

Person Identification using Shadow Analysis

Yumi Iwashita¹
yumi@ait.kyushu-u.ac.jp

Adrian Stoica²
adrian.stoica@jpl.nasa.gov

Ryo Kurazume¹
kurazume@ait.kyushu-u.ac.jp

¹ Information Science and Electrical
Engineering
Kyushu University
Fukuoka, Japan

² Jet Propulsion Laboratory
California Institute of Technology
Pasadena, USA

Abstract

We introduce a novel person identification method for a surveillance system of much wider area than conventional systems using CCTV cameras. In the proposed system, we install cameras to rooftops of buildings or a low altitude airship, and identify people by gait features extracted from shadows, which are projected on the ground by the sun in the daytime or lights in the evening. Since conventional systems extract gait features from actual body area, the correct classification ratio is reduced due to the lack of information of body area, in case that images are captured by overhead cameras. On the other hand, the proposed system enables to identify people by gait features which are extracted from shadows projected on the ground, even if images are captured by overhead cameras. In the proposed system, shadow areas projected on the ground are extracted automatically from captured images, and then analyze dynamics of shadow areas by the spherical harmonics. Experiments of person identification using actual outside images revealed that the proposed method showed the best performance than conventional methods, and the results indicate the feasibility of person identification based on shadow analysis.

1 Introduction

People identification has numerous applications. At one end of the spectrum are surveillance applications for intelligence and security operations, while at the other end are personal robots that need to recognize their owners and users. Identification is based on classification using human biometrics features, such as face or gait. While traditional biometrics rely on direct observation (e.g. image of the face/body), it may be the case that a projection may have more information than the direct signal. For example, the shadow of a person observed from a higher point, or from overhead, as from the airship. The exploitation of biometrics information in human shadow silhouettes (shadow biometrics) has been proposed. Centimeter-level resolution of airborne sensing systems enables people identification from high altitudes. Although two individuals seen from above may appear indistinguishable, in particular if wearing similar head covers and robes, their shadow silhouette often offers a larger area and almost certainly more accurate body movement details head and shoulder

top view. Figure 1 shows an image above a city; what appears to be the shape of a human body is in fact the shape of its shadow, a body projection. Shadows offer the connecting link between aerial observation and gait/biometrics/gesture and behavior classification. Shadow biometrics technology [13] (simply defined as biometrics using information from shadows) opens a new research area of applications in "overhead biometrics", which includes the remote observations from airborne/space platforms of biometric characteristics.

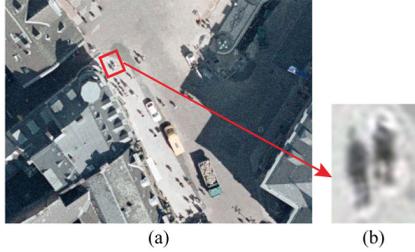


Figure 1: (a) Remote sensing imagery (Google [14]), (b) magnification of smaller window. What looks like humans are shadows.

To identify people based on their shadow, gait may be used to extract human characteristics. Although there is a large diversity of gait recognition algorithms, a majority have focused on canonical (side) viewing point using silhouettes for human detection or identification, with several public databases available. Silhouette gait recognition approaches generally fall into two main categories [15]: (1) model-based structural analysis, and (2) appearance-based analysis. Structural model-based approaches include parameterization of gait dynamics, such as stride length, cadence, and stride speed [6] [5] [6]. Static body parameters, such as the ratio of sizes of various body parts, can be considered in conjunction with these parameters. Traditionally, these approaches have not reported high performances on common databases, partly due to their need for 3D calibration information and the self occlusion caused by legs and arms crossing.

Appearance-based analysis uses measurements of spatio-temporal features of the silhouette [2] [9] [16] [17]. To characterize shape and its variations different measures may including size (width, height, area), angles between lines (e.g. foot/ankle, upper arm-lower arm), higher order moments around the centroid, measures of symmetry, or other shape representations, and temporal variations such as cyclic oscillations at the stride frequency. Shape-based approaches have been shown very effective for human silhouette detection and have been used with good results on human identification. Affine moment invariants-based method [8] and 1D Fourier Transform-based method [14] revealed their strong discrimination power on common database. However, in case that a camera is placed on a high place, such as on a rooftop of a tall building and an airship, the correct classification ratio degrades due to the reduced image resolution.

