Appears in *International Conference on Audio- and Video-based Biometric Person Authentication*, June 2001, Halmstad, Sweden. pp. 295-300.

Visual Categorization of Children and Adult Walking Styles

James W. Davis

Motion Recognition Laboratory
Dept. of Computer & Information Science
Ohio State University
Columbus, OH 43210 USA
jwdavis@cis.ohio-state.edu

Abstract. We present an approach for visual discrimination of children from adults in video using characteristic regularities present in their locomotion patterns. The framework employs computer vision to analyze correlated, scale invariant locomotion properties for classifying different styles of walking. Male and female subjects for the experiments include six children (3–5 yrs) and nine adults (30–52 yrs). For the analysis, we coordinate a minimalist point-representation of the human body with a space-time analysis of head and ankle trajectories to characterize the modality. Together the properties of relative stride length and stride frequency are shown to clearly differentiate children from adult walkers. The highly correlated log-linear relationships for the stride properties are exploited to reduce the categorization problem to a linear discrimination task. Using a trained two-class linear perceptron, we were able to achieve a correct classification rate of 93–95% on our dataset. Our approach emphasizing the natural modal behavior in human motion offers a useful and general methodology as the basis for designing efficient motion recognition systems using limited visual features.

1 Introduction

We can easily perceive a person walking in the distance simply from viewing the characteristic pattern of human motion. We can even identify a close friend from the way he or she walks. Remote-sensing computer monitoring and security systems designed to visually interpret our moving world will also require similar abilities to recognize human movements. In this paper we describe a categorical recognition system that addresses one of the most fundamental classes associated with human movement — locomotion. Human locomotion is subject to a variety of physical and dynamical constraints, which together create a tight region – or "mode" – in some multi-dimensional feature space. Our belief is that such correlated *visual* properties in human locomotion can be used within computer vision systems for the reliable classification of walking people. In support of this claim, we describe a categorical vision approach for distinguishing children from adults using the inherent modal nature associated with the properties of relative stride length and stride frequency during locomotion.

Much research has appeared in the computer vision literature pertaining to people walking, including several model-based tracking approaches [8, 1, 14, 3], static-based pedestrian detection methods [12, 5], and trajectory-based recognition systems [10, 11, 4]. Our approach differs from the aforementioned recognition methods in that we seek dynamical correlations between movement features to infer *categories* of people (e.g. child, adult) from their walking motions. The applied significance of the child-adult recognition paradigm impacts those automated visual surveillance and monitoring systems interested in identifying child and adult behaviors. Monitoring systems in locations such as shopping malls and airports place special importance on the detection of a lone or wandering child. Also, a smart car able to detect crossing pedestrians [5] would have a further safety advantage to avoid accidents if the car were able to have a specific awareness of children walking nearby.

2 Human Locomotion Features

Measurements of time and distance for each walking cycle represent the most basic descriptions that determine a particular gait [9]. Such parameters include stride length (distance between footfalls of the same foot), stature (body height), and cycle time (time of one complete cycle of one leg).

A commonly used variable in describing locomotion is *relative stride* L', calculated by normalizing the person's stride length by stature. The result is a dimensionless number with no issue of inches or pixels, and thus it can be used to compare the relative spatial configuration of children and adults. Another descriptive temporal feature is *stride frequency* f (strides/min), computed from the inverse of the cycle time. We list the locomotion features in Table 1 for convenience. These stride properties are visual features and can therefore be used, in part, by computer vision systems for the analysis of locomotion. As we will show, two distinctive modal relationships for the conjunction of the stride-based properties can be used to classify child and adult locomotion.

Description	**Symbol**	**Units**	**Obtained**
Stride length	L	pixels or inches	measured
Stature	S	pixels or inches	measured
Leg cycle time	T_c	sec.	measured
Relative stride	L'	–	L/S
Stride frequency	f	min.$^{-1}$	$60/T_c$

Table 1. Fundamental locomotion features.

