Gait identification using invisible shadows: robustness to appearance changes

Yumi Iwashita †, Ryo Kurazume †, Adrian Stoica ‡
†*Information Science and Electrical Engineering, Kyushu University, Fukuoka, Japan*
‡*Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA*
Email: yumi@ieee.org

***Abstract*—This paper presents a person identification technique that uses information from person's shadow, and is robust to appearance changes caused by variations of clothes and carried objects. The technique uses invisible lights and resulting shadows and has advantages from undetected sensing. The shadows on the ground obtained through illumination by multiple lights can be considered as silhouettes captured by multiple virtual cameras placed at light positions. Thus, a single camera, e.g. in the ceiling, is able to obtain multiple silhouettes, equivalent to a multi-camera system. If the person's appearance changes compared to the training cases in the database, e.g. by wearing different clothes or carrying a/another bag, then the identification performance gets worse. To deal with this problem, we introduce a new shadow-based identification technique robust to appearance changes. Firstly, we divide each shadow area into several parts, and estimate the discrimination capability for each part based on gait features between gallery datasets and probe dataset. Next, according to the estimated capability, we adaptively control the priorities of these parts in the person identification method. We constructed a new shadow database with a variety of clothes and bags, and carried out successful experiments to verify the effectiveness of the proposed technique.**

I. Introduction

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 [10] [3]. In applications related to security in controlled spaces, such as at the airport corridors leading to an airport security passport control, 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).

Gait identification methods based on appearance information use measurements of gait features from silhouettes, obtained by feature extraction methods, such as gait energy image(GEI) [5], active energy images (AEI), [13], Fourier transforms [11] , and affine moment invariants [6]. These methods 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 the reduced cost and fewer technical difficulties including (i) easier installation in real environments, less cameras and no need to synchronize cameras, (ii) a reduction of computation costs.

As one of options to deal with these problems, Iwashita $et\ al.$ proposed a person identification system [9] using shadow biometrics [7] and a shadow database. In this system, they placed multiple lights obliquely to project shadows of a subject on the ground, and placed a camera on the ceiling to capture the shadows. Shadow areas, which are projections of the subject's body on the ground by multiple lights, can be considered as body areas captured by multiple cameras from different viewpoints, so the system enabled to capture multiple body areas from only one camera. Furthermore, in the system they used infrared lights, with an advantage of 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.

The work described in [9] relied on extracted affine moment invariants, calculated from an average image, as gait features. The average image was created from sequential silhouette images, which were aligned with a center position of the subject area in each silhouette image. The performance was not high enough, due to the fact that when the subject position with respect to the light source changes compared with the one in the database, the shape of the subject area changes. This causes variations of center positions of the subject's areas in sequential silhouette images, which make different average images compared with the ones in the database, and the discrimination capability of the average image decreases. Besides, in case that the appearance of the subject changes due to variations of clothes and bags, the performance of the system [9] gets worse.

In this paper we propose an identification method, which achieves higher performance compared with the previous method [9] and is robust to appearance changes. The proposed method is based on the following two methods; (i) a voting-based method which achieves high performance [1] and (ii) a person identification method robust to appearance changes [8]. Both the voting-based method [1] and the person identification method robust to appearance changes [8] are for identifying walking people using captured images of their bodies, and the methods did not use any shadow images. In the present paper, we propose a method which integrates both previous methods [1] [8]. In the proposed method, firstly we divide a shadow area into several parts, and estimate the discrimination capability for each part based on gait features between gallery datasets and probe dataset. Next, according to the estimated capability, the priorities of these parts for

978-1-4799-7007-0/14 $31.00 © 2014 IEEE
DOI 10.1109/EST.2014.18

the person identification are controlled adaptively. We show the effectiveness of the proposed method through experiments with a new shadow database including 54 people.

This paper is organized as follows. Section 2 describes a new shadow database with a variety of clothes and bags, and section 3 describes the details of the proposed person identification method. Section 4 explains the experiments using the database. Conclusions are presented in section 5.

