Gait identification from invisible shadows

Yumi Iwashita[a b], Koji Uchino[a], Ryo Kurazume[a] and Adrian Stoica[b]

[a]Kyushu University, 744 Motooka, Nishi-ku, Fukuoka, Japan;
[b]Jet Propulsion Laboratory, 4800 Oak Grove Drive, Pasadena, CA, USA

ABSTRACT

This paper introduces a person identification system that uses as input the shadow images of a walking person, as projected by multiple lights(in this application invisible/infrared lights); the system uses a database of examples of shadows images of a number of people who walk. While it is accepted that personal identification has a higher correct classification rate if views from multiple cameras are used, most systems use only one camera, mainly because (i) Installation in real-world environments is easier, less cameras and no need to synchronize cameras, (ii) Computational cost is reduced. In the proposed system, we obtain the advantages of multiple viewpoints with a single camera and additional light sources. More specific, we install multiple infrared lights to project shadows of a subject on the ground and a camera with an infrared transmitting filter mounted in the ceiling inside of a building. Shadow areas, which are projections of one's body on the ground by multiple lights, can be considered as body areas captured from different viewpoints; thus, the proposed system is able to capture multiple projections of the body from a single camera. We explored in other papers the use of sun-produced shadow for identification of people walking freely in the outdoor. In this paper the application scenario is a system installed at the airport in the areas that precedes the immigration checkpoint. Japan already has health monitoring cameras focused on approaching individuals, to determine their health condition; the here described system would also be installed in such a controlled area with restricted walk corridors of walk and controlled lighting. Gait is a *remote* biometrics and can provide early warning; on another hand it can be used as corroborating evidence in a multi-modal biometrics system. A database of images including shadows for a set of 28 walking people was collected, and the features extracted from shadow areas by affine moment invariants, after which identification of the subject followed. The experiments using the database show the effectiveness of the proposed method and further prove the superiority of using multiple viewpoints compared to a single viewpoint.

Keywords: Person identification, gait, shadow biometrics, infrared light

1. INTRODUCTION

Person identification systems have been used for a wide variety of applications, such as secure access into buildings, and is expanding further to new applications such as service robots that coexist with human and provide various personalized services in daily life. Gait is a powerful *remote* biometric, offering the advantages of identification from a distance, and of being unobtrusive, as body-invasive sensing is not needed to capture gait information.

Gait identification relies usually in one of the two approaches, using: (1) model-based analysis, or (2) appearance-based analysis. Model-based approaches include parameterization of gait dynamics, such as stride length, cadence, and joint angles[1,2,3] . Traditionally, these approaches have not reported high performances on common databases, partly due to their needs for 3D calibration information and self-occlusion caused by legs and arms crossing.

Appearance-based analysis[4,5,6] uses measurements of gait features from silhouettes by feature extraction methods, such as gait energy image(GEI)[7] , fourier transforms[8,9] , and affine moment invariants[10] . Appearance-based approaches have been used with good results on human identification. Correct classification rates in person identification are generally better when multiple cameras from different viewpoints are used, yet most of conventional methods have used one camera, because of (i) easy installation in real environments, less cameras and no need to synchronize cameras, (ii) a reduction of computation costs.

Y.I.: yumi@ait.kyushu-u.ac.jp, she is a visiting researcher in JPL.

Sensing Technologies for Global Health, Military Medicine, Disaster Response, and Environmental Monitoring II; and Biometric Technology for Human Identification IX, edited by Sárka O. Southern, et al., Proc. of SPIE Vol. 8371, 83711S
 · doi: 10.1117/12.919605

In past work, we proposed shadow biometrics which is a projected body area of a subject on the ground by either the sun in daytime or lights during the night[12].[13] In that research a camera was used to capture imagery from a high point in a building, and images of walking people, both shadow and body areas of subjects were acquired. Since the shadow area can be considered as the area taken from a viewpoint of the sun, That was equivalent to obtaining images from 2 different viewpoints with one camera.

