [7–10].

Joint Angles: A very popular feature for gait analysis in the medical and the psychophysics community is the joint angles – i.e., the angles made at each of the limb joints such as the knee, elbow ankle, wrist etc. There have been a few gait based biometrics algorithms that use joint angles as the feature for matching [11, 12]. The advantage of using joint angles as a feature is the fact that view-invariance is automatically achieved while using joint angles as a feature. Nevertheless, the essential problem with using joint angles is the fact that it is very challenging to robustly estimate them from uncontrolled monocular video sequences.

Algorithms for Matching

Most of the features described above have incorporated modest forms of view-invariance (atleast scale and translational invariance) as a part of the feature. Therefore the essential task of the algorithm for matching would be to model the dynamics of the feature during gait and use this to perform matching in a manner that is fairly insensitive to the speed of walking.

Dynamic Time Warping (DTW): Dynamic time warping is an algorithm for estimating the non-linear time synchronization between two sequences of features. The two sequences could be of differing lengths. Experiments indicate that the intra-personal variations in gait of a single individual can be better captured by non-linear warping rather than by linear warping [13]. The DTW algorithm which is based on dynamic programming computes the best non-linear time normalization of the test sequence in order to match the template sequence, by performing a search over the space of all allowed time normalizations. The space of all time normalizations allowed is cleverly constructed using certain temporal consistency constraints. Several gait-based biometrics algorithms have used the Dynamic time warping algorithm in order to time synchronize and match gait sequences [7, 8]. Recently, the DTW algorithm has also been extended so as to learn the warping constraints in a class-specific manner in order to improve discrimination between individuals [9].

Hidden Markov Model (HMM) The Hidden Markov Model (HMM) is a statistical state space model in which the observed shape sequence is modeled as outputs of a hidden states whose transitions are assumed to be Markovian. The model parameters of the HMM encode both the transition probabilities between the hidden states and the outputs of hidden states. The advantage of using a HMM is that there exists a wealth of literature on learning the parameters of the HMM and to perform inference using the HMM. Typically, the model parameters for each individual in the gallery is learnt and stored during the training phase. During the test phase, the probability of the observation sequence conditioned on the model parameters is maximized in order to perform recognition. The HMM [2, 3] and its many variants [14] have been successfully used for gait based person identification.

Autoregressive Moving Average Model (ARMA): Matching gait biometrics essentially is a problem of matching time-series data where the feature at each time instant is a silhouette or shape or joint angles. Therefore traditional time series modeling approaches such as the autoregressive model (AR) and the autoregressive moving average (ARMA) model have also been successfully used for gait based person identification. The model parameters of the ARMA model are learnt from the training sequences and stored. Given a test sequence, the model parameters for the test sequence are learnt and the distance between the model parameters is used in order to perform recognition [7].

Model Based Approaches

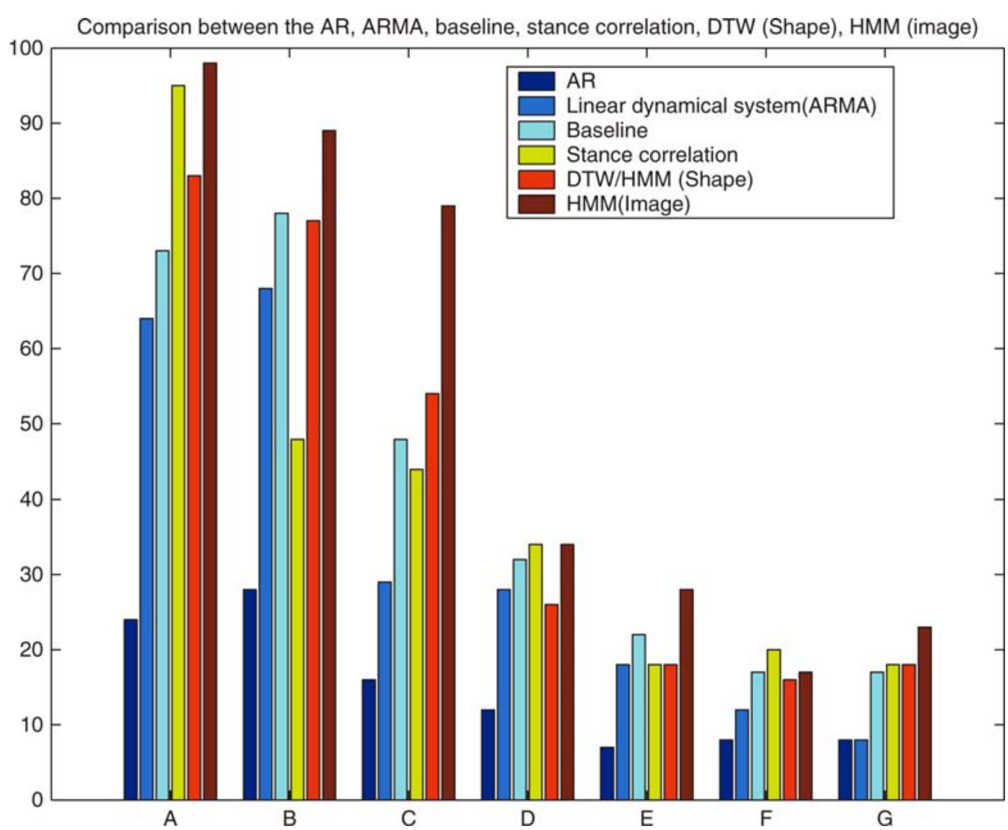
Typical feature based approaches first compute a sequence of features from each video and then match the sequence of features obtained in the test video to those stored in the gallery. Model-based approaches are different in the sense that they fit the sequence of features to a physical model of the human body and its inherent dynamics. For example, a model-based feature extraction process guided principally by biomechanical analysis for gait-based person identification is proposed [15]. The shape model for human subjects is composed of an ellipse to describe the head and the torso, quadrilaterals to describe the limbs and rectangles to describe the feet. Anatomical data is first used in order to derive shape and motion models that are consistent with normal human body proportions. Prototype gait motion models are then adapted to individuals using the specific characteristics of the extracted features. These individual specific shape and motion models are then used for gait recognition. A systematic analysis of the model-based approach also showed that cadence and static shape parameters of the human body account for most of the recognition performance.

Experiments on the USF Gait Data

In order to quantitatively test the performance and the viability of gait based biometrics a challenging gait database of 122 individuals was collected at the University of South Florida [4] as part of the DARPA Human Identification at a Distance (HID) program. The entire dataset containing over 1,200 videos was separated into 12 different experiments with varying levels of difficulty. The different challenge experiments amounted to varying different covariates during gait, like viewpoint, clothing, surface type, shoe type, and time etc. A bar plot of the recognition performance of various algorithms on the USF dataset (Experiments A-G) is shown in Figure 2. Experiments A,B and C correspond to changes in "view", "shoe type" and "view + shoe type" respectively without any change in the surface of walking, while challenege experiments D,E,F and G correspond to changes in the surface type from grass to concrete. The experiments indicate that changes in the surface type has significant impact on the recognition performance while view, shoe type affects recognition performance to a much lesser degree.

Summary

Gait is thus a novel biometric that provides significant operational advantages over several other biometrics such as face, fingerprint, iris etc. Unlike traditional biometrics like fingerprint, gait does not require the active cooperation of the subjects. Moreover, gait is a medium range biometric in the sense that acquisition distances can be as large as tens of meters. Moreover, in most operational scenarios, it is non-intrusive and does not require the subject to wear any special clothing. Preliminary experiments into gait as a biometric seem to indicate that the discriminative power of gait is not as strong as that of traditional biometrics such as fingerprints or iris. Therefore, several successful investigations for fusing the gait biometric with other



Gait Biometrics, Overview. Figure 2 Comparison of various algorithms on the USF gait database. (Courtesy [1]).

traditional biometrics in order to boost the identification performance have been performed and this seems to be an area of immense potential [16].

Related Entries

- ▶ Biometrics, Overview
- ▶ Covariates
- ▶ Multibiometrics
- ▶ Surveillance

References

1. Veres, G., Gordon, L., Carter, J., Nixon, M.: What image information is important in silhouette-based gait recognition? In: Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision Pattern Recognition (CVPR 2004), Washington, DC, 27 June 2004, vol. 2 (2004)
2. Kale, A., Sundaresan, A., Rajagopalan, A., Cuntoor, N., Roy-Chowdhury, A., Kruger, V., Chellappa, R.: Identification of humans using gait. IEEE Trans. Image Process. **13**(9), 1163–1173 (2004)
3. Sundaresan, A., RoyChowdhury, A., Chellappa, R.: A hidden markov model based framework for recognition of humans from gait sequences. In: Proceedings of 2003 International Conference on Image Processing, Barcelona, Spain, vol. 2, pp. 93–96 (2004)
4. Sarkar, S., Phillips, P., Liu, Z., Vega, I., Grother, P., Bowyer, K.: The human ID gait challenge problem: data sets, performance, and analysis. Pattern Anal. Mach. Intell. IEEE Trans. **27**(2), 162–177 (2005)
5. Man, J., Bhanu, B.: Individual recognition using gait energy image. IEEE Trans. Pattern Anal. Mach. Intell. **28**(2), 316–322 (2006)
6. Dryden, I., Mardia, K.: Statistical shape analysis. Wiley, Chichester (1998)
7. Veeraraghavan, A., Roy-Chowdhury, A., Chellappa, R.: Matching shape sequences in video with applications in human movement analysis. Pattern Anal. Mach. Intell. IEEE Transactions on **27**(12), 1896–1909 (2005)
8. Veeraraghavan, A., Chowdhury, A., Chellappa, R.: Role of shape and kinematics in human movement analysis. In: Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, (CVPR 2004), Washington DC, USA, 27 June-2 July 2004, vol. 1, pp. I–730 – I–737 (2004)
9. Veeraraghavan, A., Chellappa, R., Roy-Chowdhury, A.: The function space of an activity. In: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Washington, DC, USA, vol. 1, pp. 959–968 (2006)
10. Wang, L., Ning, H., Hu, W., Tan, T.: Gait recognition based on procrustes shape analysis. In: Proceedings of the Nineth IEEE International Conference on Image Processing (ICIP, Poster), Rochester, New York, USA, vol. III (2002)
11. Cunado, D., Nash, J., Nixon, M., Carter, J.: Gait extraction and description by evidence-gathering. In: Proceedings of the International Conference on Audio and Video Based Biometric Person Authentication, vol. 48 (1995)
12. Bissacco, A., Chiuso, A., Ma, Y., Soatto, S.: Recognition of human gaits. In: Conference on Computer Vision and Pattern Recognition, Hawaii, USA, vol. 2, pp. 52–57 (2001)
13. Forner-Cordero, A., Koopman, H., van der Helm, F.: Describing gait as a sequence of states. J. Biomech. **39**(5), 948–957 (2006)
14. Liu, Z., Sarkar, S.: Improved gait recognition by gait dynamics normalization. IEEE Trans. Pattern Anal. Mach. Intell. **28**(6), 863–876 (2006)
15. Wagg, D., Nixon, M.: On automated model-based extraction and analysis of gait. In: Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition, 2004, Seoul, Korea, pp. 11–16 (2004)
16. Kale, A., Roychowdhury, A., Chellappa, R.: Fusion of gait and face for human identification. In: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, 2004,(ICASSP'04), Montreal, Quebec, Canada, vol. 5 (2004)

Gait Models for Biometrics

▶ Gait Recognition, Model-Based

Gait Recognition

▶ Evaluation of Gait Recognition
▶ Gait Biometrics, Overview

Gait Recognition, Model-Based

Chew-Yean Yam, Mark S. Nixon
University of Southampton, Southampton, UK

Synonyms

Gait models for biometrics; Knowledge-based gait recognition

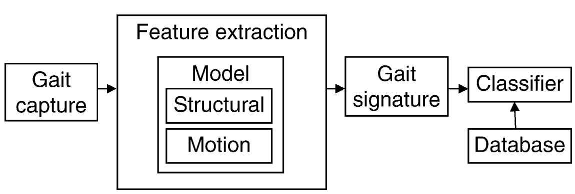
Definition

Model-based gait recognition relates to the identification using an underlying mathematical construct(s) representing the discriminatory gait characteristics (be they static or dynamic), with a set of parameters and a set of logical and quantitative relationships between them. These models are often simplified based on justifiable assumptions, e.g., a system may assume a pathologically normal gait. Such a system normally consists of gait capture, a model(s), a feature extraction scheme, a gait signature, and a classifier (Fig. 1). The model can be a 2- or 3-dimensional ► structural (or ► shape) ► model and/or ► motion model that lays the foundation for the extraction and tracking of a moving person. An alternative to a model-based approach is to analyze the motion of the human silhouette deriving recognition from the body's shape and motion. A gait signature that is unique to each person in the database is then derived from the extracted gait characteristics. In the classification stage, many pattern classification techniques can be used, such as the *k*-nearest neighbor approach.

The main advantages of the model-based approach are that it can reliably handle occlusion (especially self-occlusion), noise, scale and rotation well, as opposed to silhouette-based approaches.

Practical issues that challenge the model-based approach can be divided into two categories, which relate to the *system* and to the *person*. One of the systems-related challenges is viewpoint invariance, whilst person-related challenges include the effects of *physiological* changes (such as aging, the consistency of gait taken/enrolled at different times, whether our walking pattern changes over a longer period of time), *psychological* changes (mood), and *external factors* (load, footwear, and the physical environment).

The first model-based approach to gait biometrics was by Cunado et al. in 1997 [1, 2], featuring the ability to reliably accommodate self-occlusion and occlusion by other objects, noise, and low resolution. Also, most of the time, the parameters used within the model and their relationship to the gait are obvious, i.e., the mathematical construct may itself contain implicit/explicit meaning of the gait pattern characteristics. Though, it often suffers from high computational cost, this can be mitigated by optimization tools or increased computing power. Gait sequences are usually acquired when the subject is walking in a plane normal to the image capture device since the side view of a moving person reveals most information, though it is possible to use other views.



Gait Recognition, Model-Based. Figure 1 Components of a typical model-based gait recognition system.

Models

In a typical model-based approach, often, a ► structural model and a motion model are required to serve as the basis for tracking and feature (moving human) extraction. These models can be 2- or 3- dimensional, though most of the current approaches are 2-dimensional and have shown the capability to achieve promising recognition results on large databases (>100 subjects). A structural model describes the topology or the shape of human body parts such as head, torso, hip, thigh, knee, and ankle by measurements such as the length, width, and position. This model can be made up of primitive shapes (cylinders, cones, and blobs), stick figures, or arbitrary shapes describing the edge of these body parts. On the other hand, a motion model describes the kinematics or the dynamics of the motion of each body part. Kinematics generally describe how the subject changes position with time without considering the effect of masses and forces, whereas dynamics account for the forces that act upon these body masses and the resulting motion. When developing a motion model, the constraints of gait such as the dependency of neighboring joints and the limit of motion in terms of range and direction has to be understood.

Bobick et al. used a structural model to recover static body and stride parameters (Fig. 2a) determined by the body geometry and the gait of a person [3]. Lee et al. fit ellipses to seven regions representing the human body (Fig. 2b), then derived two types of features across time: mean and standard deviation, and magnitude and phase of these moment-based region features [4].

Cunado et al. proposed an early motion-model-based approach, based on the angular motion of the hip and thigh [1, 2], where the angular motion of the hip and the thigh is described by a Fourier series. For this method, a simple structural model was used and the angular rotation as defined in Fig. 3. Although the motion model is for one leg, assuming that gait is symmetrical, the other leg can be modeled similarly, with a phase lock of ½-period shift (Fig. 4).

Cunado et al. modeled the angular motion of the thigh by

$$\theta_T = a_0 + 2\sum_{1}^{N}[b_k \cos k\omega_0 t - c_k \sin k\omega_0 t],$$

where N is the number of harmonics, ω_0 is the fundamental frequency, and a_0 is the offset. In application, the frequency data was accumulated from a series of edge-detected versions of the image sequence of the walking subject. The gait signature was derived by the multiplication of the phase and magnitude component of the Fourier description.