3 Methods

Male and female subjects having normal gaits used for the experiments included nine adults 30–52 years old and six children 3–5 years old. This particular age range for children was motivated by the reported biomechanical difference in the walking style of children 3–5 years old as compared with adolescents and adults[1].

To calculate the proposed stride-based features (L', f) for the walkers, only the locations of the head and feet in each video frame are required. These extremity points are much more attainable than either joint angles or limb lengths/poses from images. We opted to initially record subjects walking while adorned with reflective markers on the head and ankles for the experiments. We are currently examining automatic methods (similar to [11, 7, 15, 12, 5]) for identifying these body locations from multiple viewpoints in un-marked video.

The imaging configuration positioned an infrared video camcorder (Sony CCD-TRV87) at a distance of 12.5 feet fronto-parallel to the walker (at 3-foot elevation for adults and 2-foot elevation for children). The adult subjects were recorded walking on a motorized treadmill (See Fig. 1.a) at speeds ranging from 1.5–4.5 MPH, increasing in 0.2 MPH increments (individuals had different maximum speeds). Treadmill and overground walking strides are not significantly different for the major range of the walking speeds [2]. The six child subjects were recorded walking at different speeds across a room in front of the camera (See Fig. 1.b). Due to the nature of using young children as experimental subjects, only those natural-looking walking sequences were retained. The trials were recorded onto a VCR, and later digitized and de-interlaced to achieve a faster frame-rate of 60 Hz at 320×240 resolution.

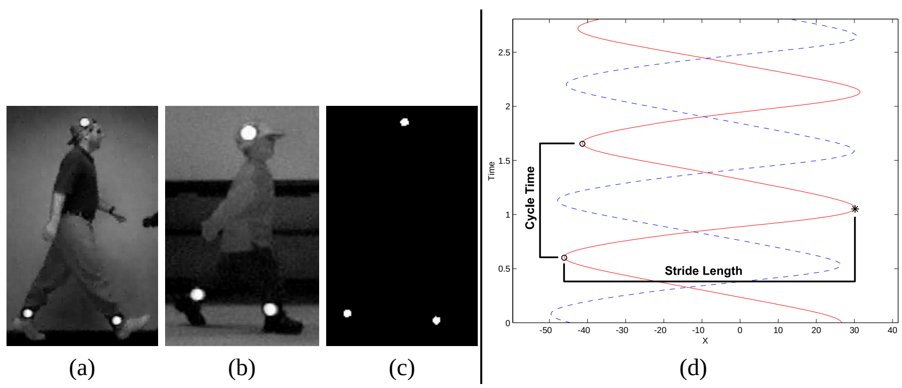

Fig. 1. Walkers and point-light motions. (a) Adult walking on a treadmill at 4.5 MPH. (b) Child walking across a room. (c) Thresholding image in (a) produces point-lights to be tracked. (d) Ankle x-trajectories with stride length and cycle time noted.

[1] For children 3–5 years old, their leg swing time remains relatively constant over various walking speeds, while for adults the swing time is negatively correlated with increasing walking speed [6].

3.1 Automatic feature extraction

We first threshold the images in each walking sequence to highlight the reflective markers on the head and ankles (See Fig. 1.c). After using a region-growing algorithm to identify the three centroids, trajectories were automatically created using the point tracking and correspondence approach of [13] which minimizes a *proximal uniformity function* δ for points q, r using three frames X^{t-1}, X^t, X^{t+1} with

$$\delta_{qr}^{t+1}(X) = \frac{\| \overline{X_q^{t-1} X_q^t} - \overline{X_q^t X_r^{t+1}} \|}{\sum_i \sum_j \| \overline{X_i^{t-1} X_i^t} - \overline{X_i^t X_j^{t+1}} \|} \quad + \quad \frac{\| \overline{X_q^t X_r^{t+1}} \|}{\sum_i \sum_j \| \overline{X_i^t X_j^{t+1}} \|} \tag{1}$$

where the first and second terms represent smoothness and proximity constraints, respectively. We then removed the translation component in the resulting trajectories using the horizontal location of the head point as a reference in each frame, followed by lowpass filtering.