In applications related to security in controlled spaces, such as at the airport corridors leading to an airport security passport control (Fig. 1), etc there is an interest to enhance the capability of person identification, for advanced warning (hence the need for a remote biometric) as well as for the purpose of achieving a multi-modal system for increased correct classification rate (i.e. working in addition to face biometrics, for example). In this case the idea is to use multiple lights to generate multiple shadow areas and to capture them with a single camera. (Multiple cameras have cost, placement and synchronization penalties). For being inconspicuous we decided to use invisible lights(infrared). Thus, in this paper, we propose a novel person identification system from shadow biometrics projected on the ground by infrared lights. In the proposed system, at first we install multiple infrared lights at high places, which project the subject's body area on the ground, and a camera with an infrared transmitting filter in the ceiling to capture all shadow areas. Next, we collect a database of (invisible to the human eye) shadows, for people walking along the corridor, and extract features from the shadow areas by affine moment invariants[10] , followed by people identification.

Shadow areas, which are projections of one's body on the ground by multiple lights, can be considered as body areas captured from different viewpoints, so the proposed system enables to capture multiple body areas from only one camera. The advantage of using infrared lights is that these lights cause less stress on subjects compared with normal visible lights; in addition it is an observation that is transparent to the subjects, and offers no reason for falsifying the gait. Figure 2 shows examples of captured images in which the subject's shadow was projected by infrared light. Figure 2 (a) shows an example of captured images with a visible light transmitting filter, which are equivalent to images people see, and Fig. 2 (b) shows an example of images with the infrared transmitting filter.

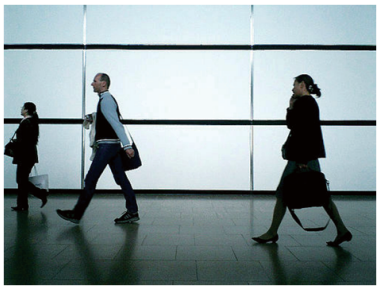
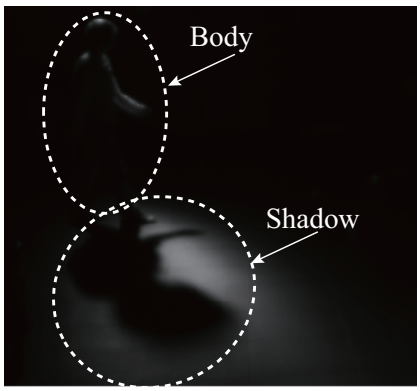


Figure 1. An example image at the airport corridors.[11]

(a) An example of captured images with a visible light transmitting filter



(b) An example of captured images with an infrared transmitting filter

Figure 2. Examples of captured shadows projected by infrared light, (a) the image without filter, which is equivalent to how people see, (b) the image with filter.

This paper is organized as follows. Section 2 describes the shadow database, invisible to people, and section 3 describes the details of the person identification method. Section 4 explains the experiments using the database. Conclusions are presented in section 5.

