Analysis and Recognition of Walking Movements*

J.W. Davis S.R. Taylor
Dept. of Computer and Information Science
Ohio State University
Columbus, OH 43210
{jwdavis,taylors}@cis.ohio-state.edu

Abstract

We present an approach for recognizing human walking movements using low-level motion regularities and constraints. The features for classification are automatically extracted from video sequences of walkers. A multiplicative classification rule using statistical distances is then used to determine whether an unknown motion is consistent with normal walking patterns. Recognition results are shown distinguishing typical walking examples across multiple speeds from other non-walking locomotions.

1. Introduction

We can immediately recognize an *atypical* walking pattern (e.g., limping) from inconsistencies of that motion with our strong perceptual category for *normal* human walking. Furthermore, the normal walking category itself contains variations (e.g., stride length increases with walking speed [8]) which do not affect the classification. Most approaches to motion recognition have overlooked the dynamic changes exhibited within a category of movement, having only examined a small number of walkers each moving at a single pace.

In this paper, we describe a system for distinguishing normal walking movements (across multiple speeds) from other atypical or non-walking locomotions. Such a system would be particularly useful for automated visual surveillance and monitoring systems in helping to identify people that may be injured or that may require assistance by recognizing their "atypical" motion patterns (e.g., a person limping). By modelling low-level dynamic regularities and constraints in normal walking patterns, a sub-space of human walking motions is constructed for classification.

We begin with a review of related research (Sect. 2) and briefly describe the underlying motion category representation (Sect. 3). The method used to track the person in video (Sect. 4) and the motion features used to characterize the walking gait are then described (Sects. 5-6). We lastly present recognition results (Sects. 7-8) and conclude with a summary of the research (Sect. 9).

2. Related Work on Recognizing Gait

The most common approach to recognition of gait is the analysis of trajectory information derived from features of the walking body. In [13], the curvature trajectory of positions on a stick figure representation was examined using the Fourier transform to recognize cyclic walking motion. The frequency-based approach of [9] examined the phase relationships of periodic elements for the task of person identification from walking. Also addressing identification, [10] used the spatio-temporal braided patterns of the legs within the XYT volume to detect walking, and then extracted features of the pattern to identify individuals. To distinguish children from adults, [4] used correlated stride-based properties of their spatio-temporal walking styles for classification. Additionally, [2] incorporated a sequence of body signature skeletons into an HMM framework to determine a posture transition path for recognition.

Certain types of motion recognition can benefit from a model-based tracking pre-processor, but there are situations where direct motion recognition (with no part tracking) is applicable [5]. In [11], the periodic motion of pixels (as created from walking) was analyzed throughout a sequence using Fourier techniques. Another image-based periodicity approach is described in [3], where time-frequency analysis of a self-similarity measure between images in a sequence was used to detect and characterize periodic (walking) motion. In [7], individual walkers were recognized by combining an Eigenspace transformation and a Fisher Linear Discriminate function on background-subtracted silhouettes.

Our interests focus on the *categorization* of motion to classify the *type* of movement. We present a low-level ap-

*OSU Dept. of Computer and Information Science Technical Report OSU-CISRC-3/02-TR05.

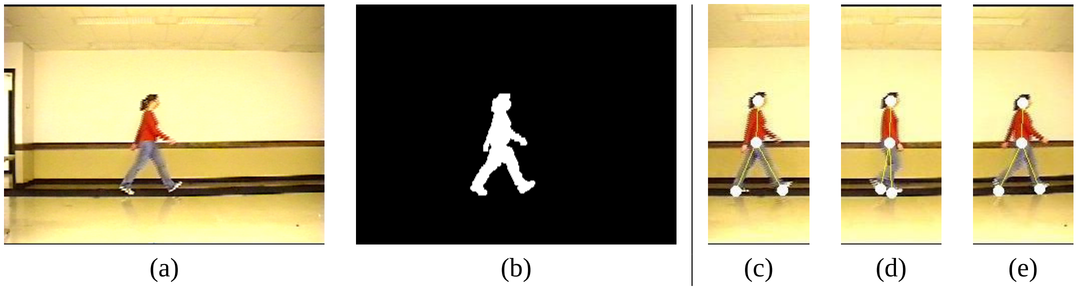


Figure 1. Person tracking. (a) Image of walker. (b) Background-subtracted silhouette. (c)-(e) Automatic identification of head, waist, and feet of walker throughout the sequence.

proach to distinguish walking from other non-walking locomotions.