To identify people based on shadow biometrics, a method based on gait is introduced in [9], and the feasibility of people identification based on shadow biometrics was shown. However, in this method shadow areas were extracted manually and comparison with this method and conventional methods based on actual body areas were not shown. So in this paper, we introduce a method to extract shadow area automatically, and present details of the person identification method. In experiments we show that the proposed method is robust to the change of image resolution, and keep the high correct classification ratio even if the

image resolution degrades. Moreover, we compare the proposed method with conventional methods and then show the effectiveness of the proposed method.

This paper is organized as follows. Section 2 introduces the new shadow extraction method and the details of the method for person identification based on shadow biometrics. Section 3 describes the database we shall use and the experiments performed. Conclusions are presented in section 4.

2 Methodology

This section describes a method for shadow extraction and the methodology for person identification. To summarize, the main steps of processing are: (1) background subtraction and segmentation of people and shadows, (2) extract gait features, (3) analyze dynamics of features and classification. The following offers more details.

2.1 Segmentation

To extract shadow silhouettes from captured images, firstly a background subtraction [15] is used. This method is based on the probability density function estimated by Parzen density estimation and the evaluation of the local texture at pixel-level resolution. The method performs robust object detection under varying illuminations. Figure 2 (b) shows an example of extracted silhouettes, from the original image shown in Fig. 2 (a).

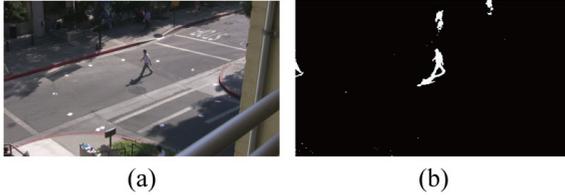


Figure 2: (a) A sample image, (b) its extracted target region.

Then we separate body area and shadow area automatically. To separate these areas colour information may be used, but it would be difficult in case that the person wears clothes whose colours are similar to one's shadow. So we separate body area and shadow area from silhouette images directly. The details of this method is described below.

In the first step, an average image $I_{average}(x, y)$ is created from sequential silhouette images according to $I_{average}(x, y) = \frac{1}{T} \sum_{t=1}^T I(x, y, t)$, where T is the number of frames. Figure 3 (a1) and (a2) show examples of silhouette images, and Fig. 3 (b) shows an example of average images. Then extract an axis S_1 along which the variance of the average image is the smallest as shown in Fig.3 (b), and calculate the histogram of the average image by adding pixel values along the axis S_2 perpendicular to the axis S_1 (Fig.3 (d)). Here, since foot touching the ground remains in the same pixel of images for several frames, so the value of this pixel in the average image is bigger than other pixels. Moreover, the shadow can be considered as the image captured from the position of the sun, so in general the body area overlaps legs in the shadow area. So the histogram value in the area between foot and body is bigger than that in foot area. On the other hand, since the moving speed of shin is fast and it's width is narrow compared with other body area, so the histogram value in this area is minimum

locally. The coordinate values of these two areas where the histogram values are maximum and local minimum are named Y_a and Y_b , respectively. So the foot area can be considered as the area between Y_a and Y_b as shown in Fig.4 (a), and the body area and the shadow area are separated by $Y' = \frac{(Y_a+Y_b)}{2}$. To separate shadow area and body area, silhouette images are converted to images on S_1 - S_2 coordinate, and then silhouettes are separated by the value Y' . Examples of the separated areas are shown in Fig.4. After extraction of shadow area, the shadow area is scaled to a uniform height, set to 100 pixels.