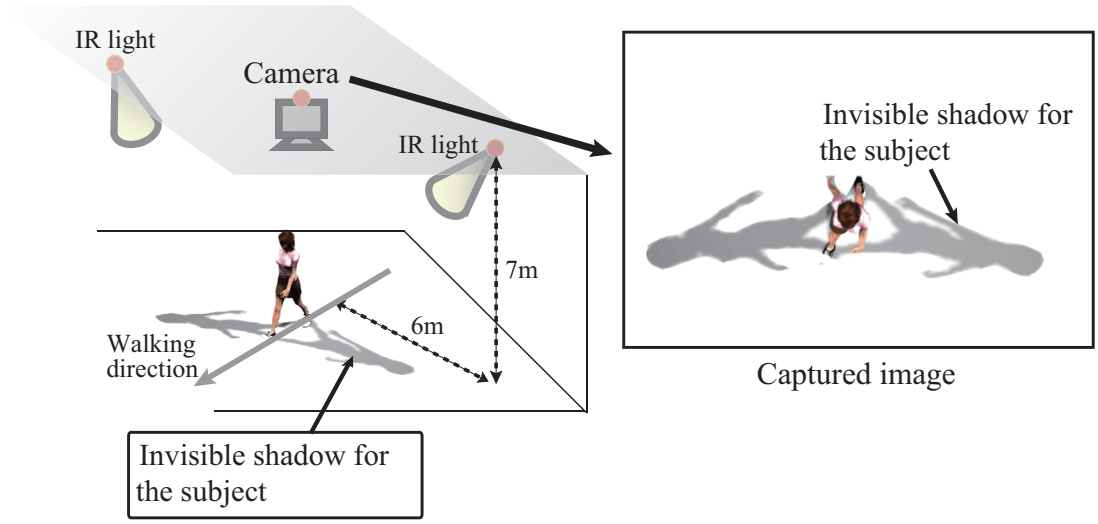


Figure 3. Experimental setting.

2. INVISIBLE SHADOW DATABASE

This section describes the shadow database. Two infrared lights (Bosch, IR Illuminator 850 nm, UFLED30-8BD) were placed on both sides, opposite, in respect to the subject, at positions (height 7 $[m]$ and depth 6 $[m]$), and a camera (PointGrey Research Inc., Grasshopper2 M/C) was placed in the ceiling, perpendicular to the ground as shown in Fig. 3. Areas for acceleration before image capturing area and for deceleration after that capture were defined (such that the subject has relatively constant speed during capture). Subjects contacted their right heel at a certain position in the image capturing area after their walking speed reached to their normal one. The image resolution was 1600×1200 and the frame rate was 30 Hz. This database records experiments with 28 subjects, each with 5 sequences of walk. Figure 4 shows examples of captured images.

3. PERSON IDENTIFICATION USING INVISIBLE SHADOW

This section explains the details of person identification.[10] To identify people, we extract gait features by the following steps: (1) an average image over a gait cycle is calculated and the subject's area is divided into multiple areas, and (2) gait features are calculated by affine moment invariants. Finally, a subject is identified by comparing his features, with the ones in the database.

3.1 Definition of average image and division of subject's area

Figure 5 shows examples of silhouette images of Fig. 4, which are threshold-segmented foreground images, after separation from the background. The silhouette area was then scaled to a uniform height, which was set to 512 pixels, and was aligned such that the position of the center of the silhouette area along the x-axis coincided with the center position of the image along the x-axis. The average image from aligned images of one gait cycle is defined as follows:

$$I^{average}(x,y) = \frac{1}{T}\sum_{t=1}^{T} I(x,y,t), \tag{1}$$

where T is the number of frames in one gait cycle and $I(x,y,t)$ represents the intensity of the pixel (x,y) at time t. Here, one gait cycle is a fundamental unit to describe the gait during ambulation, which is defined as an interval from the time when the heel of one foot strikes the ground to the time at which the same foot contacts the ground again. Figure 6 shows an example of average images.

The human body area was then divided into K areas equally according to the height to extract gait features not only from the whole area, but also from separated local areas.

Figure 4. Examples of captured images.

Figure 5. Silhouette images of Fig. 4.

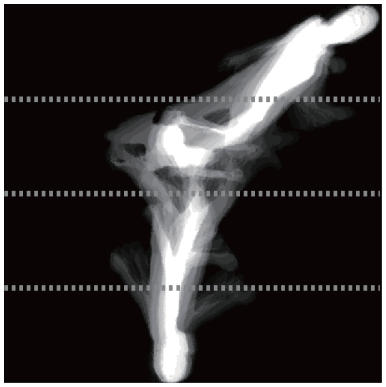


Figure 6. An example of average images (K=4).