3. Motion Category Representation

Our overall goal is to represent various types of human motion with categories to permit efficient recognition from partial (or limited) visual input. We define a motion category as a parameterization of dynamic regularities together with any constraints:

Motion Category = *Dynamic Regularities* + *Constraints*

The smoothly varying regularities describe the "genericity" (or acceptable variations) of the class. The structural constraints allow for any fixed values to be included in the representation. Multiple categories are separated by discontinuities in the regularity and constraint parameterizations.

To illustrate a simple motion category, consider the class of pendular motion for a swinging particle at the end of a light inextensible cord. The *structural constraints* include the choices for the particle mass M, cord length L, and gravity G. The *dynamic regularity* (for small amplitudes) consists of the correlation of the period of movement T to various ratios of L/G, as defined by $T = 2\pi\sqrt{L/G}$. We can then predict (or verify) the period of movement T for a pendulum from the observed length of the cord L (mass does not effect the period). We seek to categorize human motion in a similar manner to classify different types of movements from a small number of correlated features and constraints.

4. Person Tracking in Video

To track a person in video, we use a method based on the W4 approach [6] to locate the head, waist, and feet. The method is best suited to fronto-parallel views of the walker, but can accommodate slightly different viewpoints. We begin by extracting the person's silhouette in each video frame with a standard background-subtraction technique using RGB pixel differences, dilation, and removal of small pixel regions (See Fig. 1.a,b). After applying a bounding box to the silhouette, we vertically separate the head, torso, and leg regions using average anatomical proportions.

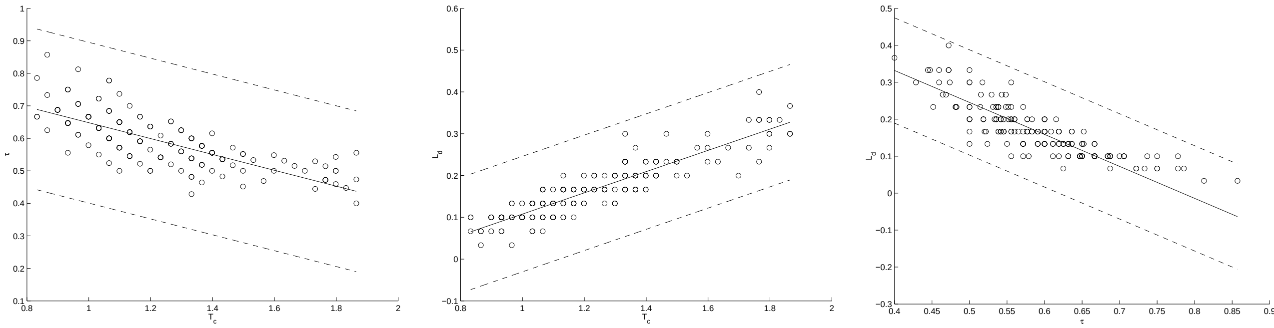
The head point is found as the centroid of the silhouette pixels in the head region. The waist location within the torso region is determined from the mean x-value of the silhouette pixels in the torso region and the expected y-coordinate of the waistline. To locate the feet, the region of the bounding box below the waistline is divided into equal left and right halves, where each half is assumed to contain one leg. Within each leg region, an Eigenvector line fitting process is used to fit a line to the silhouette pixels within that region. The foot points are determined to be the most extreme silhouette pixels in the direction of those lines away from the waist. Results of the method are shown in Fig. 1.c-e. Trajectories of the points are then lowpass filtered and stabilized (relative to the head point).

5. Low-Level Gait Features

Rather than attempting to match joint angles, limb lengths, or poses, we employ four motion properties of the tracked feet locations to demonstrate the approach. Three features together comprise the dynamic regularities and the remaining feature is a structural constraint. Additional features showing biomechanical regularities and constraints in locomotion to extend the model can be found in [8].

