Stride and Cadence as a Biometric in Automatic Person Identification and Verification

Chiraz BenAbdelkader†, Ross Cutler‡,
and Larry Davis†
† University of Maryland, College Park
{chiraz,lsd}@umiacs.umd.edu
‡Microsoft Research
rcutler@microsoft.com

Abstract

We present a correspondence-free method to automatically estimate the spatio-temporal parameters of gait (stride length and cadence) of a walking person from video. Stride and cadence are functions of body height, weight, and gender, and we use these biometrics for identification and verification of people. The cadence is estimated using the periodicity of a walking person. Using a calibrated camera system, the stride length is estimated by first tracking the person and estimating their distance travelled over a period of time. By counting the number of steps (again using periodicity), and assuming constant-velocity walking, we are able to estimate the stride to within 1cm for a typical outdoor surveillance configuration (under certain assumptions). With a database of 17 people and 8 samples of each, we show that a person is verified with an Equal Error Rate (EER) of 11%, and correctly identified with a probability of 40%. This method works with low-resolution images of people, and is robust to changes in lighting, clothing, and tracking errors. It is view-invariant though performance is optimal in a near fronto-parallel configuration.

1 Introduction

There is an increased interest in gait as a biometric, mainly due to its non-intrusive and arguably non-concealable nature [6]. Consequently, considerable research efforts are being devoted in the computer vision community to characterize and extract gait dynamics automatically from video.

That each person seems to have a distinctive, idiosyncratic, way of walking is in fact easily understood from a biomechanics standpoint. Human ambulation consists of synchronized integrated movements of hundreds of muscles and joints in the body. Although these movements follow the same basic pattern for all humans, they seem to vary from one individual to another in certain details such as their relative timing and magnitudes. Much research in biomechanics and clinical gait analysis (among others) is devoted to the study of the inter-person and intra-person variability of gait (albeit not for the purpose of recognition, but rather to determine normal vs. pathological ranges of variation). The major sources of inter-person variability are attributed to physical makeup, such as body mass and lengths of limbs, while the sources for intra-person variability are things like walking surface, footwear, mood and fatigue [17, 29, 23]. However, the gait of any one person is known to be fairly repeatable when walking under the same conditions.

That gait is at once repeatable and defined by individual physical characteristics is encouraging. However, what makes this problem challenging and novel from a computer vision viewpoint, is that automatic extraction and tracking of gait features (i.e. such as joint positions) from marker-less video is still a very ambitious prospect. Most existing video-based gait analysis methods rely on markers, wearable instruments or special walking surfaces [23].

In this paper, we propose a robust correspondence-free method to estimate the spatio-temporal parameters of gait, i.e. cadence and stride length from low-resolution video based solely on the periodicity of the walking person and a calibrated camera. By exploiting the fact that the total distance walked by a person is the sum of individual piecewise contiguous steps, we are able to accurately estimate the stride. We then use a parametric Bayesian classifier that is based on the known linear relationship between stride length and cadence.

This method is in principle view-invariant, since it uses stride and cadence (which are inherently view-invariant) for classification. Its performance is optimal in a near-fronto-parallel configuration, which provides better estimates of both stride and cadence.

1.1 Assumptions

Our technique makes the following assumptions:

- People walk on a known plane with constant velocity (i.e. in both speed and direction) for about 10-15 seconds (i.e. the time for 20-30 steps).
- The camera is calibrated with respect to the ground plane.
- The frame rate is greater than twice the walking frequency.

Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition (FGR'02)
0-7695-1602-5/02 $17.00 © 2002 IEEE

2 Background and Related Work

Several approaches already exist in the computer vision literature on automatic person identification from gait (termed gait recognition) from video [22, 21, 19, 16, 15, 14, 2, 7, 30, 18]. Closely related to these are the methods for human detection in video, which essentially classify moving objects as human or non-human [31, 8, 27], and those for human motion classification, which recognize different types of human locomotion, such as walking, running, limping, etc. [4, 20].