3.2 Affine moment invariants

This section provides details of the use of 2D affine moment invariants. Affine moment invariants are moment-based descriptors, which are invariant under a general affine transform. The moments describe shape properties of an object as it appears. For an image the centralized moment of order $(p+q)$ of an object O is given by

$$\mu_{pq} = \sum\sum_{(x,y)\in O}(x-x_g)^p(y-y_g)^q I(x,y). \tag{2}$$

Here, x_g and y_g are the center of the object. In our method we used twenty two affine moment invariants $\boldsymbol{I} = \{I_1, I_2, \ldots, I_{22}\}$, and we show six of them[14,15] .

$$\begin{aligned}
I_1 &= \frac{1}{\mu_{00}^4}(\mu_{20}\mu_{02} - \mu_{11}^2) \\
I_2 &= \frac{1}{\mu_{00}^{10}}(\mu_{30}^2\mu_{03}^2 - 6\mu_{30}\mu_{21}\mu_{12}\mu_{03} + 4\mu_{30}\mu_{12}^3 + 4\mu_{03}\mu_{21}^3 - 3\mu_{21}^2\mu_{12}^2) \\
I_3 &= \frac{1}{\mu_{00}^7}(\mu_{20}(\mu_{21}\mu_{03} - \mu_{12}^2) - \mu_{11}(\mu_{30}\mu_{03} - \mu_{21}\mu_{12}) + \mu_{02}(\mu_{30}\mu_{12} - \mu_{21}^2)) \\
I_4 &= \frac{1}{\mu_{00}^{11}}(\mu_{20}^3\mu_{03}^2 - 6\mu_{20}^2\mu_{11}\mu_{12}\mu_{03} - 6\mu_{20}^2\mu_{02}\mu_{21}\mu_{03} + 9\mu_{20}^2\mu_{02}\mu_{12}^2 + 12\mu_{20}\mu_{11}^2\mu_{21}\mu_{03} \\
&\quad +6\mu_{20}\mu_{11}\mu_{02}\mu_{30}\mu_{03} - 18\mu_{20}\mu_{11}\mu_{02}\mu_{21}\mu_{12} - 8\mu_{11}^3\mu_{30}\mu_{03} - 6\mu_{20}\mu_{02}^2\mu_{30}\mu_{12} + 9\mu_{20}\mu_{02}^2\mu_{21}^2 \\
&\quad +12\mu_{11}^2\mu_{02}\mu_{30}\mu_{12} - 6\mu_{11}\mu_{02}^2\mu_{30}\mu_{21} + \mu_{02}^3\mu_{30}^2) \\
I_5 &= \frac{1}{\mu_{00}^6}(\mu_{40}\mu_{04} - 4\mu_{31}\mu_{13} + 3\mu_{22}^2) \\
I_6 &= \frac{1}{\mu_{00}^9}(\mu_{40}\mu_{04}\mu_{22} + 2\mu_{31}\mu_{22}\mu_{13} - \mu_{40}\mu_{13}^2 - \mu_{04}\mu_{31}^2 - \mu_{22}^3)
\end{aligned} \tag{3}$$

3.3 Person identification

The nearest neighbor method was chosen as classifier. In the training phase, gait features were extracted from training sequences by affine moment invariants, and a database was built. During the identification phase, gait features are extracted in the same way as in the training phase. Features were extracted from each area, and features of all areas of a sequence were fused by feature vector concatenation. Finally, the person was identified by the classifier.

4. EXPERIMENTS

This section shows the results of person identification experiments using the shadow database in Section 2. The database contains 28 people with 5 sequences for every subject. The correct classification rate was estimated with the leave-one-out cross validation. Here, in the leave-one-out cross validation a sequence from the database is selected for a test dataset and the rest of sequences in the database are used as training datasets, and then the subject in the test dataset is identified. This process is repeated for all combination of test datasets and training datasets.

At first, we used gait features from whole area (K=1), and the correct classification rate (CCR) was 87.1 %. Next, we changed the division number to K=2, and the CCRs of upper area and lower area were 90.7 % and 78.6 %, respectively. The reason why the CCR of upper area is better than that of lower area may be: the upper area includes both the subject's body area and one's shadow area and the lower area includes only one's shadow area as shown in Fig. 6, so the discrimination capability of the upper area is better than that of the lower area. Then, we combined both upper and lower areas, and the CCR increased to 94.3 %.