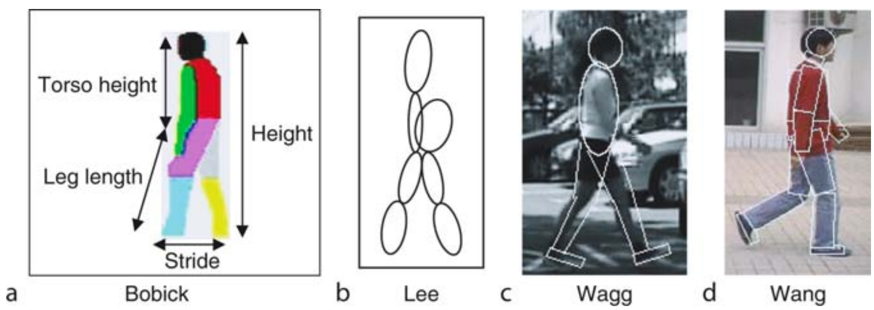
Later, Yam et al. [5] extended the approach to describe the hip, thigh, and knee angular motion of both walking and running gaits first by an empirical motion model, then by an analytical model motivated by coupled pendulum motion. Similarly, the gait signature is the phase-weighted magnitude of the Fourier description of both the thigh and knee rotation.

Bouchrika et al. [6] have proposed one of the latest motion-model-based gait feature extraction using a parametric form of elliptic Fourier descriptors to describe joint displacement.

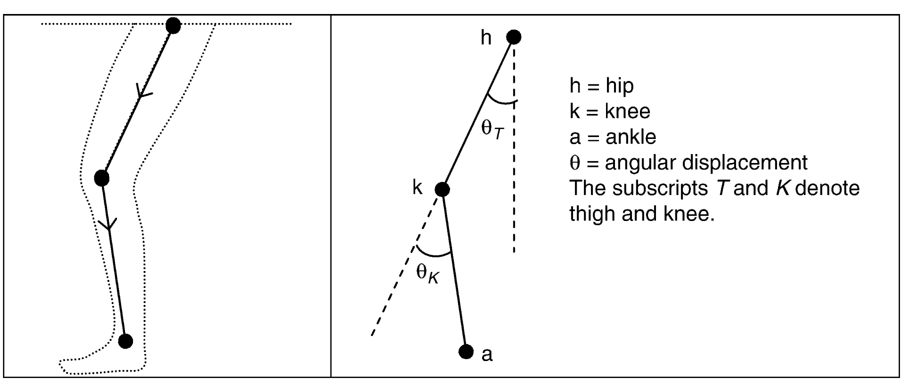
$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} a_0 \\ b_0 \end{bmatrix} + \begin{bmatrix} \cos(\alpha) - \sin(\alpha) \\ \sin(\alpha) \ \cos(\alpha) \end{bmatrix} \begin{bmatrix} X(t) * S_x \\ Y(t) * S_y \end{bmatrix},$$

where α is the angle, S_x and S_y are the scaling factors, and $X(t)$ and $Y(t)$ are Fourier summation. The joint trajectory is then fitted to the image sequence by optimizing a_0, b_0, α, S_x and S_y; the motion model fit is implemented by the Hough Transform.

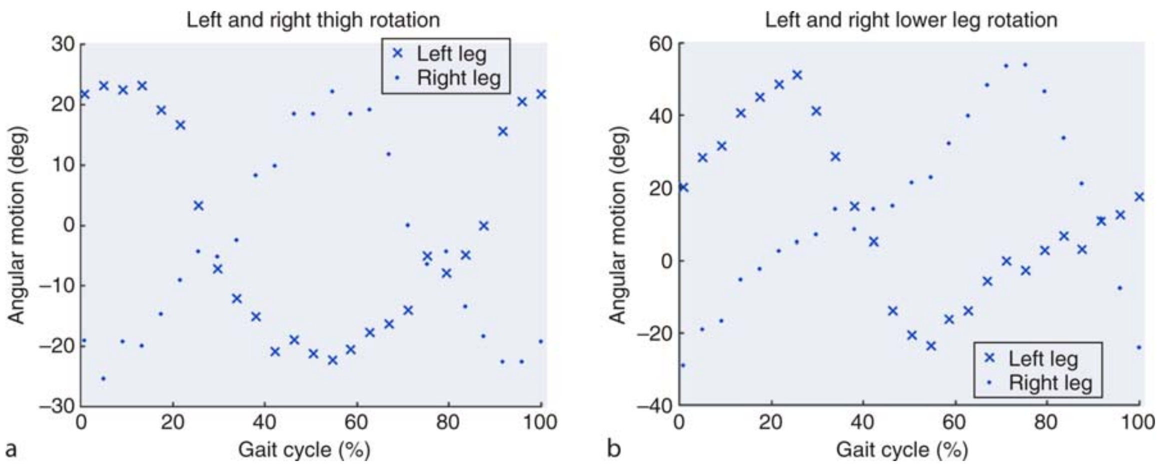
Wagg et al. (Fig. 2c) and Wang et al. (Fig. 2d) used a combination of both structural and motion models to



Gait Recognition, Model-Based. Figure 2 Example body parameters that are used in structural models. (**a**) Bobick (**b**) Lee (**c**) Wagg (**d**) Wang.



Gait Recognition, Model-Based. Figure 3 Structural model of a lower limb: upper and lower pendulum represents the thigh and the lower leg, respectively, connected at the knee joint.



Gait Recognition, Model-Based. Figure 4 Thigh and lower leg rotation of the left and right leg. (**a**) Left and right thigh rotation (**b**) Left and right lower leg rotation.

track and extract walking human figures [7, 8]. Wagg introduced a self-occlusion model whilst Wang used the conditional density propagation framework [9] to aid feature extraction.

Beyond the 2D models, Urtasun et al. developed a 3D gait motion model derived from a small group of subjects [10]. The joint motion is approximated by a weighted sum of the mean motion and the Eigenvectors of sample angular motion vectors. This approach also shows that it is capable of approximating running motion as well.

Feature Extraction

Feature extraction segments interesting body parts for a moving human, and extracts static and/or dynamic gait characteristics. The process normally involves model initialization, segmentation, and tracking (estimation) of the moving human from one image to the next. This is a significant step that extracts important spatial, temporal, or spatial-temporal signals from gait. Feature extraction can then be carried out in a concurrent [1, 2, 5, 8], or iterative/hierarchical [7] manner.

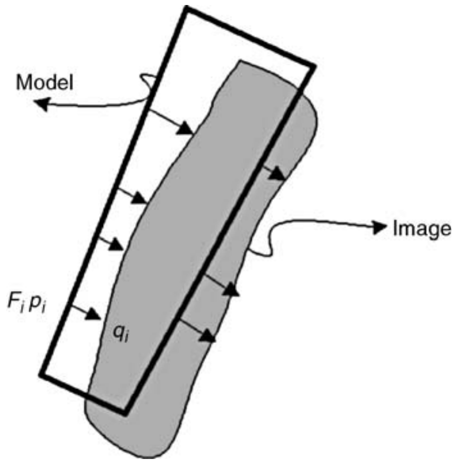
A conventional starting point of a gait cycle is the heel strike at the stance phase, although any other stage within a gait cycle can be used. Earlier techniques determine the gait cycle manually, later, many have employed automatic gait cycle detection. A gait cycle can be detected by simply identifying the stance phase; if using a bounding box method, the width of the box has the highest value during the stance phase. Other alternatives are counting the pixels of the human figure, using binary mask (Fig. 5) by approximating the outer region of the leg swing [7].

Quality of Feature Extraction

A good model configuration is defined as one that yields a high correlation between the model and the subject's image. Useful measures for computing model and image data correlation include *edge correspondence* and *region correspondence* [8]. Edge correspondence is a measure of how closely model edges coincide with image edges, whilst region correspondence is a measure of similarity between the image region enclosed by the model and that corresponding to the image of the subject. These two measures are used together. A high edge correspondence indicates that the model is closely aligned with image edges; however, it does not guarantee that the model matches the correct edges. If the initial model configuration is poor, or the subject is occluded, the match may be coincidental. For this reason, region correspondence is also required.

Another measure is a pose evaluation function (PEF) which combines the boundary (edge) matching error and the region matching error to achieve both accuracy and robustness. For each pixel, p_i, in the boundary of the projected human model, the corresponding pixel in the edge image along the gradient

Gait Recognition, Model-Based. Figure 5 Binary mask to detect gait cycle. The sum edge strength within the mask varies periodically during the subject's gait and the heel strike being the greatest.



Gait Recognition, Model-Based. Figure 6 Measuring the boundary matching error.

direction at p_i (Fig. 6) is searched. In other words, the pixel nearest to p_i and along that direction is desired. Given that q_i is the corresponding pixel and that $\boldsymbol{F}_i$ stands for the vector $\overline{p_i q_i}$, the matching error of pixel p_i to q_i can be measured as the norm $\|\mathbf{F}_i\|$. Then the average of the matching errors of all pixels in the boundary of the projected human model is defined as the boundary matching error

$$E_b = \frac{1}{N}\sum_{i=1}^{N} \|\mathbf{F}_i\|,$$

where N is the number of the pixels in the boundary.

In general, the boundary matching error measures the similarity between the human model and image data, but it is insufficient under certain circumstances, as illustrated in Fig. 7a, where a model part falls into the gap between two body parts in the edge image. Although it is obviously badly-fitted, the model part may have a small boundary matching error. To avoid such ambiguities, region information is further considered. Figure 7b illustrates the region matching. Here the region of the projected human model that is fitted into the image data is divided into two parts: P_1 is the model region overlapped with the image data and P_2 is the rest of the model region. Then the matching error with respect to the region information is defined by

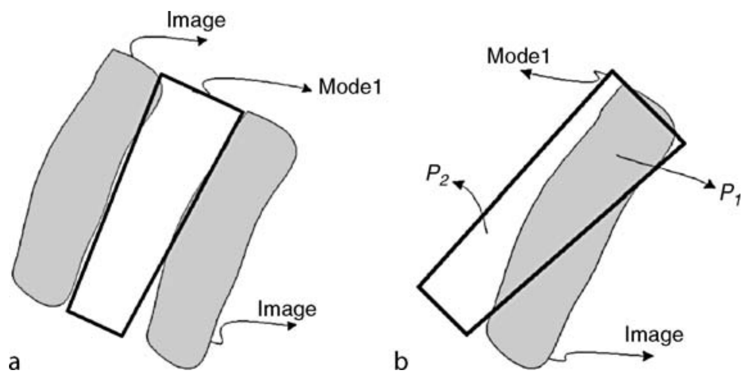
$$E_r = |P_2|/(|P_1| + |P_2|)$$

where $|P_i|$, $(i = 1, 2)$ is the area, i.e., the number of pixels in the corresponding region.

Recognition

A gait signature is a discriminatory feature vector that can distinguish individual. These signatures have invariant properties embedded in a person such as stride length, person's height/width, gait cycle and self-occlusion, and that related to the imaging system such as translation, rotation, scale, noise, and occlusion by other objects. These signatures can be of static [3], dynamic [2, 5] or a fusion of static and dynamic [7, 8] characteristics of gait or with other biometrics [11, 12]. The fusion can happen either at the feature extraction stage or at the classification stage. On the Southampton datasets of 115 subjects filmed indoors (in controlled conditions) and outdoors (with effects of shadows, background objects, and changing illumination) Wagg's approach achieved an overall CCR of 98.6% on the indoor data and 87.1% on the outdoor data.

In the case of 3D approach [10], experiments show that the first six coefficients of that motion model can



Gait Recognition, Model-Based. Figure 7 Illustrating the necessity of simultaneous boundary and region matching. (**a**) A typical ambiguity: a model part falls into the gap between two body parts (**b**) Measuring region matching error.

characterize 90% gait patterns of the database used. This resulted in a very compact gait signature, which requires only the first three coefficients to form separate clusters for each subject. It is interesting that this study found that the first few coefficients could represent physiological characteristics like weight, height, gender or age, while the remaining ones can be used to distinguish individual characteristics. Another interesting finding is that the nature of the gait signature for running derived from this 3D motion model is similar to that of Yam et al., that is, signature clusters are more dispersed within subject, and span more widely within the signature space, as compared to that of walking. Both studies were based on data collected by having subjects running on the treadmill.

Conclusions and Outlook

Using a model is an appealing way to handle known difficulty in subject acquisition and description for gait biometrics. There is a selection of models and approaches which can handle walking and running. Clearly, the use of a model introduces specificity into the feature extraction and description process, though this is generally at the cost of increased computation. Given their advantages, it is then likely that model-based approaches will continue to play a part in the evolution of systems which deploy gait as a biometric. Currently, practical advantages of three-dimensional (3D) approaches have yet to be explored and investigated. Given that human motion occurs in space and time, it is likely that much information is embedded within the 3D space. Further, 3D approaches may provide a more effective way to handle issues like occlusion, pose, and view point. Therefore, 3D model-based gait recognition may be a good way to move forward.

Related Entries

- ► Gait Recognition, Model-Based
- ► Human Detection and Tracking
- ► Markerless 3D Human Motion Capture from Images
- ► Multibiometrics
- ► Silhouette-Based Recognition

References

1. Cunado, D., Nixon, M.S., Carter, J.N.: Using gait as a biometric, via phase-weighted magnitude spectra. In: First International Conference on Audio- and Video-Based Biometric Person Authentication, pp. 95–102 Crans-Montana, Switzerland, (1997)
2. Cunado, D., Nixon, M.S., Carter, J.N.: Automatic extraction and description of human gait models for recognition purposes. Comput. Vis. Image Underst. **90**(1), 1–41 (2003)
3. Bobick, A.F., Johnson, A.Y.: Gait recognition using static, activity-specific parameters. In: Proceedings of IEEE Computer Vision and Pattern Recognition Conference (CVPR'01), pp. 423–430 Kauai, Hawaii, (2001)
4. Lee, L., Grimson, W.E.L.: Gait analysis for recognition and classification. In: Proceedings of Automatic Face and Gesture Recognition, pp. 148–155 Washington, DC, (2002)
5. Yam, C.Y., Nixon, M.S., Carter, J.N.: Automated person recognition by walking and running via model-based approaches. Pattern Recognit. **37**, 1057–1072 (2004)
6. Bouchrika, I., Nixon, M.S.: Model-based feature extraction for gait analysis and recognition. In: Mirage: Computer Vision/Computer Graphics Collaboration Techniques and Applications, INRIA Rocquencourt, France, March (2007)

7. Wagg, D.K., Nixon, M.S.: On automated model-based extraction and analysis of gait. In: Proceedings of Sixth International Conference on Automatic Face and Gesture Recognition, pp. 11–16 Seoul, South Korea, (2004)
8. Wang, L., Tan, T., Ning, H., Hu, W.: Fusion of static and dynamic body biometrics for gait recognition. IEEE Trans. Circuits Syst. Video Technol. (Special Issue on Image- and Video-Based Biometrics) **14**(2), 149–158 (2004)
9. Isard, M., Blake, A.: CONDENSATION – conditional density propagation for visual tracking. Int. J. Comput. Vis. **29**(1), 5–28 (1998)
10. Urtasun, R., Fua, P.: 3D tracking for gait characterization and recognition. In: Proceedings of Sixth IEEE International Conference on Automatic Face and Gesture recognition, Seoul, South Korea, pp. 17–22 (2004)
11. Kale, A., RoyChowdhury, A.K., Chellappa, R.: Fusion of gait and face for human identification. In: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Montreal, Canada, May 2004, vol. 5, pp. 901–904 (2004)
12. Shakhnarovich, G., Lee, L., Darrell, T.: Integrated face and gait recognition from multiple views. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 439–446 Hawaii, (2001)

Gait Recognition, Motion Analysis for

Ahmed Elgammal
Department of Computer Science, Rutgers University, Piscataway, NJ, USA

Synonyms

Appearance-based gait analysis; Silhouette analysis for gait recognition

Definition

The appearance of gait in an image sequence is a spatiotemporal process that characterizes the walker. The spatiotemporal characteristics of gait contain rich perceptual information about the body configuration, the person's gender, the person's identity, and even the emotional states of the person. Motion analysis for gait recognition is a computer vision task that aims to capture discriminative spatiotemporal features (signature) from image sequences in order to achieve human identification. Such a signature ought to be invariant to the presence of various viewing conditions, such as viewpoint, people clothing, etc. In contrast to Model-based gait analysis systems, which is another article, the goal here is to capture gait characteristics without fitting a body model or locating the body limbs, rather by analyzing the feature distribution over the space and time extent of the motion.

Human Gait as a Biometric

Human gait is a valuable biometric cue that has the potential to be used for human identification similar to other biometric features, such as faces and fingerprints. Gait has significant advantages compared to other biometric features since it is easily observable in an unintrusive way, it does not require collaborative subjects, and it is difficult to disguise [1]. Therefore, using gait as a biometric feature has a great potential for human identification in public places for surveillance and for security. A fundamental challenge in gait recognition is to develop robust algorithms that can extract visual gait features invariant to the presence of various conditions that affect people's appearance, as well as conditions that affect people's gait. That includes, viewpoint, clothing, walking surface, shoe type, object carried, etc. [2].