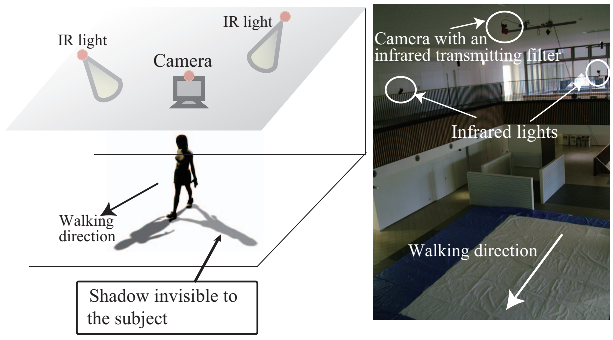
II. Invisible shadow database

Figure 1 shows examples of captured images in which a subject's shadow was projected by an infrared light (Bosch, IR Illuminator 850 nm, UFLED30-8BD). Figure 1 (a) shows an example of captured images with a visible light transmitting filter, which are equivalent to images people see with their eyes, and Fig. 1 (b) shows an example of images with an infrared transmitting filter. These captured images show that a shadow by the infrared right is not visible to human.

(a) (b)

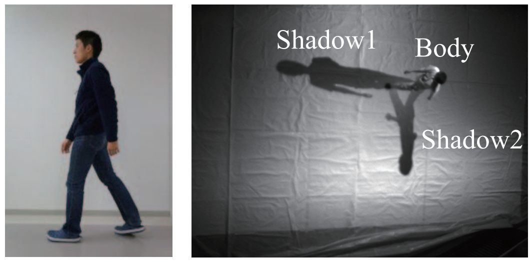
Figure 1. Examples of captured shadows projected by an infrared light, (a) the image with a visible light transmitting filter (this image is equivalent to how people see), (b) the image with a infrared transmitting filter.

To collect a shadow database, we used two infrared lights and a camera (PointGrey Research Inc., Grasshopper2 M/C). The infrared lights were placed obliquely and the camera was placed on the ceiling perpendicular to the ground as shown in Fig. 2. Previous papers showed that front-view gait images had the best performance compared with the rest of viewpoints [12], thus we placed one of the cameras to capture front-view images of walking people. We placed the other camera to get lateral-view images which produce a big difference of their appearance compared with the frontal images. The velocity of the subject in the observation area is constant because the initial acceleration of walking occurred outside the observation area. The image resolution and the frame rate are 1600 $\times$ 1200 and 30 Hz, respectively. The number of subjects is 54, 6 gait sequences for each, which consists of 4 standard gait sequences (IRSD-ST), 1 carrying-bag sequence (IRSD-BG), and 1 changing-cloths sequence (IRSD-CL). Figure 3 (a) shows the standard cloth and an example of captured images. The sequences in the IRSD-BG are categorized into 3 categories, (i) backpack (18 people), (ii) hand bag (18 people), (iii) traveling bag (18 people), as shown in Fig. 4 (a) $\sim$ (c). The sequences in the IRSD-CL are categorized into 3 categories, (i) down jacket (18 people), (ii) laboratory coat (18 people), (iii) coat (18 people), as shown in Fig. 4 (d) $\sim$ (f).



(a) (b)

Figure 2. (a) Experimental setting, (b) actual scene.



(a) (b)

Figure 3. (a) Standard cloth, (b) an example of captured images.