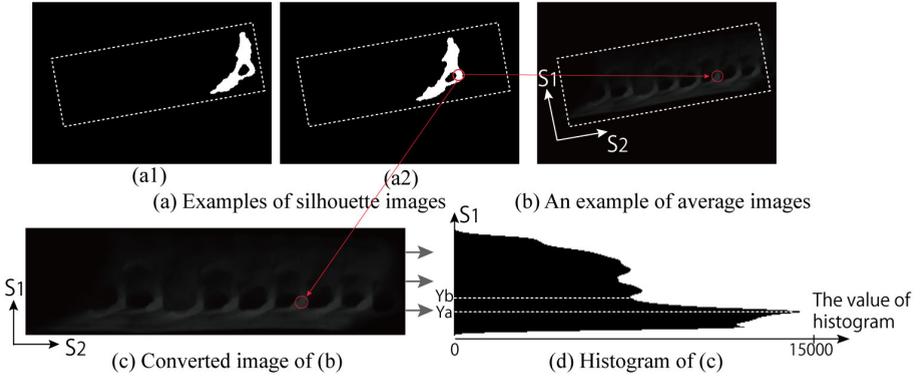


Figure 3: Calculation of the histogram.

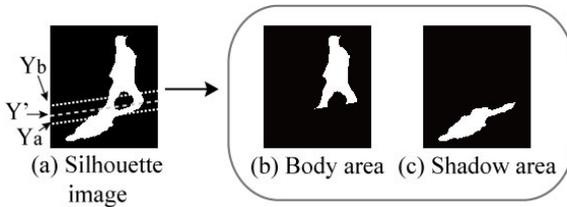


Figure 4: Examples of body and shadow separation.

2.2 Extraction of gait features

After extracting image sequences of shadow silhouettes of one gait cycle, which is used to partition the sequences, dynamic features of gait sequence are obtained as follows: in each row of the image $I(x, y, t)$ the largest distance $H(y, t)$, which we name here "gait stripe", is determined as the absolute value of the distance between extremities of this shape, illustrated in Fig. 5 (a) as a horizontal stripe measured from left to right.

2.2.1 Frequency analysis

Since the gait motion repeats in time, the gait stripe $H(y, t)$ is also repetitive as shown in Fig. 5 (b). We decided to apply spherical harmonics transforms [2] because of its robustness to noise. Spherical harmonic transforms are essentially Fourier transforms on the sphere. Spherical harmonic functions $\{Y_l^m(\theta, \phi) : |m| \leq l \in \mathcal{N}\}$ are special functions defined on the

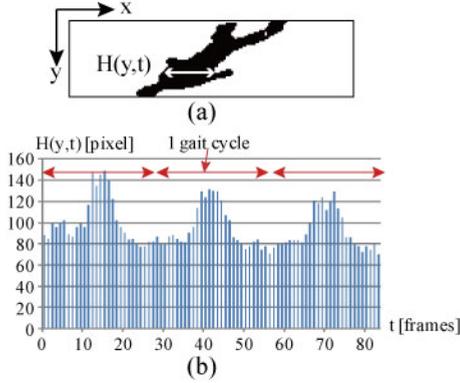


Figure 5: (a) Determination of $H(y,t)$, (b) An example of gait stripes ($H(y,t)$).

unit sphere S^2 [1] as:

$$Y_l^m(\theta, \phi) = \begin{cases} \sqrt{2}K_l^m \cos(m\phi) P_l^m(\cos\theta) & m > 0 \\ \sqrt{2}K_l^0 P_l^0(\cos\theta) & m = 0 \\ \sqrt{2}K_l^m \sin(-m\phi) P_l^{-m}(\cos\theta) & m < 0 \end{cases} \quad (1)$$

where $\theta \in [0, \pi]$, $\phi \in [0, 2\pi]$, K_l^m is a scaling factor defined $K_l^m = \sqrt{\frac{(2l+1)}{4\pi} \frac{(l-m)!}{(l+m)!}}$, and P_l^m is the associated Legendre polynomial. The spherical harmonic functions are projected into spherical harmonic coefficients c_l^m as follows:

$$c_l^m = \int_S f(\theta, \phi) Y_l^m(\theta, \phi) ds. \quad (2)$$

Here, f is an original function. The range of l is $0 \leq l < n$, and m is $|m| \leq l$, so the total number of coefficients c_l^m are n^2 .