Johansson's seminal psychophysical experiments [3] showed that humans can recognize biological motion, such as gait, from Moving Light Displays (MLD). Cutting and Kozlowski [4] showed that humans can also identify friends from their gait using MLD. Motivated by these results, many researchers in different disciplines, have shown that the spatiotemporal characteristics of gait contain rich perceptual information about the body configuration, the person's gender, the person's identity, and even the emotional states of the person. That motivated extensive recent computer vision research on extracting features from gait.

Vision-based human motion tracking and analysis systems have promising potentials for many applications, such as visual surveillance in public area, activity recognition, sport analysis, video retrieval, and human–computer interaction. Extensive research has been done in this area in the last two decades with lots of promising results. For excellent literature surveys in the subject, the reader can refer to [5, 6]. The human body is an articulated object with a large number of degrees of freedom. This fact makes the problems of tracking the body configuration and extracting biometrics very challenging. Besides the articulation nature of the body, the variability in people's appearance adds to the problems. Human gait is a special case of the general problem of human motion analysis, and to some extent, is easier. This is because of the physical constraints on such a motion as well as the periodic nature of it.

The appearance of gait in an image sequence is a spatiotemporal process that characterizes the walker. Gait recognition algorithms, generally, aim to capture discriminative spatiotemporal features (signature) from image sequences in order to achieve human identification. Gait analysis approaches can be categorized according to the way the gait features are extracted for classification. There are two broad categories of approaches: model-based approaches and appearance-based approaches. Model-based approaches, e.g., [1], fit 3D body models or intermediate body representations to body limbs in order to extract proper features (parameters) that describe the dynamics of the gait (see the related entry on "Model-based Gait Recognition" for details). Model-based approaches typically require a large number of pixels on the tracked target to fit their model, i.e., high resolution zoomed-in images are required on the tracked person. In contrast, appearance-based approaches aim to capture a spatiotemporal gait characteristic directly from input sequences without fitting a body model. The appearance-based approaches are mainly motivated by the psychophysical experiments, mentioned earlier, e.g., [3, 4], which showed that spatiotemporal patterns such as Moving Light Displays could capture important gait information without the need of finding limbs. Appearance-based approaches do not require high resolution on subjects, which makes them more applicable in outdoor surveillance applications where the subjects can be at a large distance from the camera.

Characteristics and Challenges of Gait Motion

Gait is a 3D articulated periodic motion that is projected into 2D image sequences. Therefore, the appearance of a gait motion in an image sequence is a spatiotemporal pattern, i.e., a spatial distribution of features that changes over time. Researchers have developed several algorithms for capturing gait signature from such spatiotemporal patterns by looking at the space-time volume of features. The observed shapes of the human body, in terms of the occluding contours of the body (silhouettes), are examples of such spatiotemporal patterns, which contain rich perceptual information about the body configuration, the motion performed, the person's gender, the person's identity, and even the emotional states of the person. Objects occluding contours, in general, have a great role in perception [7] and have been traditionally used in computational vision, besides other appearance cues, to determine object category and pose.

The objective of any gait tracking and analysis system is to track the global deformations of contours over time and to capture invariant gait signature from such contours. There are several challenges to achieve this goal. An observed person's contour in a given image is a function of many factors, such as the person's body build (tall, short, big, small, etc.), the body configuration, the person's clothing and the viewpoint. Such factors can be relevant or irrelevant depending on the application. Modeling these sources of variabilities is essential to achieve successful trackers and to extract gait biometric features. Modeling the human body dynamic shape space is hard, since both the dynamics of shape (different postures) and the static variability in different people's shapes have to be considered. Such shape space lies on a nonlinear ► manifold.

Figure 1 shows an example of a walking cycle from a side view where each row shows half a walking cycle. The shapes during a gait cycle temporally undergo deformations and self-occlusion. The viewpoint from which the gait is captured imposes self-similarity on the observed shapes over time. This similarity can be noticed by comparing the corresponding shapes at the two rows in Fig. 1. This right part of the figure shows the correlation between these shapes. The similarity between the corresponding shapes in the two half cycles is exhibited by the dark diagonally parallel bands in the correlation plot. The similarity in the observed shapes indicates a nonlinear relation between the observed gait and the kinematics of the gait. This can be noticed by closely inspecting the two shapes in the middle of the two rows in Fig. 1. These two shapes correspond to the farthest points in the walking cycle kinematically (the top has the right leg in front while the bottom has the left leg in front). In the Euclidean visual input space (observed shapes) these two points are very close to each other as can be noticed from the distance plot on the right of Fig. 1. This nonlinear relation between the observed shapes

Gait Recognition, Motion Analysis for. Figure 1 Twenty sample frames from a walking cycle from a side view. Each row represents half a cycle. Notice the similarity between the two half cycles. The right part shows the similarity plot: each row and column of the plot corresponds to one sample. Darker means closer distance and brighter means larger distances. The two dark lines parallel to the diagonal show the similarity between the two half cycles.

and the kinematics poses a problem to gait tracking and analysis systems. However, such similarity can be useful in extracting gait features. For example, the temporal self-similarity characteristic has been exploited in the work of BenAbdelkader et al. [8] for gait recognition.

Extracting Gait Signature from Motion

There have been extensive research on appearance-based extraction of gait signatures. Typical preprocessing steps for gait analysis include detecting and tracking the human subject in order to locate a bounding box containing the motion and/or extracting the body silhouette (see the related entry on human detection and tracking).

One of the early papers on gait analysis using spatiotemporal features is the work of Niyogi and Adelson [9] where a spatiotemporal pattern (corresponding to leg motion) was used to detect gait motion in an image sequence represented as an XYT volume. Gait was then parameterized with four angles for recognition. Murase and Sakai [10] used a parametric eigenspace representation to represent a moving object using Principle Component Analysis (PCA). In their work, the extracted silhouettes were projected into an eigenspace where a walking cycle forms a closed trajectory in that space. Spatiotemporal correlations between a given trajectory and a database of trajectories were used to perform the recognition. Huang et al. [11] extended the method using Canonical space transformation (CST) based on Canonical Anaylsis (CA), with eigenspace transformation for feature extraction.

Little and Boyd [12] exploited the spatial distribution of optical flow to extract spatiotemporal features. From dense optical flow, they extracted scale-independent features capturing the spatial distribution of the flow using moments. This facilitates capturing the spatial layout of the motion, or as they call it "the shape of the motion." Periodicity analysis was then done on these features to capture gait signatures for recognition. BenAbdelkader et al. [8] used image self-similarity plots (similar to Fig. 1) to capture the spatiotemporal characteristics of gait. Given bounding boxes around a tracked subject, correlation is used to measure self-similarity between different time frames in the form of similarity plots. PCA analysis was used to reduce the dimensionality of such similarity plots for recognition. Hayfron-Acquah et al. [13] used spatial symmetry information to capture gait characteristics from silhouettes. Given a walking cycle, a symmetry operator was used to extract a symmetry map for each silhouette instance in the cycle. Fourier transform was used to extract descriptors from such symmetry maps for recognition.

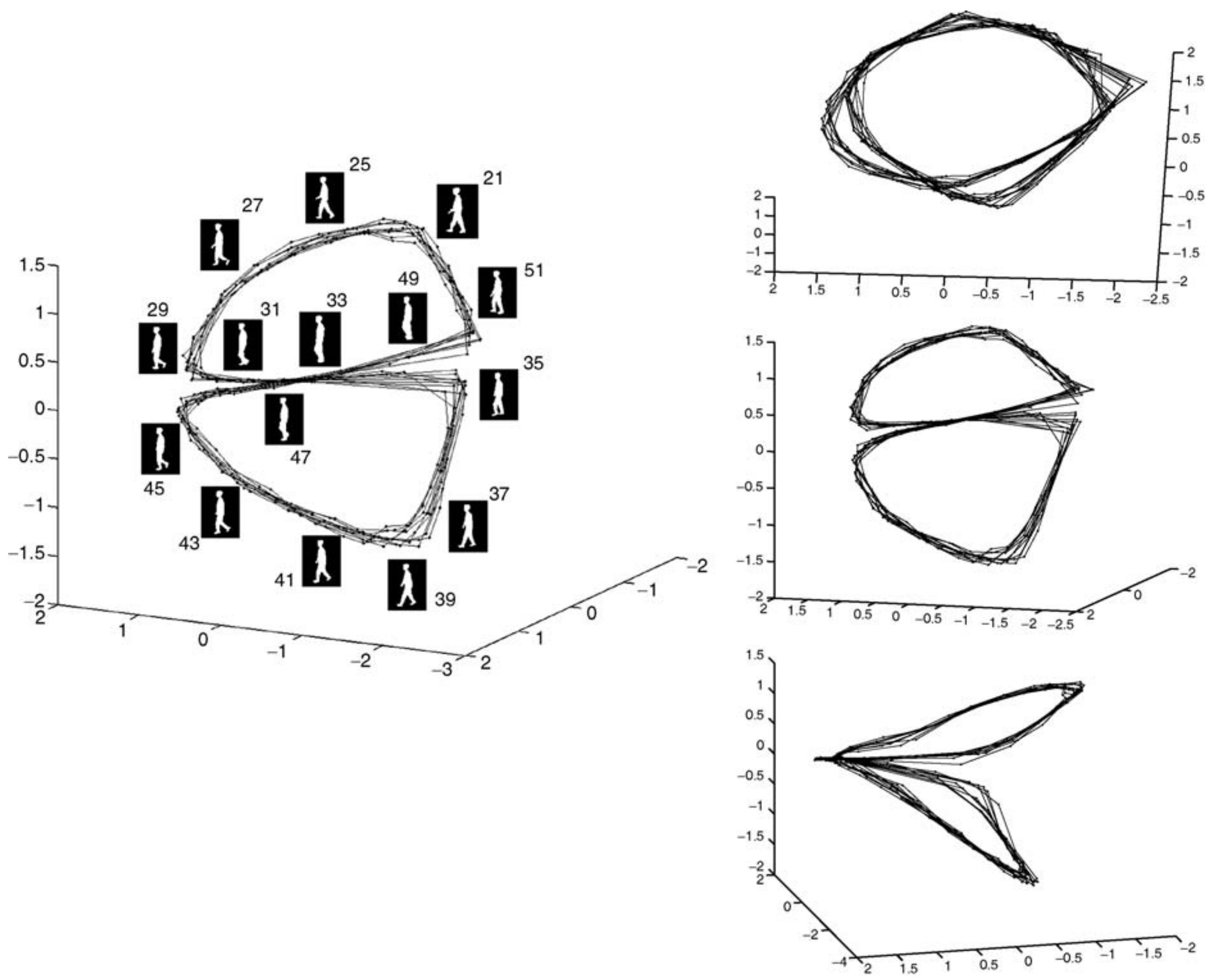
Since gait is a temporal sequence, researchers have investigated the use of Hidden Markov Models (HMM) to represent and capture gait motion characteristics. HMMs have been successfully used in many speech recognition systems, as well as gesture recognition applications. Typically a left-right HMM with a small number of states (three to five) is sufficient to model the gait of each subject in the database, where the HMMs are trained from features extracted from silhouettes. In [14], HMM was used to capture gait dynamics from quantized Hu moments of silhouettes. HMM was also used in [15] with features representing silhouette width distribution.

Lee and Elgammal [16] used bilinear and multilinear models to factorize the spatiotemporal gait process into gait style and gait content factors. A nonlinear mapping was learned from a unit circle (representing a gait cycle) to the silhouettes' shape space. The unit circle represents a unified model for the gait manifold of different people, therefore, any spatiotemporal characteristics of the gait of a specific person should exist on the mapping space. Bilinear and multilinear models were used to factorize such mapping to extract gait signatures.

Manifold-based Representation for Gait Analysis

Despite the high dimensionality of the human body configuration space, any body motion is constrained by the physical dynamics, body constraints, and the motion type. Therefore, many human activities lie intrinsically on low dimensional manifolds. This is true for the body kinematics, as well as for the observed motion through image sequences. For certain classes of motion like gait, facial expression, and simple gestures, considering a single person and factoring out other sources of variability, the deformations will lie on a one-dimensional manifold. Recently many researchers have developed techniques and representations for gait analysis that exploit such manifold structure, whether in the visual space or in the kinematic space, e.g. [17, 18]. Modeling the gait manifold was earlier used for gait recognition in [10].

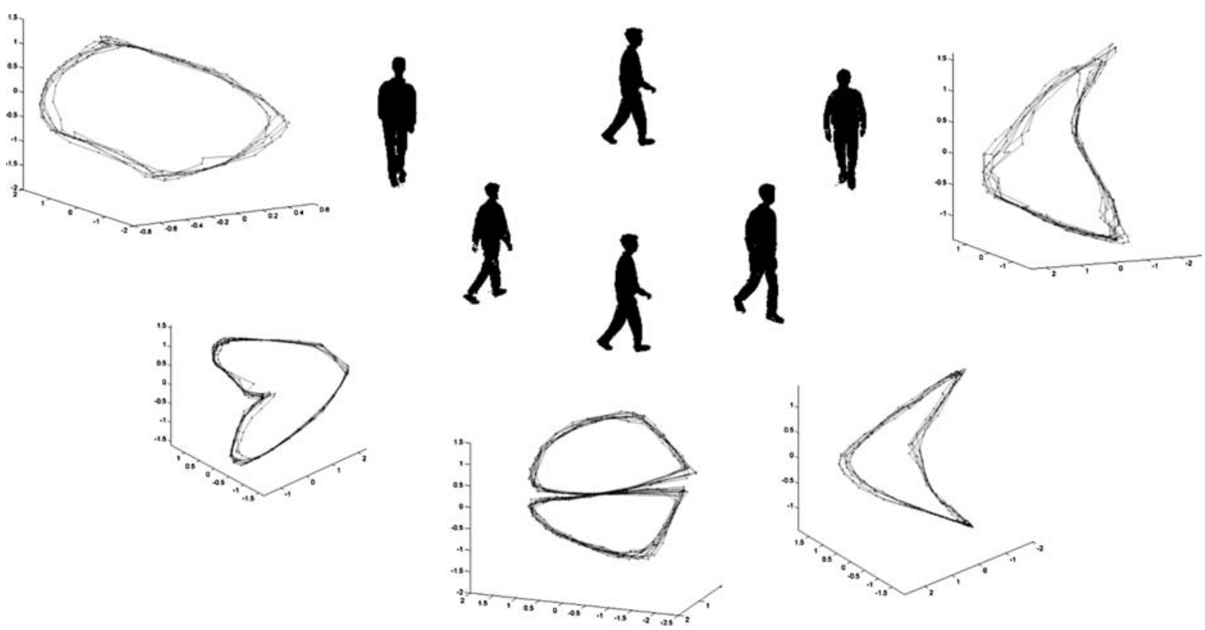
Intuitively, the gait is a one-dimensional closed manifold that is embedded in a high dimensional visual space. Such a manifold can twist and self-intersect in such high dimensional visual space. This can be



Gait Recognition, Motion Analysis for. Figure 2 Embedded gait manifold for a side view of the walker. *Left*: sample frames from a walking cycle along the manifold with the frame numbers shown to indicate the order. Ten walking cycles are shown. *Right*: three different views of the manifold. © IEEE.

noticed by considering the human silhouette through the walking cycle, (as shown in Fig. 1) as points in a high dimensional visual input space. Given the spatial and the temporal constraints, it is expected that these points will lay on a closed trajectory. In order to achieve a low dimensional embedding of the gait manifold (▶ manifold embedding), dimensionality reduction techniques can be used. Linear dimensionality reduction can be used to achieve an embedding, as in [10]. However, in such a case the two half cycles would be collapsed to each other because of the similarity in the shape space. Nonlinear dimensionality reduction techniques such as LLE [19], Isomap [20], GPLVM [21], and others can successfully embed the gait ▶ manifold in a way that separates the two half cycles. As a result of nonlinear dimensionality reduction, an embedding (and a visualization) of the gait manifold can be obtained in a low-dimensional Euclidean space [17]. Figure 2 illustrates an example embedded manifold for a side view of the walker. The data used are from the CMU Mobo gait data set which contains 25 people from six different view points. Data sets of walking people from multiple views are used in this experiment. Each data set consists of 300 frames and each containing about 8–11 walking cycles of the same person from a certain view points. The walkers were using treadmill which might result in different dynamics from the natural walking. Figure 3 illustrates the embedded manifolds for five different view points of the walker. For a given view point, the walking cycle evolves along a closed curve in the embedded space, i.e., only one degree of freedom controls the walking cycle, which corresponds to the constrained body pose as a function of the time. Such a conclusion conforms to the intuition that the gait manifold is one-dimensional.