III. Person identification using invisible shadow

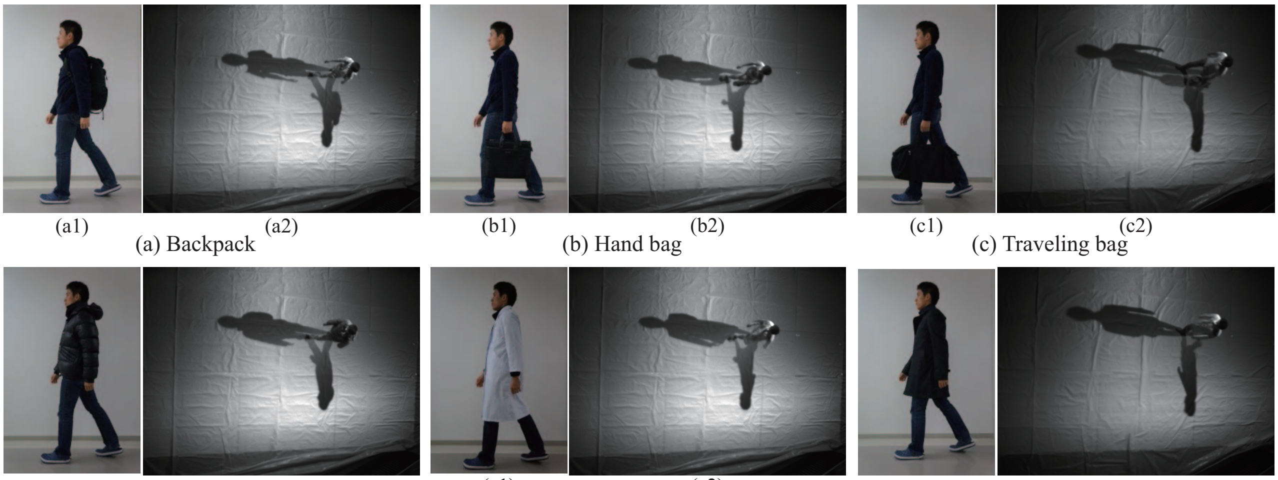
In this section, we describe the details of the proposed method. At first, we describe a method of feature extraction from gait images, and we introduce the person identification method which achieves higher performance compared with the previous method [9] and is robust to appearance changes.

A. Extraction of a subject's area and gait features

First, a background subtraction method is applied to captured images to a subject's area including his body and shadow areas. Figure 5 shows results of background subtraction of Fig. 3 (b) and one of Fig. 4 (a2) (backpack). There is noise and deficit in the silhouette images.

Next, affine moment invariants are extracted as gait features [6]. 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

$$\mu_{pq} = \sum\sum\nolimits_{(x,y)\in O}(x-x_g)^p(y-y_g)^q I(x,y). \quad (1)$$



(a1) (a2) (a) Backpack (b1) (b2) (b) Hand bag (c1) (c2) (c) Traveling bag

(d1) (d2) (d) Down jackaet (e1) (e2) (e) Laboratory coat (f1) (f2) (f) Coat

Figure 4. (a1) $\sim$ (f1) Variations of clothes and bags, (a2) $\sim$ (f2) examples of captured images.

(a) (b)

Figure 5. (a) Extracted subject area of Fig. 3 (b), (b) extracted subject area of Fig. 4 (a2) (backpack).

Here, x_g and y_g are the center of the object and $I(x, y)$ represents the intensity of the pixel (x, y). The number of affine moment invariants ($\boldsymbol{A} = (A_1, A_2, \ldots, A_M)^T$) is M; we show two of them in the following [4].

$$
\begin{aligned}
A_1 &= \frac{1}{\mu_{00}^4}(\mu_{20}\mu_{02} - \mu_{11}^2) \\
A_2 &= \frac{1}{\mu_{00}^{10}}(\mu_{30}^2\mu_{03}^2 - 6\mu_{30}\mu_{21}\mu_{12}\mu_{03} + \\
&\quad 4\mu_{30}\mu_{12}^3 + 4\mu_{03}\mu_{21}^3 - 3\mu_{21}^2\mu_{12}^2) \qquad (2)
\end{aligned}
$$

B. Method robust to appearance changes

The proposed method is based on the voting-based method [1] and the person identification method robust to appearance changes [8]. We briefly explain both methods in this section, and introduce the proposed method.