These approaches are typically either holistic [22, 21, 19, 16, 14, 2] or model-based [4, 31, 20, 7, 30, 9, 18]. In the former, gait is characterized by the statistics of the spatiotemporal patterns generated by the silhouette of the walking person in the image. That is, a set of features (the gait signature) is computed from these patterns, and used for classification. Model-based approaches use a model of either the person's shape (structure) or motion, in order to recover features of gait mechanics, such as stride dimensions [31, 9, 18] and kinematics of joint angles [20, 7, 30].

Yasutomi and Mori [31] use a method that is almost identical to the one described in this paper to compute cadence and stride length, and classify the moving object as 'human' based on the likelihood of the computed values in a normal distribution of human walking. Cutler and Davis [8] use the periodicity of image similarity plots to estimate the stride of a walking and running person, assuming a calibrated camera. They contend that stride could be used as a biometric, though they have not conducted any study showing how useful it is as a biometric. In [9], Davis demonstrates the effectiveness of stride length and cadence in discriminating the walking gaits children and adults, though he relies on motion-capture data to extract these features.

Perhaps the method most akin to ours is that of Johnson and Bobick [18], in which they extract four *static* parameters, namely the body height, torso length, leg length and step length, and use them for person identification. These features are estimated as the distances between certain body parts when the feet are maximally apart (i.e. at the double-support phase of walking). Hence, they too use stride parameters (step length only) and height-related parameters (stature, leg length and torso length) for identification. However, they consider stride length to be a static gait parameter, while in fact it varies considerably for any one individual over the range of their free-walking speeds. The typical range of variation for adults is about 30cm [17], which is hardly negligible. This is why we use *both* cadence and stride length. Also, their method for estimating step length does not exploit the periodicity of walking, and hence is not robust to tracking and calibration errors.

3 Method

The algorithm for gait recognition via cadence and stride length consists of three main modules, as shown in Figure 1. The first module tracks the walking person in each frame, extracts their binary silhouette, and estimates their 2D position in the image. Since the camera is static, we use a non-parametric background modeling technique for foreground detection, which is well suited for outdoor scenes where the background is often not perfectly static (such as occasional movement of tree leaves and grass) [11]. Foreground blobs are tracked from frame to frame via spatial and temporal coherence: based on overlap of their respective bounding boxes in consecutive frames [13].

Once a person has been tracked for a certain number of frames, the second module first estimates the period of gait (T, in frames per cycle) and distance (W, in meters) travelled, then computes the cadence (C, in steps[1] per minute) and stride length (L, in meters) as follows [23]:

$$C = \frac{120 \cdot Fs}{T} \tag{1}$$

$$L = \frac{W}{n/T} \tag{2}$$

where n is the number of frames and F_s is the frame rate (in frames per second), and n/T is the (possibly non-discrete) number of gait cycles travelled over the n frames.

Finally, the third module either determines or verifies the person's identity based on parametric Bayesian classification of the cadence and stride feature vector.

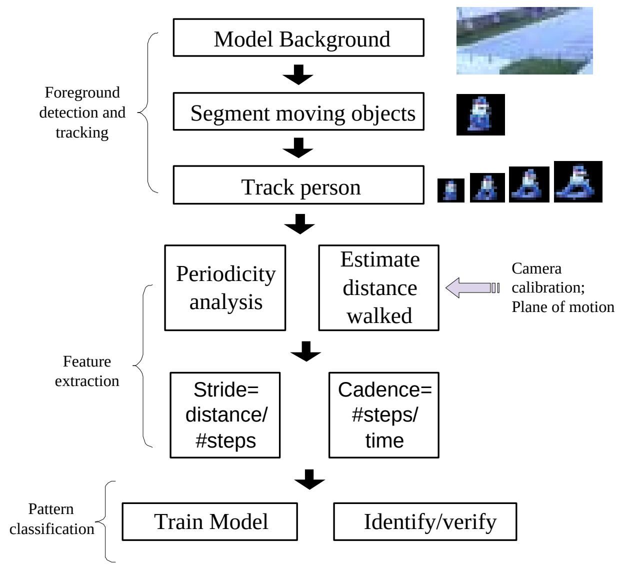


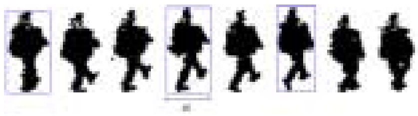
Figure 1. Overview of Method.