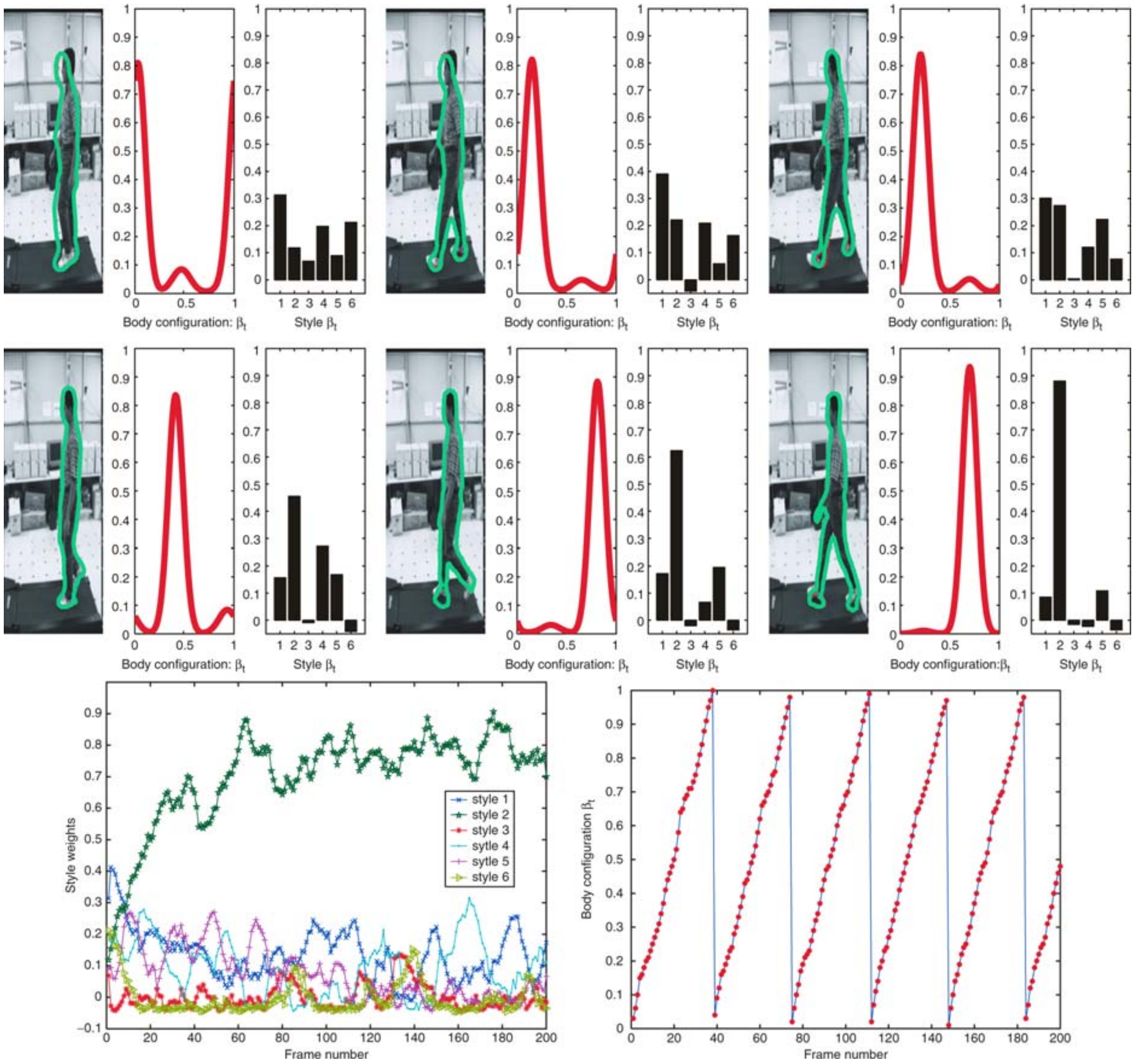
As can be noticed in Fig. 3, The manifold twists in the embedding space given the different viewpoints, which impose different self occlusions. The least twisted manifold is the manifold for the back view as this is the least self occluding view (left most manifold in Fig. 3. In this case the manifold can be embedded in a two dimensional space. For other views, the curve starts to twist to be a three-dimensional space curve. This is primarily because of the similarity imposed by the view point which attracts far away points on the manifold closer. The ultimate twist happens in the side view manifold where the curve twists to get the shape of the numeral 8 where each cycle of the eight (half eight) lies in a different plane. Each half of the “eight” figure corresponds to half a walking cycle. The cross point represents the body pose where it is



Gait Recognition, Motion Analysis for. Figure 3 Embedded manifolds for five different views of the walkers. Frontal view manifold is the right most one and back view manifold is the leftmost one. The view of the manifold that best illustrates its shape in the 3D embedding space is visualized. © IEEE.

totally ambiguous from the side view to determine from the shape of the contour which leg is in front, as can be noticed in Fig. 2. Therefore, in a side view, a three-dimensional embedding space is the least that can be used to discriminate the different body poses. Embedding a side view cycle in a two-dimensional embedding space results in an embedding similar to that shown in top right of Fig. 2 where the two half cycles lie over each other. Interestingly, despite that the side view is the most problematic view of the gait, most gait recognition systems seem to favor such view for recognition! Different people are expected to have different manifolds. However, such manifolds are all topologically equivalent.

The example embeddings shown here are for silhouette data, i.e., the *visual manifold* of the gait is



Gait Recognition, Motion Analysis for. **Figure 4** Adaptive Contour Tracking of Gait: (**a**) tracking through sample frames. (**b**) adapting to the target style. (**c**) the tracked body configuration showing a constant speed dynamic system. From [22].

embedded. Similar embedding can be obtained for kinematic data, in such a case the *kinematic manifold* of the gait is embedded. In such a case PCA would be sufficient to achieve an embedding. The importance of such embedded representations is that they provide a low dimensional representation for tracking the gait motion. Only a one-dimensional parameter is needed to control and track the gait motion. This leads to a simple constant speed dynamic model for the gait. Figure 4 shows an example of gait contour tracking system [22] that uses an embedded representation of the gait manifold. As a result, a constant speed linear dynamics is achieved (Fig. 4b). The tracker can also adapt to the tracked person shape style and identify that style from a database of styles (Fig. 4c).

Explicit manifold representation for gait is not only useful for tracking and pose estimation, but also can be used in gait recognition systems. Different people are expected to have different manifolds for the appearance of their gait. However, such manifolds are all topologically equivalent to a unit circle. A person's gait manifold can be thought of as a twisted circle in the input space. The spatiotemporal process of gait is captured in the twist of a given person's manifold. Therefore, a person's gait signature can be captured by modeling how a unit circle (an ideal manifold) can deform to fit that person's gait manifold. This can be achieved by fitting a nonlinear warping function between a unit circle and a given person's silhouette sequence. In [23] this approach was used to capture gait signatures by factorizing the warping functions' coefficient space to obtain a low-dimensional gait signature space for recognition.

Summary

Appearance-based analysis of gait is motivated and justified by psychophysical experiments. Appearance-based approaches for gait recognition aim to extract a gait signature from the spatial and temporal distribution of the features on a tracked subject without the need to fit a body model or to locate limbs. Such approaches have proved very successful in gait recognition and are applicable in scenarios where the gait biometric features can only be extracted from a distance. There are many limitations to the current gait recognition systems including achieving invariant to viewing conditions, such as viewpoint invariant. Recent progress in manifold-based representation of gait, as well as factorized models, such as multilinear tensor models provides potential solutions to such problems.

Related Entries

► Gait Recognition, Model-Based
► Human detection and tracking

References

1. Cunado D., Nixon M.S., Carter J.: Automatic extraction and description of human gait models for recognition purposes. Compt. Vision Image Understand. **90**, 1–41 (2003)
2. Sarkar S., Phillips P.J., Liu Z., Vega I.R., Grother P., Bowyer K.W.: The humanid gait challenge problem: Data sets, performance, and analysis. IEEE Trans. PAMI **27**(2), 162–177 (2005)
3. Johansson G.: Visual motion perception. Sci. Am. **232**, 76–88 (1975)
4. Cutting J.E., Kozlowski L.T.: Recognizing friends by their walk: gait perception without familiarity cues. Bull. Psychonomic Soc. **9**(5), 353–356 (1977)
5. Gavrila D.M.: The visual analysis of human movement: a survey. Comput. Vis. Image Understand. **73**(1), 82–98 (1999). DOI http://dx.doi.org/10.1006/cviu.1998.0716
6. Moeslund T.B., Hilton A., Krüger V.: A survey of advances in vision-based human motion capture and analysis. Comput. Vis. Image Underst. **104**(2), 90 –126 (2006). DOI http://dx.doi.org/10.1016/j.cviu.2006.08.002
7. Palmer, S.E.: Vision Science, Photons to Phenomenology. MIT, Cambridge, MA (1999)
8. BenAdbelkader, C., Cutler, R., Davis, L.: Motion-based recognition of people using image self-similarity. In: Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition, pp. 254–259 (2002)
9. Niyogi, S., Adelson, E.: Analyzing and recognition walking figures in xyt. In: Proceedings of IEEE CVPR, pp. 469–474 (1994)
10. Murase H., Sakai R.: Moving object recognition in eigenspace representation: gait analysisand lip reading. Pattern Recog. Lett. **17**, 155–162 (1996)
11. Huang P., Haris C., Nixon, M.: Recogising humans by gait via parametric canonical space. Artif. Intell. Eng. **13**, 359–366 (1999)
12. Little, J.J., Boyd, J.E.: Recognizing people by their gait: The shape of motion. Videre: J. Comput. Vision Res. **1**(2) (1998)
13. Hayfron-Aquah J.B., Nixon M.S., Carter J.N.: Automatic gait recognition by symmetry analysis. Pattern Recogn. Lett. **24**, 2175–2183 (2003)
14. He, Q., Debrunner, C.: Individual recognition from periodic activity using hidden markov models. In Proceedings of the Workshop on Human Motion (Humo) (Dec, 2007). HUMO. IEEE Computer Society, Washington, DC
15. Kale, A., Sundaresan, A., Rajagopalan, A.N., Cuntoor, N.P., Roy-Chowdhury, A.K., Kruger, V., Chellappa, R.: Identification

G

of human using gait. IEEE Trans. Image Process. **13**(9), 1163–1173 (2004)
16. Lee, C.S., Elgammal, A.: Gait style and gait content: Bilinear model for gait recogntion using gait re-sampling. In: Proceedings of FGR, pp. 147–152 (2004)
17. Elgammal, A., Lee, C.S.: Inferring 3d body pose from silhouettes using activity manifold learning. In: Proceeedings of CVPR, vol. 2, pp. 681–688 (2004)
18. Urtasun, R., Fleet, D.J., Hertzmann, A., Fua, P.: Priors for people tracking from small training sets. In: ICCV, Beijing, China pp. 403–410 (2005)
19. Roweis S., Saul L.: Nonlinear dimensionality reduction by locally linar embedding. Science **290**(5500), 2323–2326 (2000)
20. Tenenbaum, J.: Mapping a manifold of perceptual observations. In: Proceedings of Advances in Neural Information Processing (NIPS), vol. 10, pp. 682–688 (1998)
21. Lawrence, N.D.: Gaussian process models for visualisation of high dimensional data. In: Proceedings of NIPS. BMVC, Oxford, UK (2004)
22. Lee, C.S., Elgammal, A.: Style adaptive bayesian tracking using explicit manifold learning. In: Proceedings of British Machine Vision Conference (2005)
23. Lee, C.S., Elgammal, A.: Audio- and Video-based Biometric Person Authentication Conference AVBPA. Terrytown, NY, USA (2005)

Gait Recognition, Silhouette-Based

Jeffrey E. Boyd[1], James J. Little[2]
[1]University of Calgary, Calgary, AB, Canada
[2]University of British Columbia, Vancouver, BC, Canada

Definition

Silhouette-based gait recognition is the analysis of walking human figures for the purpose of biometric recognition. Gait biometrics offers the advantage of covertness; acquisition is possible without the awareness or cooperation of the subject. The analysis may apply to a single static image, or to a temporal sequence of images, i.e., video.

Introduction

The phenomenon of gait is the "coordinated, cyclic combination of movements that result in human locomotion" [1]. Gait is necessary for human mobility and is therefore ubiquitous and easy to observe.

The common experience of recognizing a friend from a distance by the way they walk has inspired the use of gait as a biometric feature. In fact, Cutting and Kozlowski [2], using ▶ moving light displays to isolate the motion stimulus, demonstrated that humans can indeed identify familiar people from gait. In their experiments, seven subjects identified the gaits of a subset of six subjects correctly at a rate of 38%. While this rate is less than adequate for biometrics, it is significantly better than random (17% in for their sample size), and validates the human source of inspiration.

To convert a gait into a feature vector suitable for biometrics, one can *characterize the motion* in the gait, e.g., by analyzing joint angles and limb trajectories, or by measuring the overall pattern of motion. Alternatively, one can *measure critical body dimensions* such as height or limb lengths. In the later approach, biometric features can be measured statically, but the motion in the gait provides a convenient mechanism to reveal joint positions, and consequently, limb lengths. McGeer's work on *passive dynamic walkers* [3, 4] reveals the extent to which gait motion relates to body mass and limb lengths: in the passive dynamic model of a human gait, the motion is a stable limit cycle that is a direct result of body mass and limb length. Factors not accounted for in McGeer's original model are muscle activation (gravity powers a passive dynamic walker), walking surface, injury, and fatigue. Intuitively, the motion in a gait is a reflection of the mass and skeletal dimensions of the walker. McGeer's passive dynamic model leads to more sophisticated models that account for some of these other factors. For example, see the work of Kuo [5, 6].

Confounding factors in gait biometrics include clothing and footwear. Clothing can change the observed pattern of motion and make it difficult to accurately locate joint positions. The effect of footwear is more complex. Some variation in footwear causes changes in muscle activation, but causes no outwardly visible change in the pattern of motion [7], whereas other footwear changes will alter gait.

▶ Silhouette-based gait recognition extracts the form of a walking subject, and then computes a feature vector that describes either the pattern of motion in the gait, or the physical dimensions of the subject. A classifier then matches the feature vector against previously acquired examples for identification or verification.

Silhouettes

Definitions of silhouette are often ambiguous: some definitions refer to the region covered by a figure, whereas other definitions refer to the boundary between a figure and its background. In the context of silhouette-based gait biometrics, we assume that the silhouette refers to the region, rather than the border. Nevertheless, there are related examples that use the boundary, e.g., see Baumberg and Hogg [8].

To form a silhouette of a walking figure requires the ▶ segmentation of image pixels into foreground (the moving figure) and background (everything else) sets of pixels. The silhouette is the set of foreground pixels. The easiest way to acquire a reliable silhouette is *chroma-keying* [9], which relies on color disparities between a backdrop and the foreground subject. The background color (usually green or blue), is chosen to make the color discrimination robust. Figure 2d shows an example of chroma-keying in gait analysis. The unusual color of the backdrop makes the subject aware that they are under surveillance, negating the covertness of gait biometrics.

▶ Background subtraction obviates the need for a colored backdrop by measuring the naturally occurring scene behind the subject. This entails estimating the statistical properties (usually in the luminance and color) of every pixel over one or more frames of video. By comparing the background estimate with subsequent frames of video, one can classify foreground pixels as those that do not match the background. The classifier can be as simple as thresholding of the absolute difference between the background and video frames. In most cases, the background estimation and subtraction are merged into an online system that continuously computes pixel differences and then updates the background for each frame of video. Background subtraction requires that the background and camera be stationary. Stauffer and Grimson [10] describe a widely used background subtraction method that uses a multimodal estimate of background statistics to produce reliable silhouettes of moving objects. Their method is robust in the presence of some background motions (e.g., rustling leaves or swaying tree branches).

The projection of motion in a scene onto a camera image plane is called a motion field. When a human figure is walking, segmenting moving from slow or stationary pixels in the motion field will extract a silhouette of the figure. Additionally, a motion field provides richer information than a simple silhouette because it indicates not only where the subject is moving, but also how fast the various body parts are moving. In general, it is not possible to measure a motion field, but one can measure ▶ optical flow, an approximation to the motion field that is sufficient for biometric gait recognition. If one imagines the luminance of pixels to be a fluid that can flow around an image, the optical flow estimates the movement of that fluid. It is, in part, related to the motion field, but is not necessarily equal to the motion field in all cases. Barron et al. [11] provide a comparative survey of some well-known optical flow algorithms. For example, see Fig. 2a.

