Person Identification using Automatic Height and Stride Estimation

Chiraz BenAbdelkader†, Ross Cutler‡,
and Larry Davis†
† University of Maryland, College Park
College Park, MD 20742
{chiraz,lsd}@umiacs.umd.edu
‡Microsoft Research
rcutler@microsoft.com

Abstract

We present a parametric method to automatically identify people in monocular low-resolution video by estimating the height and stride parameters of their gait. Stride parameters (stride length and cadence) are functions of body height, weight, and gender. Previous work has demonstrated effective use of these biometrics for identification and verification of people. In this paper, we show that performance is significantly improved by using height as an additional discriminant feature. Height is estimated by segmenting the person from the background and fitting their apparent height to a time-dependent model. With a database of 45 people and 4 samples of each, we show that a person is correctly identified with 49% probability when using both height and stride parameters, compared with 21% when using stride parameters only. Height estimates for this configuration are accurate to within $\sigma = 3.5cm$. *This method works with low-resolution images of people, and is robust to changes in lighting, clothing, and tracking errors.*

1 Introduction

An emergent behavioral biometric is gait, i.e. the use of an individual's walking style to determine or validate identity, mainly due to its non-invasive nature not requiring direct participation or even cooperation of the subject [4]. Furthermore, because it can be measured 'at a distance', there is an increased interest in using gait features for human identification in surveillance applications.

The motivation for this line of research comes from psychophysical experiments [15, 6] as well as biomechanics studies [18, 13, 21], both of which have provided evidence that gait dynamics contain a signature that is characteristic of, and possibly unique to, each individual. However, complete and accurate characterization of gait dynamics requires knowledge of the kinematics of tens, if not hundreds, of body landmarks (such as joints and extremities) [22]. Furthermore, achieving this via automatic feature extraction and tracking in typical low-resolution surveillance video is error-prone due to occlusion, insufficient texture, etc.

In this paper, we propose a correspondence-free method to automatically and robustly extract four view-invariant gait variables (parameters) from low-resolution video. The first two variables characterize the person's apparent height (the mean and amplitude of oscillation), and the other two characterize the stride dimensions (the cadence and stride length). We use the term 'apparent height' to refer to the person's height while walking, which is a time-dependent quantity, and is different from, though related to, their *stature* (i.e. standing-height), as we shall explain later.

Accurate estimation of the gait features is achieved by exploiting the periodic nature of human walking, and computing the features over many steps. Cadence is estimated based on 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, then counting the number of steps (again using periodicity). Height parameters are estimated by robustly segmenting the person from the background and fitting their apparent height to a time-dependent model.

To evaluate the discrimination power of these four gait features, we apply K-nearest neighbor classification in this 4-D feature space on a database of fronto-parallel sequences of 45 people. The obtained identification accuracy of 49% shows that height and stride parameters alone are not enough to achieve perfect recognition; an intuitively expected result. More importantly, however, we show that these features are quite effective identity filters.

This method works with low-resolution images of people, and is robust to changes in lighting, clothing, and tracking errors. It is view-invariant, albeit the estimation of height and stride parameters is typically more accurate in a fronto-parallel configuration. Our method also makes a few assumptions, mainly that people walk upright with constant velocity for at least 3-4 seconds, that the frame rate is greater than twice the frequency of walking, and that the camera is calibrated with respect to the ground plane.

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. They are typically either holistic [17, 12, 11, 1] or model-based [5, 23, 16]. In the former, gait is characterized by the statistics of the spatiotemporal patterns generated by the walking person in the image. Model-based approaches use a model of either the person's shape or motion, in order to recover features of gait mechanics, such as stride dimensions [24, 7, 16] and kinematics of joint angles [5, 23].

3 Method

The algorithm consists of three main modules (Figure 1). The first module tracks the walking person in each frame, extracts their binary silhouette and 2D position in the image. Assuming the camera is static, we segment moving objects via a non-parametric background modeling technique [8], that is quite robust to lighting changes, camera jitter, background clutter, and shadows. Tracking of foreground blobs is achieved via simple spatial and temporal coherence [10].

Once a person has been tracked for a sufficient number of frames, the second module uses the obtained sequence of binary silhouettes to estimate the height and stride parameters. Finally, the third module determines the person's identity via standard pattern classification in this 4-D feature space.

The gait period is is estimated via periodicity analysis of the silhouette bounding box width (Figure 2), and the 3D trajectory is computed from the person's 2D trajectory in the image and the calibration parameters of the camera with respect to the ground plane of walking [2].