For characterizing gait stripes, we obtain spherical harmonic coefficients of gait stripes $H(y,t)$ as features. To obtain spherical harmonic coefficients c_l^m of $H(y,t)$, we define the original function $f(\theta, \phi)$ as $f(\theta, \phi) = H(y,t)$ where $\theta = \frac{t}{T} \times \pi$, $\phi = \frac{y}{Y} \times 2\pi$. Here, T is the number of frames in one gait cycle and Y represents the height of images ($Y = 100$). Moreover, to reduce the influence of high frequency of gait stripes, the number n of coefficients for classification is less than N .

2.2.2 The use of both shadow area and body area

In case that cameras are positioned on an airship and images perpendicular to the ground are captured as shown in Fig. 1, gait features are extracted from shadow area. In case that cameras are positioned on an upper floor of a building and oblique images are captured as shown in Fig. 2 (a), gait features can be extracted from both shadow area and body area. The body area is captured from the real camera, and the shadow area can be considered as the area which are taken by a different camera placed at the light source. So when oblique image are captured, the correct classification ratio becomes high by extracting features both shadow area and body area.

2.3 Classification

In the proposed method, we use the nearest neighbour as the classifier. In the training phase, we extract spherical harmonic coefficients from training sequences, and then create database. Then in the identification phase, the same spherical harmonic coefficients are extracted from test data sets, and the subject is identified by the classifier.

2.4 Characteristics of the proposed method

The characteristics of the proposed method are as follows.

- 1 The proposed system enables to identify people by gait features which are extracted from shadows projected on the ground, even if images are captured by overhead cameras.
- 2 The proposed method can separate shadow area and body area from silhouette images automatically, even if the target wears clothes whose colours are similar to that of the shadow area.
- 3 Although the extracted target region may include noise, such as holes in the extracted region, the proposed method can extract features robustly against noise by the gait stripe, which is the distance between extremities.

3 Experiments

In this section, we describe the experiments. In our experiments, we used a shadow gait database collected in [9]. The shadow database was created based on imagery recorded with a video camera positioned on an upper floor of a building. The database contains raw image sequences, and Fig. 6 shows example entries from the database. It contains 20 video sequences, which contain 5 different subjects with 4 sequences for every subject. Moreover, one gait cycle is determined manually, and we used the leave-one-out cross validation to estimate the classification error rate.

We carried out three experiments as follows: (i) person identification with shadow area, (ii) person identification by conventional methods [14] [8] and 2D Fourier Transform, (iii) person identification with both shadow area and body area. The detail of each experiment is described in the following sections.



Figure 6: Example images of person walking, with shadow visible.

3.1 Person identification with shadow area

In the first experiment, we extracted gait features from shadow area extracted manually. We set the number of coefficients $n = 10$ (total number is 100), and Fig.7 shows examples of

calculated coefficients c_i^m of subject A and subject B. From these examples, we see that the coefficients of subject A and subject B show their own characteristics. Figure 8 represents the correct classification rate (CCR) with respect to order of coefficients. The CCR exceeded 95.0 % for more than 49 coefficients ($n=7$), so in the following experiments, we used 49 coefficients.

In the next experiment, we extracted gait features from shadow area extracted automatically by the proposed method, and the CCR was 90 %, which was also a high score. In the following experiments, to evaluate the performance of the person identification method, we separated shadow area and body area manually.

3.2 Person identification by conventional methods and 2D Fourier Transform

To compare the proposed method with conventional methods, we applied a method based on 1D Fourier Transform (1D DFT) [14], and a method based on affine moment invariants (AMIs) [8] to shadow gait images. The results are shown in Table 1 (1) and (2). From these results, the CCR by the proposed method (Table 1 (4)) is the highest compared with those by both conventional methods.