Most silhouette-based biometric gait analysis focuses on a view of the subject orthogonal to the sagittal plane of the subject, i.e., the subject walks across the field of view rather than toward or away from the camera. We believe that this preference exists because front or rear views of the subject show mostly side-to-side motion and do not reveal either joint location or the complex patterns of limb motion.

Marker-based motion capture, e.g., Johansson's moving light displays, offers a counterpoint to silhouettes that are less practical for biometrics, but are useful for gaining insight into the perceptual issues surrounding gait [12, 13].

Duration of Observation

In general, it is desirable to observe the gait as long as possible. One way to extend the duration of an observation indefinitely is to have a subject walk on a treadmill in front of a stationary camera, e.g., see Fig. 2b. However, this requires the cooperation and awareness of the subject.

Alternatively, allowing the camera to pan with the motion of the subject can extend the observation time without the subject walking on a special apparatus. However, when the camera moves, the images acquired contain both the movement of the subject, and the background. The changing background makes accurate background subtraction difficult.

Using a static camera simplifies both the apparatus and the processing to extract the silhouette, but the duration of observation is limited by the time it takes the subject to cross the field of view of the camera. The actual duration will vary with the angular width of the field of view, the distance between the subject and the camera, and the speed of the subject. The practical limit on distance to subject depends on the resolution of the camera. Higher resolutions allow the subject to be further away while maintaining enough pixel coverage to measure biometric feature vectors accurately. In examples reported in the literature that use a static cameras and subjects walking on the ground, the typical duration of observation is approximately three to six strides.

Periodicity and Synchronization

Gait is a periodic phenomenon, so the silhouette of a walker varies with position in the gait cycle. Consequently, it is necessary to synchronize measurements of the silhouette to positions in the gait cycle. In turn, this requires measurement of the frequency of the gait and establishment of a phase reference within the gait cycle.

The method used to perform the synchronization depends on the particular measurements acquired and can serve to differentiate gait analysis methods. For example, Little and Boyd [14] measure the frequency from the oscillations of the centroid of the figure. To establish a phase reference, they use the phase of an oscillating measurement. In methods that measure height, e.g., Ben-Abdelkader et al. [15], the frequency of oscillations of the figure height gives the frequency of the gait. Positions of maxima in the height correspond to the positions in the gait where the swinging leg is vertical, thus defining a phase reference.

Conversion of Silhouettes to Features

A necessary step in silhouette-based gait recognition is conversion of a temporal sequence of silhouettes into a *gait signature*, i.e., a feature vector suitable for classification. One approach is to extract features that characterize the silhouette shapes and their variation over time, as illustrated schematically in Fig. 1a.

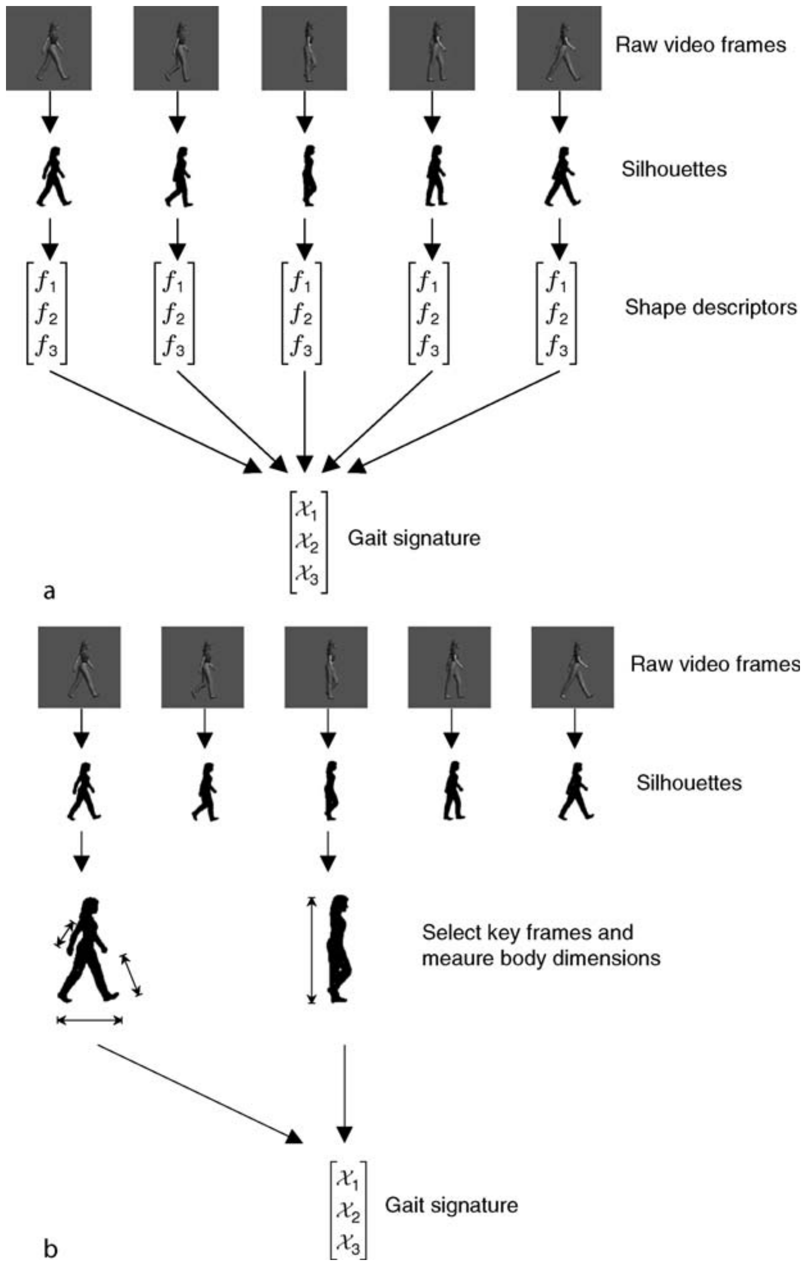
As an example, Little and Boyd [14] use geometric moments to describe a silhouette within a single frame of video. The moments include geometric centers, i.e., the average position of pixels in the silhouette, sometimes called the center of mass. Weighting the pixel positions by corresponding optical flow values gives geometric moments sensitive to rapid limb movement. Little and Boyd also use eccentricity [16], based on higher-order geometrical moments. A further step is necessary to combine the shape description for silhouettes in individual frames to a feature vector representative of the entire gait. Cyclic oscillations in the silhouette shape moments result naturally from a gait, so Little and Boyd exploit this to collect the individual shape descriptions into a single feature vector of the relative phases of the moment oscillations. Shutler and Nixon [17] describe a variation on this approach that uses Zernike moments to represent an accumulated shape over the duration of a gait cycle.

Ben-Abdelkader et al. [18] also exploit the periodic nature of a gait to form feature vectors. Periodicity and symmetry in a gait mean that similar shapes occur throughout the cycle of a gait. A feature vector built from measures of the silhouette self-similarity over period forms the basis for gait recognition. Periodicity in the self-similarity measures establishes the frequency of the gait. Hayfron-Acquah et al. [19] characterize the silhouette shape in a single frame by measuring symmetries in the outline of the silhouette to produce a *symmetry map*. The average of these symmetry maps over a gait cycle gives the gait signature used for recognition. Boyd [20] uses an array of phase-locked loops to measure the frequency, amplitude, and phase of pixel intensity oscillations due to a gait. The amplitudes and relative phases form a vector of complex phasors that acts as gait signature for recognition.

Rather than relying on the connection between gait and body structure to form a gait signature, one can use feature vectors that relate directly to body dimensions as shown in Fig. 1b. For example, Bobick and Johnson [21] measure stride and torso lengths, and Ben-Abdelkader et al. [15] measure height and stride characteristics. Collins et al. [22] identify key frames in a gait sequence for both the double-support (two feet on the ground) and mid-stride phase of a gait. From these key frames they measure cues related to height, width, and other body proportions, and movement-related characteristics such as stride length, and amount of arm swing.

Data Sets

A database of sample gaits is essential for developing a silhouette-based gait recognition system. Little and Boyd [14] provided one of the earliest databases

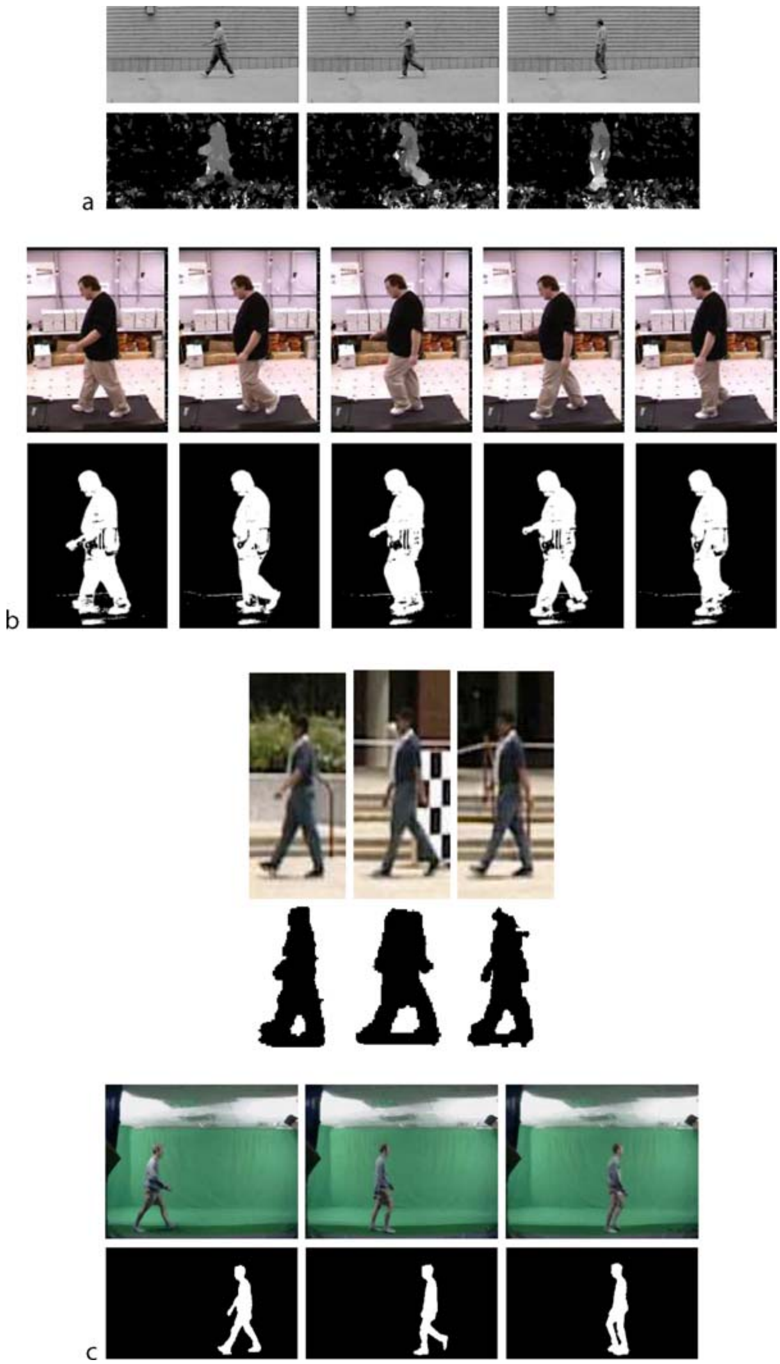


Gait Recognition, Silhouette-Based. Figure 1 Themes in silhouette-based gait recognition: (**a**) shape descriptors of the silhouette combine to form a gait signature from the motion of the gait, or (**b**) critical body dimensions are measured from key frames within the gait cycle. Existing methods use variations on both of these themes and can even combine them.

featuring seven sample gaits for each of six subjects, for a total of 42 gait sequences (Fig. 2a).

Gross and Shi [23] created the Motion of Body (MOBO) database (Fig. 2b). It features gait samples for 25 subjects. Each subject walks on a treadmill under four different conditions (slow, fast, on an incline, and carrying a ball) and from a variety of viewing angles. Segmented silhouettes are part of the database.

Sarkar et al. [24] present a large (1.2 Gigabytes) gait database as part of the *HumanID Gait Challenge Problem* associated with the Defense Advanced Research Projects Agency (DARPA) HumanID project (Fig. 2c).



Gait Recognition, Silhouette-Based. Figure 2 Example images from gait databases suitable for testing silhouette-based gait recognition: (**a**) Little and Boyd [14], (**b**) MOBO [21], (**c**) HumanID Gait Challenge [22], and (**d**) Shutler et al. [23]. All examples show raw video images in the top row and silhouettes or magnitude of the optical flow (Little and Boyd only) in the bottom row. The silhouettes shown for the Shutler et al. do not correspond to the images above.

The database contains samples for 122 subjects acquired in multiple sessions and under variable conditions. The *challenge problem* specifies a series of tests using the database as well as a reference algorithm to facilitate comparative testing by researchers.

Shutler et al. [25] created a database featuring over 100 subjects (Fig. 2d). The database contains sequences acquired over multiple sessions and features subjects walking from both left-to-right and right-to-left. Subjects walk on the ground or on treadmills, and in front of green screens (for chroma-keying) or in outdoor scenes.

Examples

Bhanu and Han [26] estimate upper bounds on the performance of gait recognition by equating gait with body dimensions, presented as plots of recognition rate versus gallery size for varying assumptions of accuracy. As one might expect with upper bounds, these rates are optimistic. Random guessing is a good lower bound on performance, but any practical biometric system must be much better. One can reasonably expect that a gait biometric system should perform at least as well as humans on moving light displays [2], i.e., 38% from a gallery of six.

Within these broad bounds, there are numerous examples of existing silhouette-based gait recognition systems. Most of these have been tested with one or more of the databases mentioned earlier. Examples include the work of Hayfron-Acquah et al. [19], Shutler and Nixon [17], Collins et al. [22], Bobick and Johnson [21], Ben-Abdelkader et al. [15, 18], Liu and Sarkar [27], Robledo and Sarkar [28]. Lee and Grimson [29], Little and Boyd [14, 20], and Wang et al. [30]. The best reported correct classification rates (CCR) are better than 90% from a gallery of approximately 100 people.

Summary

Human experience supported by psychological observation suggests that humans can be recognized by their gaits, which inspires gait biometric systems. Silhouette-based gait recognition systems convert images from a video gait sequence to silhouettes of the walker. Dynamic shape or body dimensions are measured from the silhouettes and combined to form a gait signature used for recognition. There are several databases available for testing silhouette-based gait recognition, and numerous published examples of successful recognition using these databases.