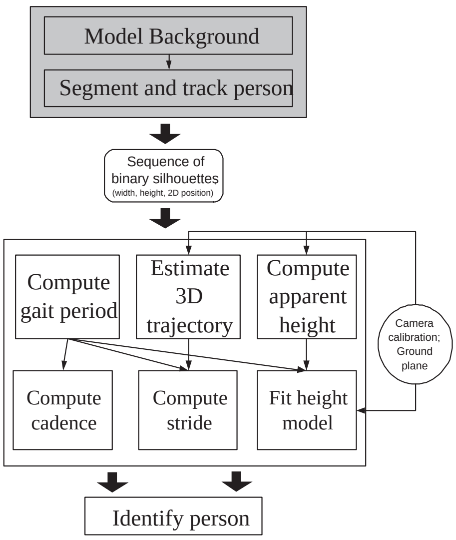


Figure 1. Overview of Method.

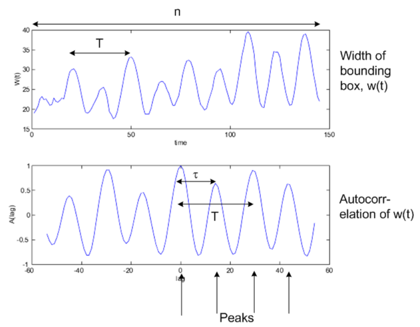


Figure 2. Computation of the period of gait.

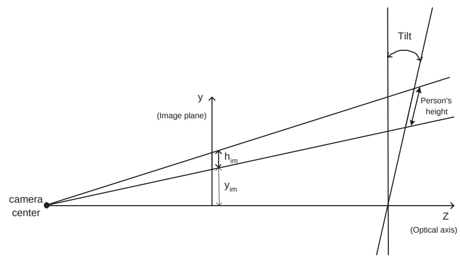


Figure 3. Estimating person's height from image height.

3.1 Estimating Stride Parameters

Given that the person has walked a distance W meters over n frames at T frames per cycle (stride), the cadence C (in steps per minute) and stride length L (in meters) are obtained by $C = 120 \cdot \frac{F_s}{T}$ and $L = \frac{W}{n/T}$, respectively, where F_s is the frame rate [19]. Note that n/T is the (possibly non-discrete) number of gait cycles travelled during n frames, and that W is computed as the distance between the initial (in first frame) and the last (in nth frame) 3D positions of the person, which hence relies on the assumption that the person is walking in a straight line.

3.2 Estimating Height Parameters

Human walking involves rhythmic up-and-down displacement of the upper body (from pelvis to head), hence the apparent bobbing of the head [13, 21]. Furthermore, these vertical movements must occur in a smooth sinusoidal manner for the conservation of energy [13]. Thus, the apparent height of a walking person can be modelled as a sinusoidal curve:

$$h(t) = \mu_h + \alpha_h \sin(\omega t + \phi) \tag{1}$$

The maximum height, $\mu_h + \alpha_h$, occurs at the *mid-stance* phase of walking (when the legs are closest together), and is slightly smaller than the person's stature (i.e. standing-height), typically within less than 1 cm. The minimum height, $\mu_h - \alpha_h$, occurs at the *mid-swing* phase of walking (when the legs are furthest apart).

Assuming the person is sufficiently far from the camera (i.e. orthographic projection applies), the apparent height at any time can be estimated from an image as:

$$h = \frac{Z \frac{h_{im}}{f}}{\cos\theta_v - \frac{y + h_{im}}{f}\sin\theta_v} \tag{2}$$

where y and h_{im} are respectively the person's vertical position and height in the image, θ_v is the tilt angle, f is the camera focal length, and Z is the distance from camera center to the person (see Figure 3). The person's height in the image is taken to be the bounding box height of the binary silhouette. Note this assumes no systematic (persistent) segmentation errors in the silhouette (such as shadows), no clothing style that modifies the person's apparent height significantly (such as a very long hat), and obviously that the person is walking upright.

Then, given the time-series h_t of apparent heights of a walking person measured over a video sequence of length n, and assuming a known frequency of gait ω (note $\omega \equiv \frac{2\pi F_s}{T}$), we estimate the three parameters of the model in Equation 1 via least squares

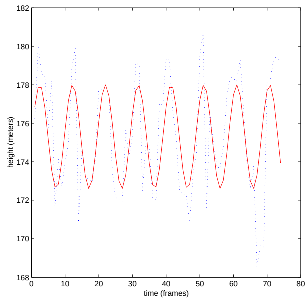


Figure 4. Height data (blue); model (red).

fitting as described in [3]. Specifically, assuming a data model

$$h_t = \mu_h + \alpha_h \cos(\omega t + \phi) + \varepsilon_t, t = 0, 1, .., n \tag{3}$$

the unknown parameters of the model that minimize the sum-square of the residuals ε are given by $\mu_h = \frac{1}{n}\sum_{t=1}^{n} x_t$ and $\alpha_h = \sqrt{A^2 + B^2}$, where $A = \frac{2}{n}\sum_{t=1}^{n}(x_t - \bar{x})\cos\omega t$ and $B = \frac{2}{n}\sum_{t=1}^{n}(x_t - \bar{x})\sin\omega t$. Figure 4 shows an example of a height series (blue dashed line) fitted to the model (red solid line) via above method. Here, $\mu_h \cong 175.3cm$ and $\alpha_h \cong 2.7cm$, and the person's real height is $177cm$.