In further tests the division number was increased to K=4 and K=8. Table 1 shows correct classification rates with respect to the change of division number. From these results, the correct classification rate of K=2 showed the highest score, while the correct classification rates of K=4 and K=8 are worse. The reason may

be the fact that: the silhouette images include noise and deficits, and these may cause alignment error in the phase of calculating average images. Hence, a near term direction of work is to implement a method for feature extraction robust to noise and deficits based on frequency analysis.[13]

Table 1. The correct classification rates (CCRs).

The division number	1	2	4	8
CCR [%]	87.1	94.3	86.4	87.1

5. CONCLUSION

We proposed a system for person identification from shadow images of a walking person projected by invisible artificial lights. We built a shadow database of walking people. An experimental set-up was built, in which multiple infrared lights were installed to project shadows of a walking subject on the ground, and a camera with an infrared transmitting filter installed in the ceiling, The proposed system enables to capture multiple body areas from only one camera. Experiments with the database showed the effectiveness of the proposed system. Our preliminary results had a 94.3% CCR on a database of 28 people. Future work will focus on developing a shadow-based gait identification method robust to appearance changes due to variations of clothes and carried objects.

REFERENCES

[1] I. Bouchrika & and M. Nixon: People Detection and Recognition using Gait for Automated Visual Surveillance, IEE Inter. Symp. Imaging for Crime Detection and Prevention, 2006.

[2] D. Cunado, M. Nixon & J. Carter: Automatic Extraction and Description of Human Gait Models for Recognition Purposes, Computer Vision and Image Understanding, vol. 90, no. 1, pp.1–41, 2003.

[3] C. Yam, M. Nixon & J. Carter: Automated person recognition by walking and running via model-based approaches, Pattern Recognition, vol. 37, no. 5, pp.1057–1072, 2004.

[4] C. BenAbdelkader, R. Cutler, H. Nanda & L. Davis: EigenGait: Motion-based Recognition of People using Image Self-Similarity, Int. Conf. Audio- and Video-Based Biometric Person Authentication, 2001.

[5] S. Lee, Y. Liu & R. Collins: Shape Variation-based Frieze Pattern for Robust Gait Recognition, Computer Vision and Pattern Recognition, pp.1–8, 2007.

[6] Y. Liu, R. Collins & Y. Tsin: Gait Sequence Analysis using Frieze Patterns, European Conf. Computer Vision, pp.657–671, 2002.

[7] J. Han & B. Bhanu: Individual Recognition Using Gait Energy Image, IEEE Trans. PAMI, vol. 28, no. 2, pp. 316-322, 2006.

[8] J. Acquah, M. Nixon & J. Carter: Automatic gait recognition by symmetry analysis, Pattern Recognition Letters, 24, pp.2175–2183, 2003.

[9] K. Sugiura, Y. Makihara & Y. Yagi: Gait Identification based on Multi-view Observations using Omnidirectional Camera, Asian Conf. on Computer Vision, Vol.1, pp.452–461, 2007.

[10] Y. Iwashita & R. Kurazume: Person identification from human walking sequences using affine moment invariants, Proc. IEEE Int. Conf. Robotics and Automation, pp.436-441, 2009.

[11] http://flickrhivemind.net/Tags/fuhlsbuettel/Interesting

[12] A. Stoica: Towards Recognition of Humans and their behaviors from Space and Airborne Platforms: Extracting the Information in the Dynamics of Human Shadows, Bio-inspired, Learning and Intelligent Systems for Security, 2008.

[13] Y. Iwashita, A. Stoica & R. Kurazume: Person Identification using Shadow Analysis, British Machine Vision Conference, pp.35.1–10, 2010.

[14] J. Flusser & T. Suk: Pattern recognition by affine moment invariants, Pattern Recognition, vol.26, no.1, pp.167–174, 1993.

[15] J. Flusser, T. Suk & B. Zitova: Moments and Moment Invariants in Pattern Recognition, Wiley & Sons Ltd., 2009.