Related Entries

- ▶ Gait Recognition, Model-Based
- ▶ Gait Recognition, Motion Analysis for
- ▶ Human Detection and Tracking

References

1. Boyd, J.E., Little, J.J.: Advanced Studies in Biometrics: Summer School on Biometrics, Alghero, Italy, June 2–6, 2003, *Revised Selected Lectures and Papers (Lecture Notes on Computer Science)*, vol. 3161/2005, chap. Biometric Gait Recognition, pp. 19–42. Springer (2005)
2. Cutting, J.E., Kozlowski, L.T.: Recognizing friends by their walk: gait perception without familiarity cues. Bull. Psychonomic Soc. **9**(5), 353–356 (1977)
3. McGeer, T.: Passive dynamic walking. Int. J. Robot. Res. **9**(2), 62–82 (1990)
4. McGeer, T.: Passive walking with knees. In: IEEE International Conference on Robotics and Automation, pp. 1640–1645 (1990)
5. Kuo, A.D.: A simple model of bipedal walking predicts the preferred speed-step length relationship. J. Biomech. Eng. **123**, 264–269 (2001)
6. Kuo, A.D.: Energetics of actively powered locomotion using the simplest walking model. J. Biomech. Eng. **124**, 113–120 (2002)
7. von Tscharner, V., Goepfert, B.: Gender dependent emgs of runners resolved by time/frequency and principal pattern analysis. J. Electromyogr. Kines. **13**, 253–272 (2003)
8. Baumberg, A.M., Hogg, D.C.: Generating spatiotemporal models from examples. In: Sixth British Conference on Machine Vision, Vol. 2, pp. 413–422. Birmingham, UK (1995)
9. Brinkmann, R.: The Art and Science of Digital Compositing. Morgan-Kaufmann, San Diego, CA (1999)
10. Stauffer, C., Grimson, W.E.L.: Adaptive background mixture models for real-time tracking. In: Computer Vision and Pattern Recognition, vol. II, pp. 246–252. Fort Collins, CO, USA (1998)
11. Barron, J.L., Fleet, D.J., Beauchemin, S.: Performance of optical flow techniques. Int. J. Comput. Vision **12**(1), 43–77 (1994)
12. Troje, N.F.: Decomposing biological motion: a framwork for analysis and synthesis of human gait patterns. J. Vision **2**, 371–387 (2002)
13. Omlor, L., Giese, M.A.: Extraction of spatio-temporal primitives of emotional body expressions. Neurocomputing **70**(10–12), 1938–1942 (2007)
14. Little, J.J., Boyd, J.E.: Recognizing people by their gait: the shape of motion. Videre **1**(2), 1–32 (1998)
15. Ben-Abdelkader, C., Cutler, R., Davis, L.: Person identification using automatic height and stride estimation. In: 16th International Conference on Pattern Recognition, pp. 377–380. Quebec, Quebec (2002)
16. Ballard, D.H., Brown, C.M.: Computer Vision. Prentice-Hall, Englewood Cliffs, NJ (1982)
17. Shulter, J.D., Nixon, M.S.: Zernike velocity moments for sequence-based descriptiion of moving features. Image and Vision Computing **24**, 343–356 (2006)
18. Ben-Abdelkader, C. Cutler, R., Davis, L.: Motion-based recognition of people in eigengait space. In: 5th IEEE International Conference on Automatic Face and Gesture Recognition, 267–272 (2002)
19. Hayfron-Acquah, J.B., Nixon, M.S., Carter, J.N.: Automatic gait recognition by symmetry analysis. Pattern Recogn. Lett. **24**, 2175–2183 (2003)
20. Boyd, J.E.: Synchronization of oscillations for machine perception of gaits. Comput. Vision Image Understand. **96**, 35–59 (2004)

21. Bobick, A., Johnson, A.: Gait recognitin using static activity-specific parameters. In: Computer Vision and Pattern Recognition 2001, Vol. I, pp. 423–430. Kauai, HI (2001)
22. Collins, R.T., Gross, R., Shi, J.: Silhouette-based human identification from body shape and gait. In: Automatic Face and Gesture Recognition, pp. 351–356. Washington DC (2002)
23. Gross, R., Shi, J.: The cmu motion of body (mobo) database. Tech. Rep. CMU-RI-TR-01-18, Robotics Institute, Carnegie Mellon University (2001)
24. Sarkar, S., Phillips, J., Liu, Z., Robledo, I., Prother, P., Bowyer, K. W.: The humanid gait challenge problem: data sets, performance, and analysis. IEEE Trans. Pattern Analy. Mach. Intell. **27**(2), 162–177 (2005)
25. Shutler, J., Grant, M., Nixon, M., Carter, J.: On a large sequence-based human gait database. In: Fourth International Conference on Recent Advances in Soft Computing, pp. 66–71. Nottingham, UK (2002)
26. Bhanu, B., Han, J.: Bayesian-based performance prediction for gait recognition. In: IEEE Workshop on Motion and Video Computing, pp. 145–150. Orlando, Florida (2002)
27. Liu, Z., Sarkar, S.: Improved gait recognition by gait dynamics normalization. IEEE Trans. Pattern Analy. Mach. Intell. **28**(6), 863–876 (2006)
28. Robledo, I., Sarkar, S.: Statistical motion model based on the change of feature relationships: human gait-based recognition. IEEE Trans. Pattern Analy. Mach. Intell. **25**(10), 1323–1328 (2003)
29. Lee, L., Grimson, W.: Gait analysis for recognition and classification. In: Automatic Face and Gesture Recognition, pp. 148–155. Washington DC (2002)
30. Wang, L., Tan, T., Ning, H., Hu, W.: Silhouette analysis-based gait recognition for human identification. IEEE Trans. Pattern Analy. Mach. Intell. **25**(12), 1505–1518 (2003)

Gait, Forensic Evidence of

Niels Lynnerup, Peter K. Larsen
Laboratory of Biological Anthropology, Institute of Forensic Medicine, University of Copenhagen, Copenhagen, Denmark

Synonyms

Gait analysis; Perpetrator identification

Definition

Forensic evidence of gait, or forensic gait analysis, may be defined as analyses of gait performed in the service of the law. Usually, this involves analyses of criminal cases with the aim to characterize the gait of a perpetrator, and often to compare the gait of a perpetrator with the gait of a suspect. The results of the analyses may furthermore need to be presented in court. The methods involved in forensic gait analyses comprise morphological assessment of single gait features and kinematic assessment of body movement, often combined with photogrammetrics. The latter means that body segment lengths, stride length, etc. may be quantified and used in direct comparisons.

Introduction

Forensic analysis of ▶ gait has a lot of common ground with biometric gait recognition, but there are also some major differences. In terms of image capture, the imagery used in forensic gait analysis is mostly always acquired from CCTV, with the perpetrator specifically trying to conceal identity. In biometric systems, image capture of a person, or registration, takes place under specific circumstances, designed to maximize data quality, and obviously a person will willingly follow a set of guidelines in order to ensure proper registration. On the other hand a bank robber might try to hide his or her face to avoid facial recognition or wear baggy clothes to blur body morphology.

Biometric gait recognition systems may operate with various false accept or reject rates, which govern how exclusive the system is, and reflect the number of "wrong" registrations that can be tolerated. For example, a relatively high false reject rate (i.e., rejecting a person who otherwise should be cleared) is not a problem if the system is meant for a screening function, where rejection simply leads to an additional identity check. It is possible to generate computer models which can identify people by their gait with more than 90% success [1, 2], but these models are still based on a small number of people and require optimal conditions seldom found outside the laboratory [3]. Alternative biometric approaches use a description of a subject's silhouette, often with reportedly improved recognition performance [4]. In forensic gait analysis, the analysis is often specifically carried out to match a perpetrator with a suspect. If the case is made that there is a match then the suspect may be sentenced. This places a certain onus on the gait analysis and the scientists carrying out the analyses, and the prosecution and the defense may well challenge the findings of the gait analysis. This also means that when presenting the results of a forensic gait analysis, one has to be familiar with the legal prerequisites for the legal statements, and how expert evidence is adjudicated.

Technical Issues

Bank CCTV systems are often set up not to capture gait specifics, but rather to give fields of view covering office spaces, teller machines, etc., and also often to supervise the bank employees. It is a not uncommon experience when perusing CCTV footage after a bank robbery that the perpetrator is seen moving behind desks and tellers, so that only the upper part of his body is filmed. Also, the CCTV system may vary quite a lot in terms of technical quality, e.g., image capture frequencies, digital versus analog data storage, color versus b/w cameras (the latter often of sharper quality), and numerous supplier – dependent video and computer systems (e.g., in terms of data compression of video images).

The recording frequency should ideally be about 15 Hz allowing the examination of dynamic features such as, e.g., lateral instability in the knee at heel strike. Others have found a similar frequency sufficient for obtaining joint angles [5] and for automatic recognition of gait [2]. Lower recording frequencies may also be sufficient to examine features that are more static, although the gait will have a "jerky" appearance. Even at a low 5 Hz recording frequency, it has proved possible to examine gait parameters such as dorsal/plantar flexion at heel strike, degree of "push-off" at toe-off, and knee flexion during stance. At even lower recording frequencies, where the images really are still image series, specific gait-related characteristics may be noticed, e.g., a perpetrator with a bow-legged left knee. This means that even just one single image of the gait can sometimes be useful, if the gait feature captured can be deemed characteristic.

Gait

The ability to recognize other individuals is fundamental to human life. Identification by gait is a part of this process. Shakespeare made use of this in his play "The Tempest" where Ceres said: "High'st queen of state, Great Juno, comes; I know her by her gait". Psychophysiological studies have proved that the human being can recognize the sex of a walker [6] and friends and colleagues [7, 8] with a success rate up to 70–80%.

The authors derive from the Institution that has conducted what are, so far, the only scientific approaches to gait analysis for evidential procedures. The essay describes how evidential analysis was derived and presented in two forensic investigations [9, 10].

Gait analyses is performed by first gaining a purely morphological, ► anthroposcopic impression of the gait of a perpetrator. We then combine the basic ability to recognize people with biomechanical knowledge in order to give statements as to whether a suspect could have the same identity as a perpetrator in a given case by comparing the suspect's posture and joint angles during gait with the perpetrator's. A checklist has been developed for forensic gait analysis (Table 1). First described are the general characteristics of the perpetrator's gait following which are analyzed each of the joint rotations and segment movements found relevant for forensic gait analysis (by trial end error). When a profile of the perpetrator has been completed, each item of the list is compared to the recording of the suspect and stated if agreement (A), no agreement (N), or comparison not possible (-) is found. An item can be incomparable because either the joint rotation/movement cannot be analyzed due to poor quality of the surveillance recordings, or the recording of the suspect differs too much in some way from the recording of the crime such as differences in shoulder angles between suspect and perpetrator because of elevated shoulders in one of the recordings.

There have been several automated assessments of feature analysis for forensic and biometric purposes which show that there is a natural match between technique and observed performance [5]. Their features include foot angle (degree of outward rotation), the step length, and the mean hip joint angle, among others. Several other characteristic features have also been identified: inversion/eversion in the ankle during stance, lateral flexion in the dorsal column of the spine, and the knee angle in the frontal plane that would show lateral instability of the knee and signs of a person being bow-legged/knock-kneed. Furthermore, some of the characteristic features found were so special, such as limping, that it was not necessarily expected to be found in the 11 randomly selected subjects.

Gait, Forensic Evidence of. Table 1 IFM Copenhagen gait description form/checklist. The rightmost column is marked up either with "A" for agreement; "N" for no agreement; and "-" for incomparable (see text). The middle column is used for notes and specific observations

General	Notes on gait of perpetrator/ suspect	
Long/short steps, stiff/ relaxed gait with Narrow/ wide distance between the feet		
Signs of pathologic gait		
Feet/ankle joint		
Outward rotation		
Inversion/eversion		
Dorsal/plantar flexion at heel strike		
Degree of "push-off" at toe-off		
Knee		
Varus/valgus		
Knee flexion during stance		
Hip/pelvis		
Pelvis Abduction/adduction		
Pelvis Rotation		
Pelvis tilt		
Upper body		
Lateral flexion of spinal column		
Forward/backward leaning		
Rotation of the upper body during walk		
Shoulders		
Angle in frontal plane		
Forward/backward rotation		
Neck/head		
Posture in sagittal plane		
Head movements in frontal plane		
Quality of recordings/ other precautions		

It should be stressed that a rather wide definition of "gait analysis" is used, so that basically all bodily movements may be studied. Posture and stance may be quite specific. For example, when standing, one leg is more often weight-bearing than the other; there may be marked lordosis; the neck and shoulders may be more or less slouched, and so on. These stance-related characteristics have a bearing on how a person initiates or stops walking, and should thus also be involved in the analysis.

All the above features may be judged purely morphologically, but it may be of great evidentiary value to attach numbers to these features. Thus, the morphological approach is combined with photogrammetry in order to acquire specific measurements of body segment lengths and heights.

Photogrammetry in Association with Gait Analysis

Photogrammetry literally means measuring by photography. Photogrammetry enables the measurement of unknown values in two-dimensional space (2D) using known values within a single image [9, 10]. Another basic application of photogrammetry is measuring objects in three-dimensional space (3D) using photographs taken from different sides and angles. Zhao et al. [11] have also worked with video sequences in this respect. Jensen and Rudin [9] used a 2D method to measure the stature and several segment lengths in two different cases and found excellent agreement between perpetrator and suspect. Lynnerup and Vedel [10, 12] used a 3D method in the investigation of a bank robbery where the perpetrator was recorded simultaneously from two different cameras and found good agreement in bodily measurements when comparing the perpetrator to the suspect.

A first step in photogrammetry is calibration of the CCTV cameras. This is done by placing frames with targets on the locations (Fig. 1). The frames are photographed with both the surveillance video cameras and a calibrated digital camera. Using the digital camera images and special software (PhotoModeler Pro®) the points are measured and subsequently imported as control points ("fiduciary points"). A feature in PhotoModeler Pro® allows determination of the internal parameters of the surveillance video cameras, e.g., focal length, and subsequent calculation of the exact placement of the cameras. After calibration, still images from the surveillance cameras are input in PhotoModeler Pro®. The photogrammetrical method described here has the advantage that there is no need to ascertain the position of the perpetrator in relation to a measuring device. After calibration by fiduciary points, the photogrammetrical analysis produces points in a 3D space, and an evaluation of the goodness of fit may be made directly in the software. This then allows measuring body segment lengths, stature, etc. of a perpetrator in various locations and with various body stances (Fig. 2). The selection of anatomical points is done by choosing specific points such as the top of the head, eyes, and joint center-points on an image. This selection is made by judging anatomical landmarks, clothing displacement, comparison with images just before and after the chosen photo, etc. When then focusing on the other images of the same situation, but from other cameras, the program will indicate the epi-lines (the "line of sight") from the first image, as well as a line connecting the two joints. After selecting the identical anatomical points in this image, it is immediately apparent how good the fit is, and whether the points selected in the first image are adequate. Thus, the 3D coordinates are calculated not only by a simple averaging of points chosen from two images, but reflect a dynamic process where the tightness of the intersections of the epi-lines is minimized. The absolute error associated with measuring using photogrammetry as described is small. For instance, the height of a desk (bolted to the floor and not moved between the incident and the analysis) was measured by photogrammetry (result: 89.3 cm) and compared to

Gait, Forensic Evidence of. Figure 1 Measuring screens put up in a department store, in order to calibrate the CCTVs [10].

Gait, Forensic Evidence of. Figure 2 Screen shots of PhotoModeler Pro® interface, showing selection of points. Simultaneous images from different CCTV cameras are used to pinpoint concurrent anatomical points (and markers) seen from different POV. The lines between the points are to scale and thus hold accurate measures of distance [10].

an actual physical measurement (result: 90.0 cm), thus the error was 7 mm or less than 1%. Intra- and interobserver tests of photogrammetric measurements of bodily segments seem to indicate that the error associated with clearly identifiable body points, such as top of the head, eyes, ear lobes, among others is small. On the other hand, if the body points are hidden or obscured by clothing, such as joint center-points, then there is some variation, which needs to be taken into account.

Currently, research focuses on implementing the possibility of performing accurate measurements of a perpetrator even though images are from only one camera. To do this, a measuring screen is used, the contours of which can be accurately measured by the software, which is physically placed near to where the perpetrator was standing (it needs to place the perpetrator on a specific point on the floor). If the screen is oriented perpendicular to the camera, then the screen can be imported as a virtual screen overlaid the crime video-footage (Fig. 3). The perpetrator can then be measured against this screen, akin to seeing a person standing in front of a light-source, and whose shadow is cast of a screen or wall behind him.

Comparing Gait and Photogrammetry

As the forensic analysis mostly pertains to comparisons of perpetrators and suspects, then gait analyses and measuring of the suspect also has to be carried out. Owing to legal exigencies, this may be performed under very different settings and conditions, comprising hidden and overt image capture for gait, and hidden and overt photogrammetry. In some cases, legal circumstances have ruled out hidden image capture; in other cases the defense counsel was invited to be present (but without the knowledge of the suspect); and finally, the suspect has sometimes been filmed completely overt. Ideally, it is felt that gait image capture should be performed hidden, so as the suspect does not know he is being filmed. This is to ensure that the gait is not "changed". Preferably, the setting for performing the image capture should to some extent mimic the crime scene. For example, if at the crime scene there was a step at the entrance, which the suspect engaged in a distinct fashion, then filming the suspect engaging a somewhat likewise step would be obvious for comparison. If the crime scene images show a perpetrator walking down a corridor, either against or away from the CCTV camera, then a setting at police offices with a long corridor may be suitable. The filming usually takes place with ordinary DV-cameras, and is done by forensic technicians, but the setting would have been discussed in advance. For instance, a policeman can be instructed to accompany the suspect, but walking at a speed that matches the velocity of the perpetrator, because the gait speed may influence some of the features. For example, a lateral instability in the knee will be more pronounced at a higher gait speed.

Gait, Forensic Evidence of. Figure 3 Using the back-projection screen method (see text).