4 Experiments and Results

The method is tested on a database of fronto-parallel sequences taken in an outdoor environment with 45 different people (7 females and 38 males) on 2 different days. Each subject walked a fixed straight path, approximately 5 meters-long, back-and-forth at their natural pace. The video sequences were captured at 20 fps and a full color resolution of 644x484 pixels (Figure 5). The distributions of the corresponding height and stride features are illustrated in Figure 6. Height estimates ($H \equiv \mu_h + \alpha_h$) for this configuration are accurate to within $\sigma = 3.5cm$.

Figure 5. Example of outdoor walking sequence used to test method.

We use KNN with $K = 1$ and leave-one-out cross validation technique [9] to estimate the classification rate. Table 1 shows two different measures of separability and the classification rate for each of the $\sum_{k=1}^{4}\binom{4}{k} \equiv 15$ feature subsets (S_b is the between-group scatter matrix, and S_w the within-group scatter matrix [14]).

Another useful performance measure is the *rank order statistic*, defined as the (cumulative) probability that the actual class of a test measurement is among its k top matches; k is called the rank [20]. Hence it effectively characterizes the 'filtering capability' of the classification features, i.e. how much of the database is eliminated as possible matches of the given person with some confidence. Figure 7(a) shows the rank order statistic for three different feature subsets. For example, it indicates that, based on all 4 gait features, a person is identified as one of 12 people with 90% confidence, which eliminates more than 2/3 of the database.

Obviously, the feature set (μ_h, α_h, C, L) achieves the best class separability and consequently the best classification accuracy. Note however that, although it is a much filter/discriminant than (C, L), it is only slightly better than (μ_h, C, L). In other words, using α_h does not 'buy' us much.

Finally, Figure 7(b) shows the classification rate as a function of the number of subjects in the database. Each point in this graph was obtained by randomly selecting different subsets of the subjects in the database and computing the average classification rate over all these subsets. Obviously classification performance degrades almost exponentially. However, we expect that it will level-off and converge asymptotically to the true classification rate.

Feature Set	$\frac{\|S_b\|}{\|S_w\|}$	$\frac{\|S_b\|}{\|S_w\|}$	Classification Rate
(μ_h)	11.6	10.6	.12
(μ_h, α_h)	24.9	12.1	.12
(C, L)	51.8	38.2	.21
(μ_h, C, L)	524.5	344.9	.41
(μ_h, α_h, C, L)	1045.7	328.4	.49

Table 1. Classification rate and separability measure for different feature subsets.

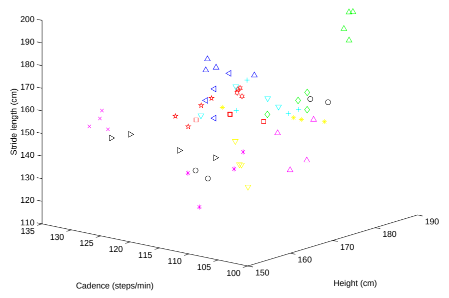
5 Conclusions and Future Work

We presented a parametric approach for human identification from low-resolution video using height and stride parameters of walking gait. It achieves its accuracy by exploiting the periodic nature of human walking, and computing the gait features over many steps. Based on a database of 45 subjects, we found a significant improvement in identification performance when using both height and stride parameters (49%) compared to using stride parameters only (21%).

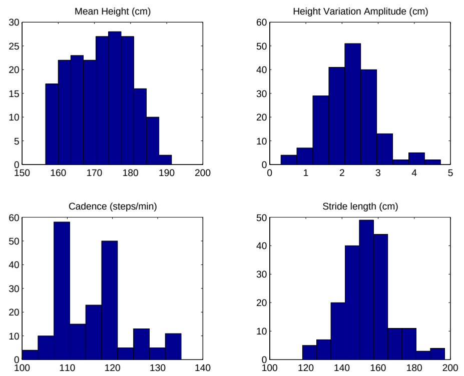
The primary goal of this work was not to achieve perfect recognition, but rather to show that stride and height are useful discriminant features for person identification, and that they contain independent pieces of information. Perhaps the best approach for achieving practical identification results (above 95%) is to combine these features with other biometrics, such as face recognition and hair color. This method can also be used to recognize asymmetric gaits (e.g. a limping person).