3.1 Estimating Period of Gait (T)

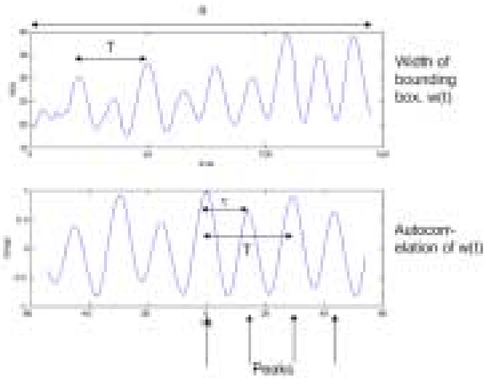
Because human gait is a repetitive phenomenon, the appearance of a walking person in a video is itself periodic. Several vision methods have exploited this fact to compute the period of human gait from image features [25, 8, 12]. In this paper, we simply use the width of the bounding box of the corresponding blob region, as shown in Figure 2, which is computationally efficient and has proven to work well with our background subtraction algorithm.

[1] Note that 1 cycle=2 steps.

Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition (FGR'02)
0-7695-1602-5/02 $17.00 © 2002 IEEE



(a)



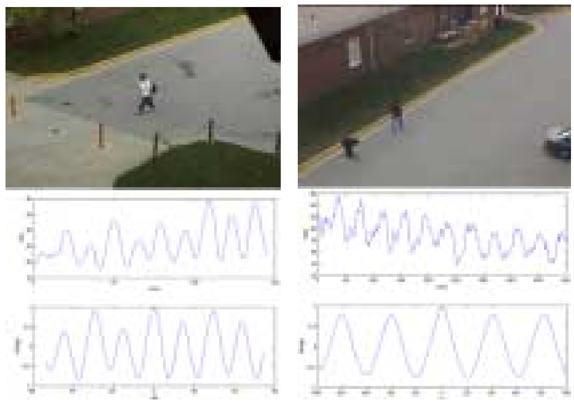
(b)

Figure 2. Computation of gait period via autocorrelation of time series of bounding box width of binary silhouettes.

To estimate the period τ of the width series $w(t)$, we first smooth it with a symmetric average filter of radius 2, then piecewise detrend it to account for depth changes, then compute its autocorrelation, $A(l)$ for $l \in [-lag, lag]$, where lag is chosen such that it is much larger than the expected period of $w(t)$. The peaks of $A(l)$ correspond to integer multiples of the period of $w(t)$. Thus we estimate τ as the average distance between every two consecutive peaks of $A(i)$.

One ambiguity arises, however, since $T = 2\tau$ for 'near' fronto-parallel sequences, and $T = \tau$ otherwise. When the person walks parallel to the camera (Figure 3(a)), gait appears bilaterally symmetrical (i.e. the left and right legs are almost indistinguishable) and we get two peaks in $w(t)$ in each gait period, corresponding to when either one leg is leading and is maximally apart from the other. However, as the camera viewpoint departs away from fronto-parallel (Figure 3(b)), one of these two peaks decreases in amplitude with respect to the other, and eventually becomes indistinguishable from noise.

While knowledge of the person's 2D trajectory in the image can help determine whether the camera viewpoint is fronto-parallel or not, we found that there is no clear cutoff between these two cases, i.e. how non-fronto-parallel the camera viewpoint can be before τ becomes equal to T. An alternative method to disambiguate these two cases is based on the fact that natural cadences of human walking lie in the range $[90, 130]$ steps/min [29], and so T must lie in the range $[0.923F_s, 1.333F_s]$ frames/cycle. Since τ and 2τ *cannot* both be in this interval, we choose the value that is.