To calculate the relative stride ($L' = L/S$) for rightward walking (the method is similar for leftward walking), the stride length L is computed as the pixel distance between consecutive minima and maxima locations in an ankle x-trajectory (See Fig. 1.d); the image stature S is calculated as the pixel distance from the head point to the ankle point of the support leg at criss-crossing locations. The cycle time for the stride frequency ($f = 60/T_c$) can be determined by measuring the time between two consecutive minima in an ankle x-trajectory (See Fig. 1.d). These features were automatically extracted for each step cycle for each ankle of the walkers in our dataset. With these features, we can now compare the walking styles of the children and adults.

4 Results

The computed ranges of stride frequencies for the children and adults were 55.3–89.8 strides/min and 36.7–73.3 strides/min, respectively. The relative strides for the children and adults were in the range 0.27–0.55, hence they share the same relative extent of spatial stride configurations. In Fig. 2.a we present a log-linear plot of relative stride vs. stride frequency as calculated from the children and adults walking at different speeds. Sub-spaces (modes) of child and adult locomotion are quite apparent in the data. A general interpretation of this plot is that whenever a child has the same (or larger) relative stride configuration as an adult, the child has a larger stride frequency.

With such strong and definitive modal sub-spaces, a simple linear decision boundary is all that is required to separate the two classes to a high degree. The data appears non-Gaussian, hence for classifying categories c1 and c2 we used a two-class linear perceptron discriminator having the general form

$$d(\mathbf{x}) = \sum_1^n w_i x_i + w_{n+1} = 0, \quad \text{with} \quad \sum_1^n w_i x_i \underset{\text{c2}}{\overset{\text{c1}}{\gtrless}} - w_{n+1}. \tag{2}$$

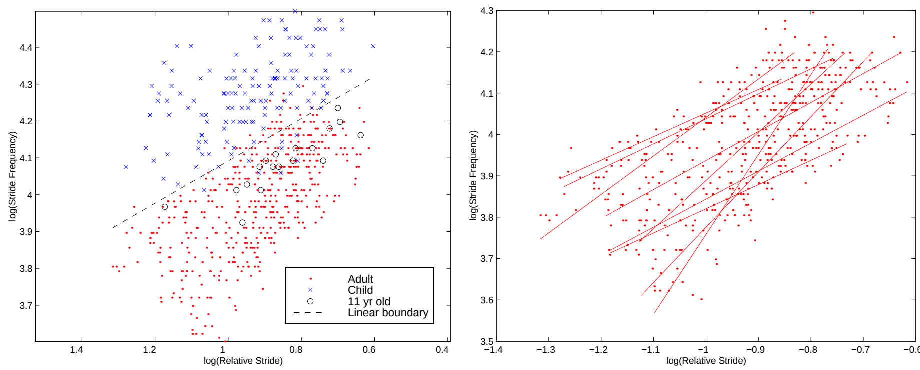

Fig. 2. (a) Walking modes for relative stride vs. stride frequency in children and adults (natural log values). A linear classification boundary is shown separating the two categories. The walking motions of an 11 year old reside entirely in the adult category. (b) Linear correlations for adult data points in (a), where each line represents an individual adult ($\bar{\rho} = 0.9232 \pm 0.0287$ SD).

A perceptron neural network was trained with all the walking examples resulting in a network having output O according to

$$O = \begin{cases} Adult & \text{if } 4376.6 \cdot \ln(L') - 7624.1 \cdot \ln(f) > -35571.8 \\ Child & \text{if } 4376.6 \cdot \ln(L') - 7624.1 \cdot \ln(f) < -35571.8 \end{cases} \tag{3}$$

determined after 30K epochs using a Matlab Neural Network Toolbox implementation (See boundary in Fig. 2.a). When the dataset is classified using this discriminator, we receive 95% correct classification for the adults and 93% correct classification for the children. Three older children 5–6 years old were also tested and shown to have more motions associated with the adult category (∼30%), suggesting the beginnings of a change in walking style. When an 11 year old was examined, his motions existed entirely within the adult category of locomotion (See Fig. 2.a).