1) Voting-based method [1]: The voting-based method [1] is proposed to achieve high performance even with variations of center positions of a subject's areas in sequential silhouette images. In the training phase, gait features of affine moment invariants are extracted from each image in the training sequences. Then, gait features are extracted from each image of a subject's sequence in the test dataset. At each frame of the subject's sequence, the distance between the feature of the subject and all features in the training sequences is calculated as follows.

$$d_{n,s,f}^{j} = \|{}^{\boldsymbol{w}}\boldsymbol{A}_{SUB}{}^{j} - {}^{\boldsymbol{w}}\boldsymbol{A}_{DBn,s,f}\| \qquad (3)$$

Here, ${}^{\boldsymbol{w}}\boldsymbol{A}_{SUB}{}^{j}$ and ${}^{\boldsymbol{w}}\boldsymbol{A}_{DBn,s,f}$ show whitened affine moment invariants of the subject and a person in the database, respectively. j, n, s, and f are $1 \leq j \leq J$ (J is the number of images of the one gait cycle of the subject), $1 \leq n \leq N$ (N is the number of people in the database), $1 \leq s \leq S$ (S is the number of sequences of each person in the database), and $1 \leq f \leq F$ (F is the number of images of the one gait cycle in the database), respectively. $\|\cdot\|$ means euclidean norm of $\cdot$. The person, who has the smallest distance, in the training dataset, is voted at each frame of the subject's sequence. Figure 6 shows examples of the voting. In the identification phase, the person who get the most votes is chosen as the subject.

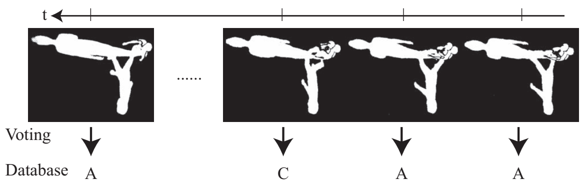


Figure 6. An example of the voting-based method.

2) Person identification method robust to appearance changes [8]: In the method robust to appearance changes, gait features are extracted from an average image created from silhouette images of one gait cycle. At first the subject's area is divided equally into K areas according to the height of the subject's area. Figure 7 shows examples of a human body area divided into 5 areas. At each area affine moment invariants are calculated. Database is built from a set of affine moment invariants of multiple people who wear standard clothes without belongings.

Next a matching weight at each area is estimated according to the similarity between the features of the

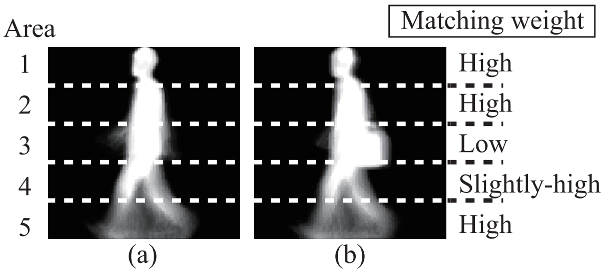


Figure 7. (a) An example of average images in the database, (b) an example of average images of subjects with a shoulder bag.

subject and those in the database. In case that the subject's appearance is different from that in the database as shown in Fig. 7, matching weights of areas with appearance changes are set to low. On the other hand, matching weights of areas with less appearance changes are set to high. This method doesn't utilize gait features extracted from areas with low matching weights, which are due to changes of clothes / belongings, but utilizes features from areas with high matching weights. Therefore the method enables person identification robust to changes in appearance.

The distance between the feature of the subject and that in the database is used to determine a matching weight at each area, and small and large distances mean high and low matching weights, respectively. The procedure to estimate the matching weight is briefly explained as follows. At each area k, a threshold $th(k)$ is set adaptively, and we select sequences from the database if $d_{n,s}(k) < th(k)$. These selected sequences are considered to have high similarities. Non-selected sequences can be considered to have low similarities, and the distance $d_{n,s}(k)$ is redefined as a value $d_{max}(k)$, which is the largest distance among all sequences at area k. The above process is done all areas, followed by people identification. For more details, please refer to [8].