(a) (b)

Figure 3. Width series and its autocorrelation function for (a) fronto-parallel, and (b) non-fronto-parallel sequences.

3.2 Estimating Distance Walked (W)

Assuming the person is walking in a straight line, the total distance traveled is simply the distance between the first and last 3D positions on the ground plane, i.e. $W = \|P_n - P_1\|$. The person's 3D position, (X_g, Y_g, Z_g), can be computed at any time from the 2D position in the image, (x_g, y_g), which is approximated as the center pixel of the lower edge of the blob's bounding box, as follows. Given the camera intrinsic (K) and extrinsic (E) matrices, and the parametric equation of the plane of motion, $P : aX + bY + cZ + d = 0$, and assuming perspective projection, then we have:

$$\begin{pmatrix} k_{11} & 0 & -x_g + k_{13} \\ 0 & k_{22} & -y_g + k_{23} \\ \hat{a} & \hat{b} & \hat{c} \end{pmatrix} E \begin{pmatrix} X_g \\ Y_g \\ Z_g \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -\hat{d} \end{pmatrix} \quad (3)$$

which is a linear system of 3 equations and 3 unknowns, where $(\hat{a}, \hat{b}, \hat{c}, \hat{d}) = (a, b, c, d) \cdot E^{-1}$ and k_{ij} is the (i, j)th element of K. Note, however, that this system does *not* have a unique solution if the person is walking directly towards or away from the camera (i.e. along the optical axis).

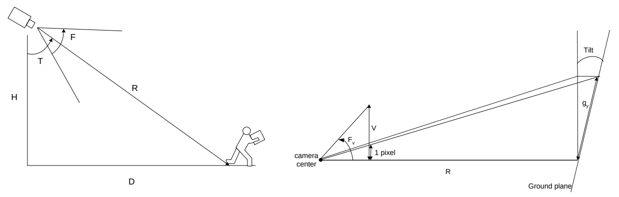
3.3 Error Analysis

According to Equations 1 and 2, the relative uncertainties in L and C satisfy: $\frac{\sigma_L}{L} \cong \sqrt{(\sigma_W/W)^2 + (\sigma_T/T)^2}$ and $\frac{\sigma_C}{C} \cong \frac{\sigma_T}{T}$, where σ_ζ generally denotes the absolute uncertainty in any estimated quantity ζ [3]. Thus to minimize both these, we need to minimize $\frac{\sigma_W}{W}$ and $\frac{\sigma_T}{T}$, which is achieved by estimating C and L over a sufficiently long sequence, as we explain below.

3.3.1 Uncertainty in T

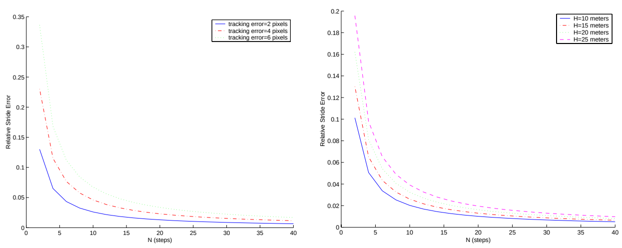
Based on the discussion in Section 3.1, $\sigma_T = \frac{\sigma_a}{N}$, where N is the number of gait cycles in the video sequence, and σ_a is the

Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition (FGR'02)
0-7695-1602-5/02 $17.00 © 2002 IEEE



(a) (b)

Figure 4. Geometry of stride error: (a) Outdoor surveillance camera configuration. (b) Estimating vertical ground sampling distance at the center of the image.