The photogrammetric measurement of the suspect is most easily performed overtly. Usually a corner in an office is identified with points fixed on the wall, and the suspect is asked to stand in the corner. Using two or three digital cameras, coupled to a computer, several sets of images of high quality for subsequent photogrammetry can be rapidly acquired. While height could be just as easily acquired using a stadiometer, it is found that the same measuring method (photogrammetry) should be used for comparing perpetrator and suspect. While at first glance stadiometer-measured stature might seem as a "gold-standard", it is also found that people almost automatically straighten themselves when asked to stand against a stadiometer, meaning in fact that a better agreement between subsequent measurings of stature by photogrammetry has been found, than between photogrammetry and a stadiometer. Of course, measuring the suspect by photogrammetry also makes it easier to measure other heights, such a floor to eye, floor to shoulder, and floor to ear-lobe.

Schöllhorn et al. [13] concluded that "identification of individuality seems to be impossible with single variables or specific parameters of single variables", so the more gait characteristics and bodily measurements of the perpetrator that can be extracted and compared to the suspect, the better.

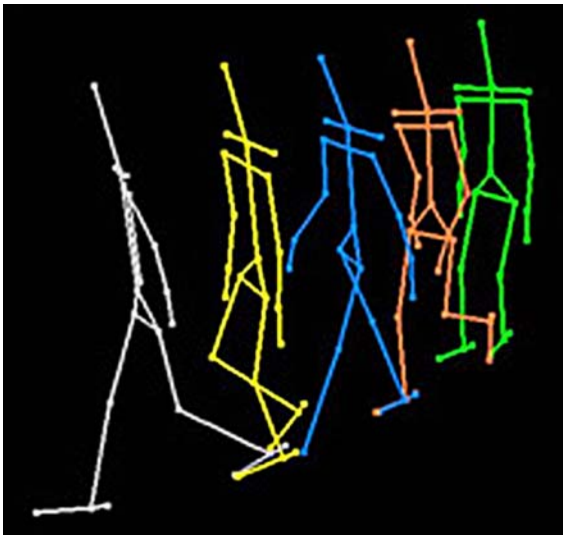
The Nature of Forensic Statements

In statements to the police it is noted what image material has been available, and what manner of image enhancing techniques had been used. The results of the above analyses are then presented, each followed by a separate conclusion, and each conclusion always summing up what features were found to indicate concordance between the suspect and the perpetrator, as well as features which seemed to indicate incongruity. Each item may therefore be seen as constituting single pieces of evidence. This renders a statistical approach, for instance the calculation of likelihood ratios for identity, based on the prevalence of certain facial and bodily traits, problematical [14].

Using the data sheets for gait analyses and photogrammetry fulfils three of the four guidelines in the ▸ Daubert Standard, a legal precedent set by Supreme Court of the United States [15], for determining whether expert witnesses' testimony is admissible as evidence: (1) the testimony in court is based on an empirically used technique, (2) the technique has been published in peer-reviewed literature and (3) it is generally accepted for use in forensic medicine. The last Daubert Guideline states that the reliability of the technique has been tested and potential error rates known.

Image based comparison will probably never achieve specific identification such as associated with DNA-typing and fingerprinting. However, analyzing gait and measuring stature and segment lengths of a perpetrator from surveillance video has the possibility of becoming a valuable forensic tool because the gait and the measures are an integrated part of the offender. At present, the methods can be used effectively to exclude a suspect if the gait and anthropometrical measures of the suspect and perpetrator are entirely different from each other. On the other hand, if the perpetrator and suspect do have a similar gait and similar measures, it can only be stated in court that the suspect cannot be excluded as the perpetrator. To give a more specific statement of the value of evidence, a database with gait characteristics and measures for a population of which the perpetrator and suspect could be referenced against. In theory, this might mean that if a perpetrator and a suspect are measured to have an unusual height, i.e., either very tall or very small, then this might in itself increase the likelihood of concordance between them, whereas very average heights would lower the likelihood (because then it might be almost anybody). In actuality, such databases are rather restricted, with often only specific subsamples of the entire population represented; populations also change in terms of e.g., immigration; and finally the perpetrator might well be from an entirely different part of the world. If comparing with such databases, it is important to stress that "given the perpetrator/suspect are drawn from the same population as the database", then their stature is more or less common, and the likelihood of concordance between them is more or less likely.

Future work will probably focus on a better integration of gait characteristics and photogrammetry in order to perform dynamic measurements of gait (basically "animating" the line models, cf. Fig. 4). This has the potential of calculating angles of flexion – extension in the major joints, step length, degree of side-to-side movement of the torso during walking, etc. These parameters may then further assist in discriminating between suspects and more specifically in identifying individual traits of gait.



Gait, Forensic Evidence of. Figure 4 Line models produced by the Photomodeler® based on the selected anatomical points, showing the gait.

Related Entries

- ► Gait Recognition, Model-Based
- ► Gait Recognition, Motion Analysis for
- ► Gait Recognition, Overview
- ► Gait Recognition, Silhouette-Based

References

1. Cunado, D., Nixon, M.S., Carter, J.N.: Automatic extraction and description of human gait models for recognition purposes. Comput. Vis. Image Underst. **90**, 1–41 (2003)
2. Urtasun, R., Fua, P.: 3D tracking for Gait Characterization and Recognition. 1*Computer Vision Laboratory.* Swiss Federal Institute of Technology, Lausanne, Switzerland. Report No. IC/2004/04(2004)
3. Nixon, M.S., Tan, T.N., Chellappa, R.: Human identification based on gait. Springer Science + Business Media, Inc. New York, NY, USA (2006)
4. Rahati, S., Moravejian, R., Kazemi, F.M.: Gait Recognition Using Wavelet Transform Information Technology: New Generations, 2008. Fifth International Conference on Information Technology. pp. 932–936. 7–9, proceedings, Universitat Triet, Germany (April 2008)
5. Wagg, D.K., Nixon, M.S.: On automated model-based extraction and analysis of gait, Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition, IEEE Seoul, Korea. 11–19. 17–19 (May 2004)
6. Kozlowski, L.T., Cutting, J.E.: Recognizing the sex of a walker from a dynamic point-light display. Percept. Psychophys. **21**, 575–580 (1977)
7. Cutting, J.E., Kozlowski, L.T.: Recognizing friends by their walk: Gait perception without familiarity cues. Bull. Psychon. Soc. **9**, 353–356 (1977)
8. Jokisch, D., Daum, I., Troje, N.F.: Self recognition versus recognition of others by biological motion: viewpoint-dependent effects. Perception **35**(7), 911–920 (2006)
9. Jensen, S.C., Rudin, L.I.: Measure: an interactive tool for accurate forensic photo/videogrammetry. In: Rudin, L.I., Bramble, S.K. (eds.) Investigative and Trial Image Processing 1995 July 13; 73–83. SPIE, San Diego, CA (1995)
10. Lynnerup, N., Vedel, J.: Person identification by gait analysis and photogrammetry. J. Foren. Sci. **50**(1), 112–118 (2005)
11. Zhao, G., Liu, G., Hua, L., Pietikainen, M.: 3D gait recognition using multiple cameras. In Automatic Face and Gesture Recognition, 2006. FGR 2006. Seventh International Conference, 10–12 April 2006: 529–534 (2006)
12. Lynnerup, N., Sejrsen, B., Vedel, J.: Identification by facial recognition, gait analysis and photogrammetry: The Anna Lindh murder. In: Brickley, M., Ferllini, R. (eds.) Forensic Anthropology: Case studies from Europe, pp. 232–244, Charles C. Thomas, Springfield, Ill., U.S.A. (2007)
13. Aitken, C.G.G.: Statistics and the evaluation of evidence for forensic scientists. Wiley and Sous, Chichester, UK (1995)
14. *Daubert v. Merrell Dow Pharmaceuticals, Inc.,* 509 U.S. 579 (1993)
15. Schollhorn, W.I., Nigg, B.M., Stefanyshyn, D.J., Liu, W.: Identification of individual walking patterns using time discrete and time continuous data sets. Gait Posture **15**(2), 180–186 (2002)

Gallery and Probe

Gallery is one of the data partitions in an algorithm-level biometric evaluation experiment. It is a collection of biometric templates that form the search dataset. Typically, these are representative of the enrolled templates in an actual biometric deployment scenario. In algorithm-level evaluations, care should be taken to have same number of representative templates per subject in the gallery. Probe is the second data partition in an algorithm-level evaluation experiment. It is a collection of biometric templates that need to be recognized or identified by matching against the gallery. In any given algorithm-level evaluation, the probes and gallery differ with respect to the covariate that is being studied. For example, to study the impact of viewpoint covariate, the gallery is chosen to be from one viewpoint and the probe is chosen to be from a different viewpoint. Since during actual operations biometric data is expected to arrive in a sequential fashion, it is not appropriate to normalize or adjust biometric matching scores over the probes. Neither is it appropriate to train on the probe data.

▶ Evaluation of Gait Recognition

Gaussian Mixture Density

▶ Gaussian Mixture Models

Gaussian Mixture Models

Douglas Reynolds
Lincoln Laboratory, MIT, Lexington, MA, USA

Synonyms

Gaussian mixture density; GMM

Definition

A Gaussian Mixture Model (GMM) is a parametric ▶ probability density function represented as a weighted sum of Gaussian component densities. GMMs are commonly used as a parametric model of the probability distribution of continuous measurements or features in a biometric system, such as vocal-tract related spectral features in a speaker recognition system. GMM parameters are estimated from training data using the iterative Expectation-Maximization (EM) algorithm or ▶ Maximum *A Posteriori* (MAP) estimation from a well-trained prior model.

Introduction

A Gaussian mixture model is a weighted sum of M component Gaussian densities as given by the equation,

$$p(\mathbf{x}|\lambda) = \sum_{i=1}^{M} w_i \, g(\mathbf{x}|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i), \tag{1}$$

where $\mathbf{x}$ is a D-dimensional continuous-valued data vector (i.e. measurement or features), w_i, $i = 1, \ldots, M$, are the mixture weights, and $g(\mathbf{x}|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$, $i = 1, \ldots, M$ are the component Gaussian densities. Each component density is a D-variate Gaussian function of the form,

$$g(\mathbf{x}|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) = \frac{1}{(2\pi)^{D/2}|\Sigma_i|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)' \, \Sigma_i^{-1} \, (\mathbf{x}-\boldsymbol{\mu}_i)\right\}, \tag{2}$$

with mean vector $\boldsymbol{\mu}_i$ and covariance matrix Σ_i. The mixture weights satisfy the constraint that $\sum_{i=1}^{M} w_i = 1$.

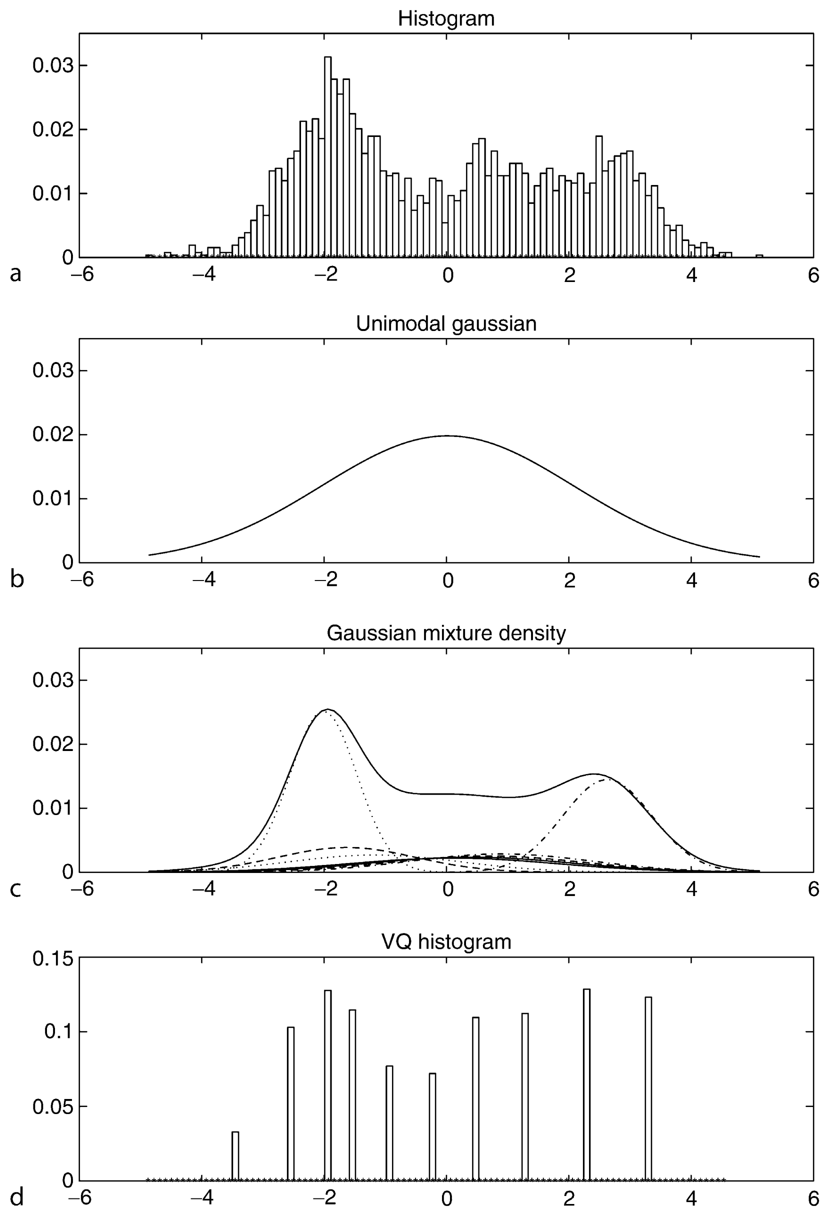
The complete Gaussian mixture model is parameterized by the mean vectors, covariance matrices and mixture weights from all component densities. These parameters are collectively represented by the notation,

$$\lambda = \{w_i, \boldsymbol{\mu}_i, \Sigma_i\} \qquad i = 1, \ldots, M. \tag{3}$$

There are several variants on the GMM shown in Eq. (3). The covariance matrices, Σ_i, can be full rank or constrained to be diagonal. Additionally, parameters can be shared, or tied, among the Gaussian components, such as having a common covariance matrix for all components, The choice of model configuration (number of components, full or diagonal covariance matrices, and parameter tying) is often determined by the amount of data available for estimating the GMM parameters and how the GMM is used in a particular biometric application.

It is also important to note that since the component Gaussian are acting together to model the overall feature density, full covariance matrices are not necessary even if the features are not statistically independent. The linear combination of diagonal covariance basis Gaussians is capable of modeling the correlations between feature vector elements. The effect of using a set of M full covariance matrix Gaussians can be equally obtained by using a larger set of diagonal covariance Gaussians.

GMMs are often used in biometric systems, most notably in speaker recognition systems [1, 2], due to their capability of representing a large class of sample distributions. One of the powerful attributes of the GMM is its ability to form smooth approximations to arbitrarily shaped densities. The classical unimodal Gaussian model represents feature distributions by a position (mean vector) and a elliptic shape (covariance matrix) and a vector quantizer (VQ) or nearest neighbor model represents a distribution by a discrete set of characteristic templates [3]. A GMM acts as a hybrid between these two models by using a discrete set of Gaussian functions, each with its own mean and covariance matrix, to allow a better modeling capability. Figure 1 compares the densities obtained using a unimodal Gaussian model, a GMM, and a VQ model. Plot (a) shows the histogram of a single feature from a speaker recognition system (a single cepstral value from a 25 second utterance by a male speaker); plot (b) shows a unimodal Gaussian model of this feature distribution; plot (c) shows a GMM and its ten underlying component densities; and plot (d) shows a histogram of the data assigned to the VQ centroid locations of a ten element codebook.



Gaussian Mixture Models. Figure 1 Comparison of distribution modeling. (**a**) histogram of a single cepstral coefficient from a 25 second utterance by a male speaker (**b**) maximum likelihood unimodal Gaussian model (**c**) GMM and its ten underlying component densities (**d**) histogram of the data assigned to the VQ centroid locations of a ten element codebook.

The GMM not only provides a smooth overall distribution fit, its components also clearly detail the multi-modal nature of the density.