3) Proposed method: In the paper [8], Iwashita et $al.$ applied the method, which is robust to appearance changes, to the CASIA gait database, and obtained the higher performance compared with conventional methods [2] [13]. In the present paper, we apply the procedure explained in section 3.2.2 to the voting-based method to calculate the matching weight at each frame, and replace the distance of Eq. 3 with the matching weight. We also divide a subject's area into multiple areas, and the voting-based method with the matching weight explained in section 3.2.2 is applied to each area, followed by person identification.

In images in the shadow database, there are two shadow areas and one body area, as shown in Fig. 3 (b). There are several options to divide a silhouette area, which includes two shadow areas and one body area, into multiple areas such as (i) areas divided according to the height of the silhouette area, using the same process with [8], and (ii) areas divided according to the height and width as shown in Fig. 8, and (iii) each shadow and body area as shown in Fig. 9 (b) $\sim$ (d). In terms of the option (i), gait information of a left shadow area in a silhouette image cannot be extracted efficiently, since to extract gait features efficiently from the left shadow area, it is desirable to divide the left area with vertical lines. Compared with the option (i), the option (ii) can efficiently extract gait information from each shadow area. In terms of the option (iii), if a subject's appearance changes, there is a chance that in all three areas the subject's appearance changes. This results in decreasing the performance of people identification. If we divide a silhouette area into more areas than the areas of the option (iii), we have bigger chance to identify the subject with higher performance. Thus in the proposed method, we use the option (ii), to divide the subject area equally according to the height and width, and we extract gait information from each area separately. Figure 8 shows examples (the number of divided areas K = 1, 4, and 6). Comparisons of the options (ii) and (iii) are shown in experiments.

The performance of the previous method using the infrared-based shadows [9] gets worse in both cases of appearance changes of a subject and variations of center positions of the subject's areas. On the other hand, the proposed method achieves better performance, which is shown in experiments. Moreover, an advantage of the proposed method is to identify the subject whose appearance changes, even if the proper number of divided areas is not known, since the priorities of divided areas for the person identification are controlled adaptively.

Figure 8. Examples of divided areas (the number of divided areas K = 1, 4, and 6).

IV. Experiments

This section shows the results of person identification experiments using the shadow database in Section 2. The database contains 54 people with 6 sequences for every subject. Six sequences consist of 4 standard gait sequences (IRSD-ST), 1 carrying-bag sequence (IRSD-BG), and 1 changing-cloth sequence (IRSD-CL). We carried out three experiments: (i) person identification with standard sequences (IRSD-ST), (ii) person identification with appearance-changed sequences (IRSD-BG and IRSD-CL), and (iii) person identification to validate discrimination capabilities of shadow and body areas. The correct classification rate was estimated based on the idea of 4-fold cross validation. The concrete procedure is as follows. Three sequences in the IRSD-ST are used as training

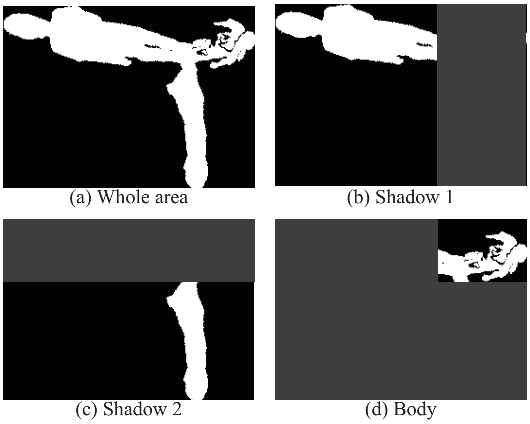


Figure 9. Examples of divided areas into each shadow and body area.

datasets, thus there are 4 kinds of combinations of training datasets. Regarding a test dataset, in case of the IRSD-ST, the rest of the sequence which is not used in the training dataset is used as a test dataset, and in case of both IRSD-BG and IRSD-CL each sequence is used. A correct classification rate is calculated by averaging results of 4-fold cross validation.