(a) (b)

Figure 5. Stride relative uncertainty as a function of (a) distance walked (W) and tracking error (σ_p) with $H = 15m$, and (b) distance walked (W) and camera height (H) with $\sigma_p = 2pixels$.

uncertainty in estimating the autocorrelation peaks. Since $N = \frac{n}{T}$ therefore $\frac{\sigma_T}{T} = \frac{\sigma_a}{n}$, and so $\frac{\sigma_T}{T}$ can be reduced by making n sufficiently large. We have empirically estimated $\sigma_a = 2$, and for example with $n = 390 \equiv 13 \cdot 30$ (which corresponds to the time it takes a person to walk 20 steps at 115 steps/min pace assuming $Fs = 30$), we get $\frac{\sigma_T}{T} = 0.5\%$.

3.3.2 Uncertainty in W

The ratio $\frac{\sigma_W}{W}$ is a decreasing function of W (assuming σ_W remains constant), regardless of whether σ_W is caused by *random* or *systematic* errors [3]. Thus, we can compensate for a large σ_W by making W sufficiently large. Since $W = \|P_n - P_1\|$, then $\sigma_W = 2\sigma_P$, and σ_P (the uncertainty in 3D position) is in turn approximated as a function of tracking error σ_p (in pixels), the ground sampling distance g (in meters per pixel), and camera calibration error σ_C (in meters) by: $\sigma_P \cong \sqrt{(g\sigma_p)^2 + \sigma_C^2}$.

Let us consider the outdoor camera configuration of Figure 4(a). The camera is at a height H, and looks down on the ground plane with tilt angle θ_v and vertical field of view F_v. $R = \sqrt{D^2 + H^2}$ is the distance along the optical axis from the camera to the ground plane, and $D = H \tan\theta_v$ is the distance from the camera base to the person. The vertical ground sampling distance is then estimated by $g_v = \frac{R}{I_v \cos\theta_v}$, where I_v is the vertical image resolution (see Figure 4(b)).

With $F_v = 12\,\text{deg}$, $\theta_v = 66\,\text{deg}$, $I_v = 240$, $\sigma_C = 2$, and $L = 1.5$, we plot $\frac{\sigma_L}{L}$ as a function of W, σ_p and H, as shown Figure 5. It is interesting to note that the stride length error is smaller than the ground sampling distance. For example, with $\sigma_p = 2$ pixels, $N = 20$ steps, and $H = 15$ m, we obtain $g_v = 82$ mm while $\sigma_L = 6.5$ mm. This is analogous to achieving sub-pixel accuracy in measurement of image features[2]. It is also important to note that our method compensates for quite a large σ_W. For example if $\sigma_p = 10$ pixels, then with $N = 25$ steps we get $\sigma_L = 45$ mm or a relative error of 4.5% (note that a person's image height in this camera configuration is typically no larger than 50 pixels).

3.4 Identification and Verification

The goal here is to build a supervised pattern classifier that uses the cadence and stride length as the input features to identify or verify a person in a given database (of training samples). We take a Bayesian decision approach and use two different parametric models to model the class conditional densities [10]. In the first model, the cadence and stride length of any one person are related by a linear regression, and in the second model they are assumed to vary as a bivariate Gaussian.

3.4.1 Model Parameter Estimation

Given a labelled training sample of a person's stride lengths and cadences, $(C_1, L_1), (C_2, L_2), .., (C_n, L_n)$, we use Maximum Likelihood (ML) estimation [10] to compute the model parameters of the corresponding class conditional densities.

- Linear Regression Model

  Stride length and cadence are known to vary approximately linearly for any one person over his/her range of natural (or spontaneous) walking speeds, typically in the range 90-125 steps/minute [17, 32]. Hence, for each class (person) ψ_i in the training set, we assume the linear regression model: $L = a_i C + b_i + \varepsilon_i$, where ε_i is random noise. The class conditional probability of a measurement $x \equiv (L, C)$ is then given by: $\Pr(x|\psi_i) = p_{\varepsilon_i}(r)$, where p_{ε_i} is the probability density of ε_i and $r = L - a_i C - b_i$ is the residual.