Moreover, we explored the degradation in classification with the reduction in spatial and temporal resolution of the images, which allows us to tailor an observational platform to the needed/sufficient classification ratio. In this experiment, the classifier is trained by using the training data sets of full resolution. For the test data sets, we decreased the spatial and temporal resolution to half, respectively. We used the proposed method, the 1D DFT-based method [14], and the AMIs-based method [8] to extract features, and Table 2 shows the result of classification. From these results, the degree of the features based on the proposed method is much less than that based on the 1D DFT-based method, and the CCR of the proposed method is almost the same with that of the 1D DFT-based method. On the other hand, the CCR based on AMIs-based method degraded, and this may be because image resolution was not enough to calculate the AMIs having sufficient discrimination power.

In the proposed method we used the spherical harmonics to analyze periodicity of gait stripes, but instead of it the Fourier Transform can be used. So we applied the 2D Fourier Transform to the gait stripe $H(y, t)$. Here, since the number of frames in one gait cycle (T) is different, it is scaled to a uniform number, set to 50. Moreover, since in high frequency the image converted to frequency domain includes noise, so we excluded high frequency for classification. The result of classification was 65 % as shown in Table 1 (3). The reason why the proposed method performed better than the 2D DFT-based method may be as follows. Each basis functions of the Fourier Transform is sine wave whose direction is along one dimension. On the other hand, each basis function of the spherical harmonics is 2D wave whose direction is along two dimension, so the correlation between two axes may be extracted by the spherical harmonics.

3.3 Person identification with both shadow area and body area

Next experiment is to classify people with gait features extracted from body area. The CCR was 95 %, which is the same with the CCR of shadow area. Moreover, we used gait features extracted from both shadow area and body area, and the CCR was 100 %. As we described in Section 2.2.2, the shadow area and the body area can be considered as images captured from

different cameras placed at different positions. So in case that gait features can be extracted from both areas, the CCR becomes high by using features from both areas.

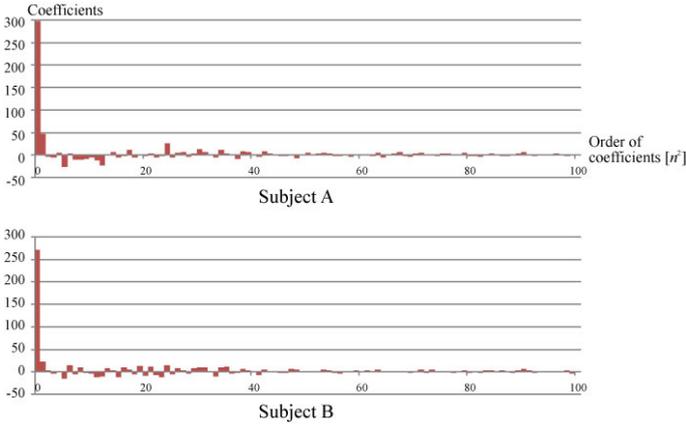


Figure 7: Examples of spherical harmonic coefficients.

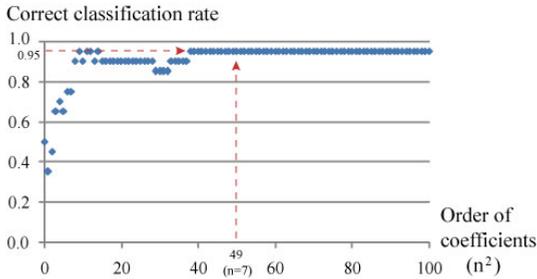


Figure 8: CCR [%] with respect to the order of the spherical harmonic coefficients.