5.1. Dynamic Regularity Features

- **Cycle Time:** T_c. One of the most fundamental descriptors of locomotion is the cycle time of a leg, which decreases with increasing walking speed. Cycle time is determined by calculating the time difference between



(a) (b) (c)

Figure 2. Dynamic feature regularities. (a) τ vs. T_c. (b) L_d vs. T_c. (c) L_d vs. τ. Linear prototypes with $\pm 5\sigma$ are shown.

neighboring minimal (or maximal) peaks in each foot's x-trajectory.

- **Swing-Stance Ratio:** τ. Our second feature is the ratio of swing and stance times of a leg. The swing time is approximated by the temporal interval of the back-to-front translation of the leg, and is determined by calculating the time difference between a minimal peak and the subsequent maximal peak in the corresponding foot's x-trajectory. Similarly, the stance time is approximated by the front-to-back interval of the leg, and corresponds to the time difference between a maximal peak and the subsequent minimal peak in the x-trajectory. The ratio of swing-to-stance increases as a person walks faster.

- **Double-support Time:** L_d. Our next feature approximates the time that both feet together are in the stance phase (double-support), and is measured by the time difference from when one leg begins its stance (onset of return time in leg 1) to when the other leg starts its swing (onset of moving forward in leg 2). The double support time is determined by the temporal difference between a minimal peak in one foot's x-trajectory and the nearest maximal peak in the other foot's x-trajectory. As a person walks faster, the double-support interval becomes increasingly smaller.

5.2. Constraints

In this paper we include only one category constraint, but other physical or motion constraints could be added.

- **Extension Angle:** θ. The extension angle is the fronto-parallel angle made from the xy position of the maximal forward extension of the leg in front of the body to the most distant extension of the leg behind the body. We calculate this feature for a given foot as the angle between the xy coordinates of maximal spatio-temporal curvature in the front and back halves of the xy-trajectory for a single walking cycle. This angle is nearly zero for people during walking.

6. Walking Category

To construct the motion category for walking, we recorded a video database of 14 male and female walkers each moving at slow, medium, and fast paces. From the person tracking output, we computed our motion features (T_c, τ, L_d, θ) for multiple cycles of each individual. Each walking cycle was analyzed separately with no averaging of features over multiple cycles. Also, each leg was analyzed separately.

The three pairwise combinations of the temporal features were used to construct the dynamic regularities. In Fig. 2.a, we present τ vs. T_c that characterizes the typical walking timings for a single leg (correlation of $\rho = -0.75$). In Fig. 2.b, L_d vs. T_c represents how the two legs move in temporal relation to one another over multiple speeds ($\rho = 0.91$). Lastly in Fig. 2.c, we see L_d vs. τ that also relates the temporal movement of the two legs ($\rho = -.84$). The constraint θ had a nearly Gaussian distribution centered around zero degrees (σ_θ = 1.479).

7. Recognition Method

To determine class membership for a new movement pattern to this walking category, statistical measures of the new features are combined into a single binary classification result.

For the regularity features, the perpendicular Mahalanobis distance of the new feature values to linear regularity prototypes are computed using

$$R_i = \frac{|\alpha_i X_i - Y_i + \beta_i|}{\sigma_i \sqrt{\alpha_i^2 + 1}} \tag{1}$$

with $\{X_1 = T_c, Y_1 = \tau, \alpha_1 = -.244, \beta_1 = .892\}$ for the single-leg regularity, and $\{X_2 = T_c, Y_2 = L_d, \alpha_2 = .254, \beta_2 = -.146\}$ and $\{X_3 = \tau, Y_3 = L_d, \alpha_3 = -.865, \beta_3 = .678\}$ for the coupled-legs regularities. The α_i, β_i parameters were determined from training using an Eigenvector line fitting process that minimizes the sum of squares of the *perpendicular* distances from the training points to the linear prototype. The standard deviations σ_i were computed as the overall deviation along each regularity ($\sigma_1 = .049$, $\sigma_2 = .028$, $\sigma_3 = .028$). We show the prototype models with $\pm 5\sigma$ class boundaries in Fig. 2.

The constraint features are examined using the standard Mahalanobis distance. The distance for a new θ to the model is computed by

$$C_\theta = \frac{|\theta - \bar{\theta}|}{\sigma_\theta} \tag{2}$$

with $\bar{\theta} = -.419$ and σ_θ = 1.479 degrees.