  Assuming ε_i is *white* noise (i.e. $\varepsilon_i \sim N(0, \sigma_i)$), the ML-estimate of the model parameters a_i and b_i are obtained via linear least squares (LSE) technique on the given training sample. Furthermore, the log-likelihood of any new measurement x with respect to each class ψ_i is obtained by: $l_i(x) = \log p_{\varepsilon_i}(r) = \frac{1}{2}(\frac{r}{s_i})^2 + \log s_i + \frac{1}{2}\log 2\pi$, where s_i is the sample standard deviation of ε_i. Since the above model only holds over a limited range of cadences $[Cmin_i, Cmax_i]$, i.e. $L = a_i C + b_i$ is not an infinite line, we set $l_i(x) = 0$ whenever C is outside $[Cmin_i - \delta, Cmax_i + \delta]$, where δ is a small tolerance (we typically use $\delta = 2$ steps/min). Since this range varies for each person, we need to estimate it from a representative training data.

[2] The following intuitive example will further elucidate this idea: suppose you are asked to measure the length of a poker card, and are given a tape ruler that is accurate to 1cm. To achieve greater accuracy, you take 20 cards from the same deck, and align them to be piecewise contiguous. You measure the length of all 20 cards and divide by the number of cards. This is 20 times the precision as when using a single card.

Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition (FGR'02)
0-7695-1602-5/02 $17.00 © 2002 IEEE

- Bivariate Gaussian Model

  A simpler model of the relationship between cadence and stride length is as a bivariate Gaussian distribution, i.e. $\Pr(x|\psi_i) \sim N(\mu_i, \Sigma_i)$ for the ith class. Although this model cannot be quite justified in nature (note for example that it implicitly assumes that cadences are not all equally probable, which is not necessarily true), we include it here for comparison purposes.

  The parameters of the model, μ_i and Σ_i, for the ith class are estimated respectively as the sample mean m_i and sample covariance S_i of the given training sample. The log-likelihood of a new observation $x \equiv (C, L)$ with respect to the ith class is then computed as $l_i(x) = \frac{1}{2}(x - m_i)^T \Sigma_i^{-1}(x - m_i) + \frac{1}{2}|\Sigma_i| + \frac{1}{2}2\pi$.

3.4.2 Performance Evaluation

We evaluate the performance of our system in *verify-mode* and *classify-mode* [5]. In the former, the pattern classifier is asked to check (or *verify*) whether a new measurement x verily belongs to some class ψ_k. For this, we use the decision rule:

$$l_k(x) \geq t \Rightarrow Accept$$

$$l_k(x) < t \Rightarrow Reject$$

where t is a decision threshold. A standard verification performance measure is the Receiver Operating Characteristic (ROC), which plots true acceptance rate (TAR) vs. the false acceptance rate (FAR) for various decision thresholds t. FAR is computed as the fraction of *impostor* attempts that are (falsely) accepted, and TAR is computed as the fraction of *genuine* attempts that are (correctly) accepted. In *identify-mode*, the classifier is asked to determine which class a given measurement x belongs to. For this, we use the Bayesian decision rule:

$$k = \max_i l_i(x) \Rightarrow \psi_k$$

A useful classification performance measure that is more general than classification error is the *rank order statistic*, denoted by $\lambda(k)$, which was first introduced by the *FERET protocol* (a paradigm for the evaluation of face recognition algorithms), and is defined as the cumulative probability that the real class of a test measurement is among its k top matches [24]. Obviously, this assumes we have a measure of the degree of match (or goodness-of-fit) of a given measurement x to each class in the database. We use the log-likelihood $l_i(x)$ as this measure. Note that the classification rate is equivalent to $\lambda(1)$.

4 Experiments and Results

The method is tested on a database of 131 sequences, consisting of 17 people with an average 8 samples each. The subjects were videotaped with a Sony DCR-VX700 digital camcorder in a typical outdoor setting, while walking at various cadences (paces). Each subject was instructed to walk on a straight line at a fixed speed a distance of about 90 feet (30 meters). Figure 6 shows a typical trajectory walked by each person in the experiment. The

Figure 6. Typical trajectory walked by each subject. Red dots correspond to repeating poses in the gait cycle.