Acknowledgment

The support of DARPA (Human ID project, grant No. 5-28944) is gratefully acknowledged.

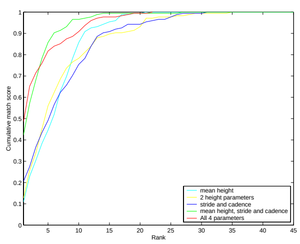


(a)

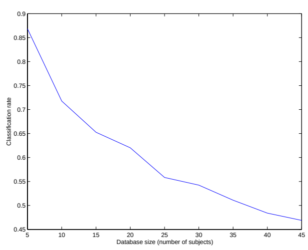


(b)

Figure 6. Distribution of data in feature space: (a) Scatter plot of μ_h, C and L. (b) Histograms of (clockwise from top-left to lower-right) μ_h, α_h, C and L, respectively.



(a)



(b)

Figure 7. (a) Identification performance in terms of rank order statistics. Note the classification rate corresponds to $rank = 1$. (b) Classification rate as a function of the number of subjects in the database.

References

[1] C. BenAbdelkader, R. Cutler, and L. Davis. Eigengait: A performance analysis with different camera viewpoints and variable clothing. In *FGR*, 2002.

[2] C. BenAbdelkader, R. Cutler, and L. Davis. Stride and cadence as a biometric in automatic person identification and verification. In *FGR*, 2002.

[3] P. Bloomfield. *Fourier Analysis of Time Series: an Introduction*. John Wiley and Sons, 1976.

[4] P. C. Cattin, D. Zlatnik, and R. Borer. Biometric system using human gait. In *Mechatronics and Machine Vision in Practice*, 2001.

[5] D. Cunado, M. Nixon, and J. Carter. Gait extraction and description by evidence gathering. In *AVBPA*, 1999.

[6] J. Cutting and L. Kozlowski. Recognizing friends by their walk: Gait perception without familiarity cues. *Bulletin Psychonomic Soc.*, 9(5):353–356, 1977.

[7] J. W. Davis. Visual categorization of children and adult walking styles. In *AVBPA*, 2001.

[8] A. Elgammal, D. Harwood, and L. Davis. Non-parametric model for background subtraction. In *ICCV*, 2000.

[9] K. Fukunaga. *Introduction to Statistical Pattern Recognition*. New York Academic Press, 1990.

[10] I. Haritaoglu, D. Harwood, and L. Davis. W4s: A real-time system for detecting and tracking people in 21/2 d. In *ECCV*, 1998.

[11] J. B. Hayfron-Acquah, M. S. Nixon, and J. N. Carter. Recognising human and animal movement by symmetry. In *AVBPA*, 2001.

[12] P. S. Huang, C. J. Harris, and M. S. Nixon. Comparing different template features for recognizing people by their gait. In *BMVC*, 1998.

[13] V. Inman, H. J. Ralston, and F. Todd. *Human Walking*. Williams and Wilkins, 1981.

[14] A. Jain and R. C. Dubes. *Algorithms for Clustering Data*. Prentice Hall Englewood Cliffs, 1988.

[15] G. Johansson. Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics*, 14(2), 1973.

[16] A. Johnson and A. Bobick. Gait recognition using static activity-specific parameters. In *CVPR*, 2001.

[17] J. Little and J. Boyd. Recognizing people by their gait: the shape of motion. *Videre*, 1(2), 1998.

[18] M. Murray. Gait as a total pattern of movement. *American Journal of Physical Medicine*, 46(1), 1967.

[19] J. Perry. *Gait Analysis: Normal and Pathological Function*. SLACK Inc., 1992.

[20] J. Philips, Hyeonjoon, S. Rizvi, and P. Rauss. The feret evaluation methodology for face recognition algorithms. *PAMI*, 22(10), 2000.

[21] J. Rose and J. G. Gamble. *Human Walking*. Williams and Wilkins, 1994.

[22] D. Winter. *The Biomechanics and Motor Control of Human Gait*. Univesity of Waterloo Press, 1987.

[23] C. Yam, M. S. Nixon, and J. N. Carter. Extended model-based automatic gait recognition of walking and running. In *AVBPA*, 2001.

[24] S. Yasutomi and H. Mori. A method for discriminating pedestrians based on rythm. In *IEEE/RSG Intl Conf. on Intelligent Robots and Systems*, 1994.