To be classified as typical walking, the regularity and constraint distances (R_i, C_θ) are thresholded and incorporated into the binary result M using the product classification rule [12] to discount those movements having any non-conforming properties:

$$M = \hat{C}_\theta \prod_i \hat{R}_i, \qquad \hat{X}_k = \begin{cases} 1 & X_k \leq 5 \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

The result is a single binary membership assignment for the movement to the walking category (0=reject, 1=accept).

8. Experiments

To test the approach, we collected several new examples of typical walking and other non-walking locomotion sequences. The set of typical walking included slow/medium/fast paces of one person filmed a week apart from her training sequences and of three new people not used in training. Additionally, a single sequence from the walking dataset of Baumberg and Hogg [1] and of a 4-year-old child were tested. The non-walking locomotions included video sequences of people limping, skipping, jogging, and marching (two styles).

The classification results using the gait features and Eqn. 3 were calculated for the sequences. All 14 normal walking sequences were correctly classified (most distances R_i, C_θ were within a 2.5σ range) and all 5 non-walking sequences were correctly rejected (most having at least one distance $> 10\sigma$).

One problem with our large acceptance region for classification (5σ was needed to accept all training examples) is that some atypical walking sequences can be wrongly accepted into the class. To demonstrate, we tested a new "fast" walking sequence created by increasing the frame rate for a slow-paced walking sequence. The sequence should fail the temporal regularity tests, but instead it had deviations within the 5σ boundary (though above 3σ). Therefore, more precise feature extraction and tighter acceptance regions are necessary to better encapsulate the motion regularities.

9. Summary

We presented an approach for representing and recognizing walking movements using a small number of low-level motion regularities and constraints. To construct the walking category, motion features were computed from real video sequences of 14 people walking at multiple speeds. A multiplicative classification rule using statistical distances was used to determine whether an unknown motion was consistent with typical walking patterns. Results using several normal walking patterns and non-walking locomotions demonstrated the ability of the approach using only a small number of motion features.

References

[1] A. Baumberg and D. Hogg. Learning flexible models from image sequences. In *Proc. Euro. Conf. Comp. Vis.*, pages 299–308, 1994.

[2] I. Chang and C. Huang. The model-based human body motion analysis system. *Image and Vision Comp.*, 18(14):1067–1083, 2000.

[3] R. Cutler and L. Davis. Robust real-time periodic motion detection, analysis, and applications. *IEEE Trans. Patt. Analy. and Mach. Intell.*, 22(8):781–796, 2000.

[4] J. Davis. Visual categorization of children and adult walking styles. In *Proc. Int. Conf. Audio- and Video-based Biometric Person Authentication*, pages 295–300, 2001.

[5] J. Davis and A. Bobick. The representation and recognition of action using temporal templates. In *Proc. Comp. Vis. and Pattern Rec.*, pages 928–934. IEEE, 1997.

[6] I. Haritaoglu, D. Harwood, and L. Davis. W4: Who? When? Where? What? A real time system for detecting and tracking people. In *Proc. Int. Conf. Auto. Face and Gesture Recog.*, pages 222–227, 1998.

[7] P. Huang, C. Harris, and M. Nixon. Recognising humans by gait via parametric canonical space. *Artif. Intell. in Eng.*, 13:359–366, 1999.

[8] V. Inman, H. Ralston, and F. Todd. *Human Walking*. Williams & Wilkins, Baltimore, 1981.

[9] J. Little and J. Boyd. Recognizing people by their gait: the shape of motion. *Videre*, 1(2):2–32, 1998.

[10] S. Niyogi and E. Adelson. Analyzing and recognizing walking figures in XYT. In *Proc. Comp. Vis. and Pattern Rec.*, pages 469–474. IEEE, 1994.

[11] R. Polana and R. Nelson. Low level recognition of human motion. In *Workshop on Motion of Nonrigid and Articulated Objects*, pages 77–82. IEEE Computer Society, 1994.

[12] E. Smith and D. Medin. *Categories and Concepts*. Harvard University Press, Cambridge, MA, 1981.

[13] P. Tsai, M. Shah, K. Keiter, and T. Kasparis. Cyclic motion detection for motion based recognition. *Pat. Rec.*, 27(12):1591–1603, 1994.