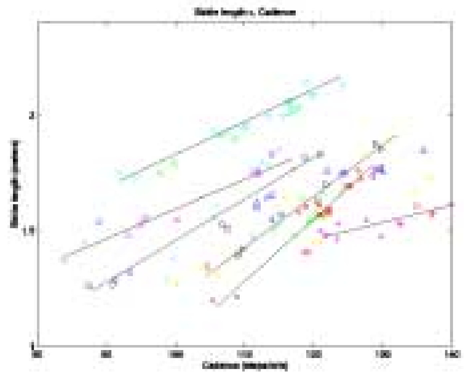


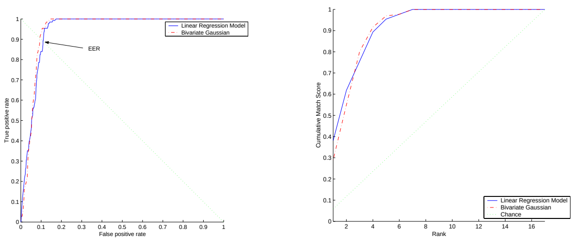
Figure 7. Stride length vs. Cadence for all 17 subjects. Note that the points corresponding to any one person (drawn with same color and symbol) are almost in a line. The best fitting line is shown for only 6 of the subjects.

same camera field of view was used for all subjects. The sequences were captured at 30 fps with an image size of 360x240. We used the technique described in this paper to automatically compute the stride length and cadence for each sample sequence. The results are plotted in Figure 7.

We estimate TAR and FAR via leave-one-out cross-validation [28, 26], whereby we train the classifier using all but one of the 131 samples, then verify the missed (or left out) sample on all 17 classes. Note that in each of these 131 iterations, there is one *genuine* attempt and 16 *impostor* attempts (since the left out sample is known a priori to belong to one of the 17 classes). Figure 8(a) shows the obtained ROC. Note that the point of Equal Error Rate (i.e. where FAR=1-TAR) corresponds to a FAR of about 11%.

We also use the leave-one-out cross-validation technique with the 131 samples to estimate the classification performance. Figure 8(b) plots the rank order statistic for the regression model, the Gaussian model, and the chance classifier (i.e. $\lambda(k) = k/17$).

Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition (FGR'02)
0-7695-1602-5/02 $17.00 © 2002 IEEE



(a) (b)

Figure 8. Performance evaluation results, based on a database of 131 samples of 17 people: (a) Receiver Operating Characteristic curve of gait classifier (b) Classification performance in terms of FERET protocol's CMC curve. Note the classification rate corresponds to $rank = 1$.

5 Conclusions and Future Work

We presented a parametric method for person identification by estimating and classifying their stride and cadence. This approach works with low-resolution images of people, is view-invariant, and robust to changes in lighting, clothing, and tracking errors. It achieves its accuracy by exploiting the nature of human walking, and computing the stride and cadence over many steps.

The classification results are promising, and are over 7 times better than chance for the bivariate Gaussian classifier. The linear regression classification can be improved by limiting the extrapolation distance for each person, perhaps using supervised knowledge of the range of typical walking speeds of each person.

Perhaps the best approach for achieving better person identification results is to combine the stride/cadence classifier with other biometrics, such as height, face recognition, hair color, and weight. We can also extend this technique to recognizing asymmetric gaits, such as a limping person.

Acknowledgment

The help of Harsh Nanda with the data collection and camera calibration, and the support of DARPA (Human ID project, grant No. 5-28944), are gratefully acknowledged.