The use of a GMM for representing feature distributions in a biometric system may also be motivated by the intuitive notion that the individual component densities may model some underlying set of *hidden* classes. For example, in speaker recognition, it is reasonable to assume the acoustic space of spectral related features corresponding to a speaker's broad phonetic events, such as vowels, nasals, or fricatives. These acoustic classes reflect some general speaker-dependent vocal tract configurations that are useful for characterizing speaker identity. The spectral shape of the *i*th acoustic class can in turn be represented by the mean $\boldsymbol{\mu}_i$ of the *i*th component density, and variations of the average spectral shape can be represented by the covariance matrix Σ_i. Since all the features used to train the GMM are unlabeled, the acoustic classes are hidden in that the class of an observation is unknown. A GMM can also be viewed as a single-state HMM with a Gaussian mixture observation density, or an ergodic Gaussian observation HMM with fixed, equal transition probabilities. Assuming independent feature vectors, the observation density of feature vectors drawn from these hidden acoustic classes is a Gaussian mixture [4, 5].

Maximum Likelihood Parameter Estimation

Given training vectors and a GMM configuration, the parameters, λ, are estimated which, in some sense, best match the distribution of the training feature vectors. There are several techniques available for estimating the parameters of a GMM [6]. By far the most popular and well-established method is ▶ maximum likelihood (ML) estimation.

The aim of ML estimation is to find the model parameters which maximize the likelihood of the GMM given the training data. For a sequence of T training vectors $X = \{\mathbf{x}_1, \ldots, \mathbf{x}_T\}$, the GMM likelihood, assuming independence between the vectors (The independence assumption is often incorrect but is needed to make the problem tractable.), can be written as,

$$p(X|\lambda) = \prod_{t=1}^{T} p(\mathbf{x}_t|\lambda). \tag{4}$$

Unfortunately, this expression is a nonlinear function of the parameters λ and direct maximization is not possible. However, ML parameter estimates can be obtained iteratively using a special case of the expectation-maximization (EM) algorithm [7].

The basic idea of the EM algorithm is, beginning with an initial model λ, to estimate a new model $\bar{\lambda}$, such that $p(X|\bar{\lambda}) \geq p(X|\lambda)$. The new model then becomes the initial model for the next iteration and the process is repeated until some convergence threshold is reached. The initial model is typically derived by using some form of binary VQ estimation.

On each EM iteration, the following re-estimation formulas are used which guarantee a monotonic increase in the model's likelihood value,

Mixture Weights

$$\bar{w}_i = \frac{1}{T}\sum_{t=1}^{T} \Pr(i|\mathbf{x}_t, \lambda). \tag{5}$$

Means

$$\bar{\boldsymbol{\mu}}_i = \frac{\sum_{t=1}^{T} \Pr(i|\mathbf{x}_t, \lambda)\ \mathbf{x}_t}{\sum_{t=1}^{T} \Pr(i|\mathbf{x}_t, \lambda)}. \tag{6}$$

Variances (diagonal covariance)

$$\bar{\sigma}_i^2 = \frac{\sum_{t=1}^{T} \Pr(i|\mathbf{x}_t, \lambda)\ x_t^2}{\sum_{t=1}^{T} \Pr(i|\mathbf{x}_t, \lambda)} - \bar{\mu}_i^2, \tag{7}$$

where σ_i^2, x_t, and μ_i refer to arbitrary elements of the vectors $\boldsymbol{\sigma}_i^2$, $\mathbf{x}_t$, and $\boldsymbol{\mu}_i$, respectively.

The *a posteriori* probability for component i is given by

$$\Pr(i|\mathbf{x}_t, \lambda) = \frac{w_i\ g(\mathbf{x}_t|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)}{\sum_{k=1}^{M} w_k\ g(\mathbf{x}_t|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)} \tag{8}$$

Maximum *A Posteriori* (MAP) Parameter Estimation

In addition to estimating GMM parameters via the EM algorithm, the parameters may also be estimated using Maximum *A Posteriori* (MAP) estimation. MAP estimation is used, for example, in speaker recognition applications to derive speaker model by adapting from a universal background model (UBM) [8]. It is also used in other pattern recognition tasks where limited labeled training data is used to adapt a prior, general model.

Like the EM algorithm, the MAP estimation is a two-step estimation process. The first step is identical to the "Expectation" step of the EM algorithm, where estimates of the sufficient statistics (These are the basic statistics needed to be estimated to compute the desired parameters. For a GMM mixture, these are the count, and the first and second moments required to compute the mixture weight, mean and variance.) of the training data are computed for each mixture in the prior model. Unlike the second step of the EM algorithm, for adaptation these "new" sufficient statistic estimates are then combined with the "old" sufficient statistics from the prior mixture parameters using a data-dependent mixing coefficient. The data-dependent mixing coefficient is designed such that mixtures with high counts of new data rely more on the new sufficient statistics for final parameter estimation and mixtures with low counts of new data rely more on the old sufficient statistics for final parameter estimation.

The specifics of the adaptation are as follows. Given a prior model and training vectors from the desired class, $X = \{\mathbf{x}_1 \ldots, \mathbf{x}_T\}$, the probabilistic alignment of the training vectors into the prior mixture components is determined (Fig. 2a). That is, for mixture i in the prior model, $\Pr(i|\mathbf{x}_t, \lambda_{\text{prior}})$ is computed as in Eq. (8).

Then compute the sufficient statistics for the weight, mean, and variance parameters $\mathbf{x}^2$ is shorthand for diag($\mathbf{xx}'$):

$$n_i = \sum_{t=1}^{T} \Pr(i|\mathbf{x}_t, \lambda_{\text{prior}}) \text{ weight,} \tag{9}$$

$$E_i(x) = \frac{1}{n_i} \sum_{t=1}^{T} \Pr(i|\mathbf{x}_t, \lambda_{\text{prior}})\mathbf{x}_t \text{ mean,} \tag{10}$$

$$E_i(\mathbf{x}^2) = \frac{1}{n_i} \sum_{t=1}^{T} \Pr(i|\mathbf{x}_t, \lambda_{\text{prior}})\mathbf{x}_t^2 \text{ variance.} \tag{11}$$

This is the same as the "Expectation" step in the EM algorithm.

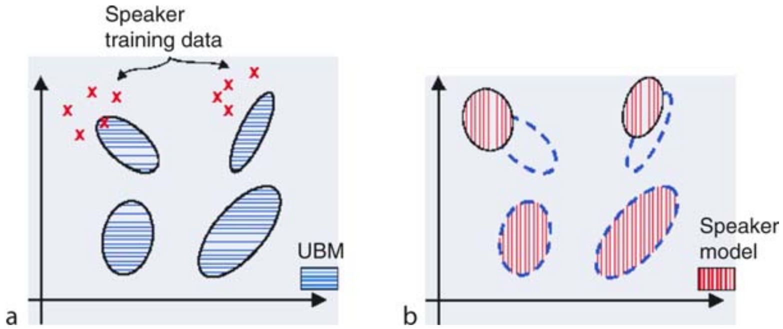
Lastly, these new sufficient statistics from the training data are used to update the prior sufficient statistics for mixture i to create the adapted parameters for mixture i (Fig. 2b) with the equations:

$$\hat{w}_i = \left[\alpha_i^w n_i / T + (1 - \alpha_i^w) w_i\right]\gamma \quad \text{adapted mixture weight,} \tag{12}$$

$$\hat{\boldsymbol{\mu}}_i = \alpha_i^m E_i(\mathrm{x}) + (1 - \alpha_i^m)\boldsymbol{\mu}_i \quad \text{adapted mixture mean,} \tag{13}$$

$$\hat{\boldsymbol{\sigma}}_i^2 = \alpha_i^v E_i(\mathbf{x}^2) + (1 - \alpha_i^v)(\boldsymbol{\sigma}_i^2 + \boldsymbol{\mu}_i^2) - \hat{\boldsymbol{\mu}}_i^2 \quad \text{adapted mixture variance} \tag{14}$$

The adaptation coefficients controlling the balance between old and new estimates are $\{\alpha_i^w, \alpha_i^m, \alpha_i^v\}$ for the weights, means, and variances, respectively. The scale factor, γ, is computed over all adapted mixture weights to ensure that they sum to unity. Note that the



Gaussian Mixture Models. Figure 2 Pictorial example of two steps in adapting a hypothesized speaker model. (**a**) The training vectors (x's) are probabilistically mapped into the UBM (prior) mixtures. (**b**) The adapted mixture parameters are derived using the statistics of the new data and the UBM (prior) mixture parameters. The adaptation is data-dependent, so UBM (prior) mixture parameters are adapted by different amounts.

sufficient statistics, not the derived parameters, such as the variance, are being adapted.

For each mixture and each parameter, a data-dependent adaptation coefficient α_i^ρ, $\rho \in \{w, m, v\}$, is used in the equations mentioned earlier. This is defined as

$$\alpha_i^\rho = \frac{n_i}{n_i + r^\rho}, \tag{15}$$

where r^ρ is a fixed "relevance" factor for parameter ρ. It is common in speaker recognition applications to use one adaptation coefficient for all parameters ($\alpha_i^w = \alpha_i^m = \alpha_i^v = n_i/(n_i + r)$) and further to only adapt certain GMM parameters, such as only the mean vectors.

Using a data-dependent adaptation coefficient allows mixture-dependent adaptation of parameters. If a mixture component has a low probabilistic count, n_i, of new data, then $\alpha_i^\rho \rightarrow 0$ causing the de-emphasis of the new (potentially under-trained) parameters and the emphasis of the old (better trained) parameters. For mixture components with high probabilistic counts, $\alpha_i^\rho \rightarrow 1$, causing the use of the new class-dependent parameters. The relevance factor is a way of controlling how much new data should be observed in a mixture before the new parameters begin replacing the old parameters. This approach should thus be robust to limited training data.

Related Entries

- ▶ Session Effects on Speaker Modeling
- ▶ Speaker Matching
- ▶ Speaker Recognition, Overview
- ▶ Universal Background Models

Acknowledgment

This work was sponsored by the Department of Defense under Air Force Contract FA8721-05-C-0002. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the United States Government.

References

1. Benesty, J., Sondhi, M., Huang, Y. (eds.): Springer Handbook of Speech Processing, vol. XXXVI. Springer, Berlin (2008)
2. Müler, C. (ed.): Speaker Classification I: Fundamentals, Features, and Methods, vol. 4343/2007. Springer: Lecture Notes in Computer Science, Berlin (2007)
3. Gray, R.: Vector qantization. In: IEEE ASSP Magazine, pp. 4–29 (1984)
4. Reynolds, D.A.: A Gaussian Mixture Modeling Approach to Text-Independent Speaker Identification. PhD thesis, Georgia Institute of Technology (1992)
5. Reynolds, D.A., Rose, R.C.: Robust text-independent speaker identification using Gaussian mixture speaker models. IEEE Trans. Acoust. Speech Signal Process. **3**(1), 72–83 (1995)
6. McLachlan, G. (ed.): Mixture Models. Marcel Dekker, New York, NY (1988)
7. Dempster, A., Laird, N., Rubin, D.: Maximum likelihood from incomplete data via the EM algorithm. J. Royal Stat. Soc. **39**(1), 1–38 (1977)
8. Reynolds, D.A., Quatieri, T.F., Dunn, R.B.: Speaker verification using adapted Gaussian mixture models. Digital Signal Process. **10**(1), 19–41 (2000)

GC

GC is an analytical chemistry separation technique which provides separation of mixtures on the basis of differential affinity between a liquid or solid stationary phase and a gas mobile phase.

▶ Odor Biometrics

Gelatin Pad

Gelatin lifting pads are designed for the lifting of fingerprints, footprints, dust marks, and trace evidences. They comprise three layers, the first layer is the carrier, which holds the second layer of thick low-adhesive gelatin in a pliable and flexible format. The thick gelatin layer is ideal for lifting evidence without sticking to the surrounding lift area. The third layer, a cover sheet, is a clear polyester film which is removed prior to lifting, and may be replaced once the lift is completed.

▶ Footwear Recognition

General Model

► Universal Background Models

Generalization

The classifier is designed to correctly classify unseen objects which are not used during the training process. Generalization represents the capacity of the classifier to respond to this task. When a classifier has a good generalization capacity, it can correctly classify unseen examples.

► Ensemble Learning
► Support Vector Machine

Generalization Error

The generalization error of a machine learning model is a function that measures how far the student machine is from the teacher machine in average over the entire set of possible data that can be generated by the teacher after each iteration of the learning process. It has this name because this function indicates the capacity of a machine that learns with the specified algorithm to infer a rule (or generalize) that is used by the teacher machine to generate data based only on a few examples.

► Image Pattern Recognition

Generative Classifier

A generative classifier is a classification algorithm that learns the full joint distribution of class and attribute values. As a result, it can generate labeled instances according to this distribution. To classify an unlabeled instance, one commonly uses the Bayes decision theory.

► Fusion, Quality-Based

Genetic Identification

Identification of a victim based on the victim's DNA samples.

► Dental Biometrics

Genuine Matching

Genuine matching is matching of two templates generated from the same finger.

► Fingerprint Matching, Automatic
► Individuality of Fingerprints

Genuine Sign

Genuine sign, also called genuine signature, is a legal sign. It is legally accepted as the registered sign.

► Signature Matching

Genuine/Impostor Attempt

In a genuine attempt, a biometric sample is compared against other biometric samples from the same subject. If similarity between the samples is not high enough, the subject will be wrongly rejected by the system. In an impostor attempt, a biometric sample is compared against biometric samples from other subjects.

If similarity between the samples is high enough, the subject will be wrongly accepted by the system. It should be noted that biometric samples from the same user are not necessarily similar (e.g., temporary injuries in the finger) and on the other hand, biometric samples from different users can be quite similar (e.g., signature forgeries).

▶ Fingerprint Databases and Evaluation

Geodesic

Geodesic is the integral curve between two points corresponding to the gradient direction of the intrinsic distance function of the manifold.

▶ Manifold Learning

Geometry Image

A geometry image is the result of representing all vertices of a 3D object (x, y, and z coordinates) as a simple 2D array of quantized points. Geometry images have at least three channels assigned to each u, v pair of coordinates, encoding geometric information (x, y, z coordinates) of a vertex in R^3, but surface normals and colors can also be stored using the same implicit surface parametrization. Creating a geometry image is accomplished by cutting an arbitrary mesh along a network of edge paths and parametrizing the resulting single chart onto a square.

▶ Face Recognition, 3D-Based

Global Fusion

Global fusion in the framework or multi-biometric score fusion refers to user-independent score fusion techniques in which a unique fusion function is used for all users, which is trained based on background data from a pool of users (both genuine and impostor scores).

▶ Fusion, User-Specific

Global Thresholding Techniques

Global thresholding technique is used to convert an image consisting of gray scale pixels to one containing only black and white pixels. Usually a pixel value of 0 represents white and the value 255 represents black with the numbers from 1 to 254 representing different grey levels. A threshold value Th is chosen in the range of 1–254 and each grey pixel P in the image is modified to either black or white according to the test.

If $P \geq Th$ then $P = 255$ (white) orelse $P = 0$ (black).

There are a number of ways to select the value of threshold Th depending on the nature of grey pixel distributions in the image.

▶ Hand Vein

Glottal Excitation

The glottal excitation corresponds to the pulsating flow of air that comes from the lungs through the vibrating vocal folds. This first process of the human speech production mechanism is named after the orifice between the vocal folds, the glottis.

▶ Speech Production

Glyph

A glyph is the shape of a handwriting sample. In Roman scripts, it may contain one letter or even a group of letters depending on the content of the sample. In oriental scripts, a glyph corresponds to a character which consists of a set of strokes.

► Signature Sample Synthesis

GMM

► Gaussian Mixture Models

Graph Matching

The configural identification of a face relating to the measurable distances between features and the relative ratios of height and width. A unique algorithm is created from the key points on the face; this algorithm is regarded as a unique biometric identifier.

► Face, Forensic Evidence of

Graphic Tablet

► Digitizing Tablet

Graphical User Interface

► User Interface, System Design

Graphometric Features

Graphometric features are intrinsic properties from an individual handwriting style, which may be employed by forensic experts during handwriting or signature recognition. These include curvature and pressure among others.

► Signature Features

Gray Scale

A continuous-tone image that has one component, which is luminescent.

► Vascular Image Data Format, Standardization

GRF (Ground Reaction Force)

The ground reaction force is, according to Newton's law of reaction, the force equal in magnitude but opposite in direction produced from the ground as the reaction to force the body exerts on the ground. The ground reaction force is used as propulsion to initiate and control the movement, and is normally measured by force sensor plates.

► Footstep Recognition

Ground-Truth

The actual facts of a situation, without errors introduced by sensors, software processing or human perception and judgment. For example the actual location of a minutia in a fingerprint image that could be used to check the accuracy of the location reported by a given automated minutiae extraction algorithm.

► SFinGe

Gummy Bear Finger

► Fingerprint Fake Detection