4 Conclusion

This paper introduced a methodology for shadow biometrics and demonstrated gait recognition from shadow analysis. We introduced a method to separate shadow area and body area automatically, and determined a set of silhouette stripes. Spherical harmonics were applied to the stripes for each gait sequence, followed by the nearest neighbor classification to spherical harmonic coefficients. A correct classification rate (CCR) of 95.0 % from shadow area extracted manually and that of 90.0 % from shadow area extracted automatically were obtained. A reduction of resolution to 50 % reduced the CCR from 95.0 % to 80.0 %, which were the best performance compared with conventional methods. Two near term directions of work are: (1) implement the equations that provide a corrective scaling with different positions of the sun, (2) study the robustness to changed direction of walk.

Table 1: Comparison of the proposed method and conventional methods.

	The number of features	Correct classification rate [%]
(1) 1D DFT [14]	Image resolution $\times 3$ (120000 in case of 100 \times 400 pixel)	85
(2) Affine moment invariants [8]	6	70
(3) 2D DFT	Image resolution (In experiments the first 800 in low frequency)	65
(4) The proposed method (manual)	49	95
(5) The proposed method (automatic)	49	90

Table 2: Correct classification rate [%] of shadow silhouettes with respect to spatial and temporal resolution.

	The number of features	(Spatial resolution, temporal resolution) [%]			
		(100, 100)	(50, 100)	(100, 50)	(50, 50)
The proposed method (4)	49	95	75	85	80
1D DFT [14]	Image resolution $\times 3$ (120000 in case of 100 \times 400 pixel)	85	80	80	80
Affine moment invariants [8]	6	70	50	80	45

References

- [1] <http://www.google.com>.
- [2] J. Acquah, M. Nixon, and J. Carter. Automatic gait recognition by symmetry analysis. *Pattern Recognition Letters*, 24.
- [3] C. BenAbdelkader, R. Cutler, and L. Davis. Stride and cadence as a biometric in automatic person identification and verification. In *Proc. IEEE Conf. Face and Gesture Recognition*, 2002.
- [4] C. BenAbdelkader, R. Cutler, H. Nanda, and L. Davis. Eigengait: Motion-based recognition of people using image self-similarity. In *3rd Int. Conf. Audio- and Video-Based Biometric Person Authentication*, 2001.
- [5] I. Bouchrika and M. Nixon. People detection and recognition using gait for automated visual surveillance. In *Proc. IEE Inter. Symp. Imaging for Crime Detection and Prevention*, 2006.

- [6] D. Cunado, M. Nixon, and J. Carter. Automatic extraction and description of human gait models for recognition purposes. *CVIU*, 90(1):1–41, 2003.
- [7] R. Green. Spherical harmonic lighting: The gritty details. In *SCEA Research and Development*, 2003.
- [8] Y. Iwashita and R. Kurazume. Person identification from human walking sequences using affine moment invariants. In *Proc. IEEE Int. Conf. Robotics and Automation*, 2009.
- [9] Y. Iwashita and A. Stoica. Gait recognition using shadow analysis. In *Int. Symp. on Bio-inspired, Learning, and Intelligent Systems for Security*, 2009.
- [10] S. Lee and Collins R. Liu, Y. Shape variation-based frieze pattern for robust gait recognition. In *CVPR*, 2007.
- [11] Y. Liu, R. Collins, and Y. Tsin. Gait sequence analysis using frieze patterns. In *ECCV*, 2002.
- [12] M. Nixon, T. Tan, and R. Chellappar. Human identification based on gait. In *Springer-Verlag New York, Inc. Secaucus, NJ, USA*, 2006.
- [13] A. Stoica. Towards recognition of humans and their behaviors from space and airborne platforms: Extracting the information in the dynamics of human shadows. In *Proc. the 2008 Bio-inspired, Learning and Intelligent Systems for Security*, 2008.
- [14] K. Sugiura, Y. Makihara, and Y. Yagi. Gait identification based on multi-view observations using omnidirectional camera. In *Proc. 8th Asian Conference on Computer Vision*, 2007.
- [15] T. Tanaka, A. Shimada, D. Arita, and R. Taniguchi. Object detection under varying illumination based on adaptive background modeling considering spatial locality. In *Proc. the Third Pacific-Rim Symposium on Image and Video Technology*, 2009.