The varying slope and positioning of the linear correlations for each adult over various speeds, as shown in Fig. 2.b, clearly show that there is no simple relationship of relative stride and stride frequency to stature. As individual mode *lines* are more unique than single points (many lines intersect), we could however use the individual modes to assess whether an observed motion is *consistent* with a hypothesized identity. To help confirm identity, one could calculate the stride properties for the person and compute the distance of that point to the mode of the proposed individual i using $D_i = |\alpha_i \ln(\hat{L}') - \ln(\hat{f}) + \beta_i| / \sqrt{\alpha_i^2 + 1}$ and verify that the motion is within tolerance. Though it is unlikely that stride-based features alone can be used for person identification, they may however be applicable to person authentication.

5 Summary

We presented an approach for the visual discrimination of children (3–5 years old) from adults using stride-based properties of their walking style. Trajectories of marked head and ankle positions for six children and nine adults were used to compute the relative stride and stride frequency for each walker at different speeds. The distinction between child and adult for these features is quite strong and reduces the task of categorization to a linear discrimination test. Using a trained two-class linear perceptron, we were able to achieve a correct classification rate of 93–95% for our dataset. Given that only two motion features were used to characterize and differentiate children from adults, the result is quite encouraging. The use of natural modes as a means of visual categorization provides a useful bottom-up framework for the classification and recognition of humans in motion.

References

1. A. Baumberg and D. Hogg. Learning flexible models from image sequences. In *Proc. European Conf. Comp. Vis.*, pages 299–308, 1994.
2. W. Boda, W. Tapp, and T. Findley. Biomechanical comparison of treadmill and overground walking. In *Proc. Can. Soc. for Biomech.*, pages 88–89, 1994.
3. C. Bregler and J. Malik. Tracking people with twists and exponential maps. In *Proc. Comp. Vis. and Pattern Rec.*, pages 8–15, 1998.
4. I. Chang and C. Huang. The model-based human body motion analysis system. *Image and Vision Comp.*, 18(14):1067–1083, 2000.
5. D. Gavrila. Pedestrian detection from a moving vehicle. In *Proc. European Conf. Comp. Vis.*, pages 37–49, 2000.
6. D. Grieve and R. Gear. The relationship between length of stride, step frequency, time of swing and speed of walking for children and adults. *Ergonomics*, 5(9):379–399, 1966.
7. 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.
8. D. Hogg. Model-based vision: a program to see a walking person. *Image and Vision Comp.*, 1(1):5–20, 1983.
9. V. Inman, H. Ralston, and F. Todd. *Human Walking*. Williams & Wilkins, Baltimore, 1981.
10. J. Little and J. Boyd. Describing motion for recognition. In *Proc. Symp. Comp. Vis.*, pages 235–240. IEEE, 1995.
11. S. Niyogi and E. Adelson. Analyzing and recognizing walking figures in XYT. In *Proc. Comp. Vis. and Pattern Rec.*, pages 469–474. IEEE, 1994.
12. M. Oren, C. Papageorgiour, P. Sinha, E. Osuma, and T. Poggio. Pedestrian detection using wavelet templates. In *Proc. Comp. Vis. and Pattern Rec.*, pages 193–99. IEEE, 1997.
13. K. Rangarajan and M. Shah. Establishing motion correspondence. *Comp. Vis., Graph., and Img. Proc.*, 54(1):56–73, 1991.
14. K. Rohr. Towards model-based recognition of human movements in image sequences. *Comp. Vis., Graph., and Img. Proc.*, 59(1):94–115, 1994.
15. C. Wren, A. Azarbayejani, T. Darrell, and A. Pentland. Pfinder: real-time tracking of the human body. *IEEE Trans. Patt. Analy. and Mach. Intell.*, 19(7):780–785, 1997.