References

[1] C. BenAbdelkader, R. Cutler, and L. Davis. Eigengait: Motion-based recognition of people using image self-similarity. In *AVBPA*, 2001.
[2] P. R. Bevington and D. K. Robinson. *Data reduction and error analysis for the physical sciences*. McGraw-Hill, 1992.
[3] L. W. Campbell and A. Bobick. Recognition of human body motion using phase space constraints. In *ICCV*, 1995.
[4] T. B. Consortium. *http://www.biometrics.org*. 2001.
[5] D. Cunado, M. Nixon, and J. Carter. Using gait as a biometric, via phase-weighted magnitude spectra. In *AVBPA*, 1997.
[6] D. Cunado, M. Nixon, and J. Carter. Gait extraction and description by evidence gathering. In *AVBPA*, 1999.
[7] R. Cutler and L. Davis. Robust real-time periodic motion detection, analysis and applications. *PAMI*, 13(2), 2000.
[8] J. W. Davis. Visual categorization of children and adult walking styles. In *AVBPA*, 2001.
[9] R. Duda, P. Hart, and D. Stork. *Pattern Classification*. John Wiley and Sons, 2001.
[10] A. Elgammal, D. Harwood, and L. Davis. Non-parametric model for background subtraction. In *ICCV*, 2000.
[11] I. Haritaoglu, R. Cutler, D. Harwood, and L. Davis. Backpack: Detection of people carrying objects using silhouettes. *CVIU*, 6(3), 2001.
[12] I. Haritaoglu, D. Harwood, and L. Davis. W4s: A real-time system for detecting and tracking people in 21/2 d. In *ECCV*, 1998.
[13] J. B. Hayfron-Acquah, M. S. Nixon, and J. N. Carter. Recognising human and animal movement by symmetry. In *AVBPA*, 2001.
[14] Q. He and C. Debrunner. Individual recognition from periodic activity using hidden markov models. In *IEEE Workshop on Human Motion*, 2000.
[15] P. S. Huang, C. J. Harris, and M. S. Nixon. Comparing different template features for recognizing people by their gait. In *BMVC*, 1998.
[16] V. Inman, H. J. Ralston, and F. Todd. *Human Walking*. Williams and Wilkins, 1981.
[17] A. Johnson and A. Bobick. Gait recognition using static activity-specific parameters. In *CVPR*, 2001.
[18] J. Little and J. Boyd. Recognizing people by their gait: the shape of motion. *Videre*, 1(2), 1998.
[19] D. Meyer, J. Psl, and H. Niemann. Gait classification with hmms for trajectories of body parts extracted by mixture densities. In *BMVC*, 1998.
[20] H. Murase and R. Sakai. Moving object recognition in eigenspace representation: gait analysis and lip reading. *PRL*, 17, 1996.
[21] S. Niyogi and E. Adelson. Analyzing and recognizing walking figures in XYT. In *CVPR*, 1994.
[22] J. Perry. *Gait Analysis: Normal and Pathological Function*. SLACK Inc., 1992.
[23] J. Philips, Hyeonjoon, S. Rizvi, and P. Rauss. The feret evaluation methodology for face recognition algorithms. *PAMI*, 22(10), 2000.
[24] R. Polana and R. Nelson. Detection and recognition of periodic, non-rigid motion. *IJCV*, 23(3), 1997.
[25] B. Ripley. *Pattern Recognition and Neural Networks*. Cambridge University Press, 1996.
[26] Y. Song, X. Feng, and P. Perona. Towards detection of human motion. In *CVPR*, 2000.
[27] S. Weiss and C. Kulikowski. *Computer Systems that Learn*. Morgan Kaufman, 1991.
[28] D. Winter. *The Biomechanics and Motor Control of Human Gait*. Univesity of Waterloo Press, 1987.
[29] C. Yam, M. S. Nixon, and J. N. Carter. Extended model-based automatic gait recognition of walking and running. In *AVBPA*, 2001.
[30] S. Yasutomi and H. Mori. A method for discriminating pedestrians based on rythm. In *IEEE/RSG Intl Conf. on Intelligent Robots and Systems*, 1994.
[31] V. M. Zatsiorky, S. L. Werner, and M. A. Kaimin. Basic kinematics of walking. *Journal of Sports Medicine and Physical Fitness*, 34(2), 1994.

Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition (FGR'02)
0-7695-1602-5/02 $17.00 © 2002 **IEEE**