A. Person identification with standard sequences (IRSD-ST)

In this experiment, to show the effectiveness of the voting-based method [1], we compared the voting-based method [1], which is not robust to appearance changes, and a conventional method which used average images to extract gait features [6] [9]. We did not divide the subject area (K, the number of divided areas, is 1) and we varied the parameter of M (the total number of affine moment invariants) from 1 to 80. Figure 10 shows the results of the experiments with respect to the change of M. From these results, it is clear that the voting-based method outperformed the conventional method.

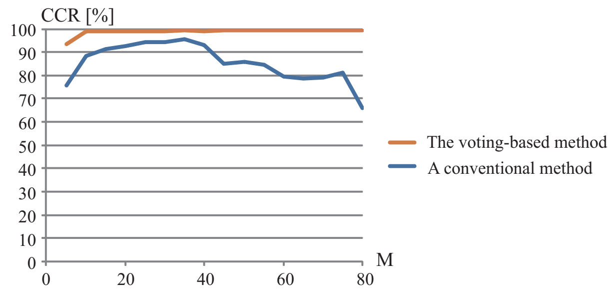


Figure 10. Correct classification rates by the voting-based method [1] and a conventional method [6] (IRSD-ST).

B. Person identification with appearance-changed sequences (IRSD-BG and IRSD-CL)

In next experiments, we applied the proposed method to IRSD-BG and IRSD-CL. In addition to that, we applied the voting-based method [1], which does not utilize the method robust to appearance changes, to see the effectiveness of the use of the method robust to appearance changes [8]. We varied the parameter of K (the number of divided areas) from 1 to 20 and M (the total number of affine moment invariants) from 1 to 80. We tested all combinations of K and M. Figures 11 and 12 show examples of CCRs (K=4, 8, 14, and 20) of IRSD-BG and IRSD-CL. From these results, it is clear that the proposed method outperformed the voting-based method and CCR increased with the parameters K and M. However, CCRs of IRSD-CL are relatively lower than those of IRSD-BG, thus we checked the CCRs for each category of IRSD-CL, as shown in Table I. From these results, the CCR of down jacket is the worst among all categories. One of the reasons is that the discrimination capability of the subject's area of down jacket decreased due to the thickness of the down jacket.

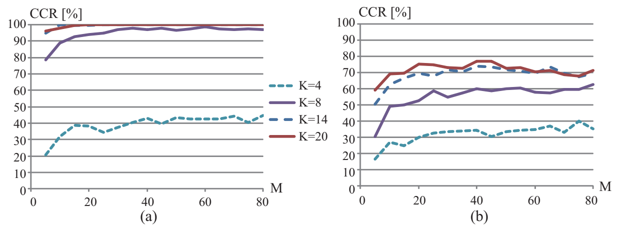


Figure 11. Comparison of correct classification rates of (a) the proposed method and (b) the voting-based method (IRSD-BG).

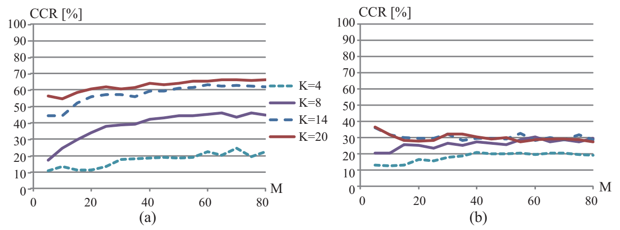


Figure 12. Comparison of correct classification rates of (a) the proposed method and (b) the voting-based method (IRSD-CL).

Table I
CORRECT CLASSIFICATION RATE OF EACH CATEGORY OF CASIA-B-BG BY THE PROPOSED METHOD[%].

	Down jacket	Laboratory coat	Coat
CCR [%]	26.4	77.8	91.7

C. Person identification to validate discrimination capabilities of shadow and body areas

As we explained in Section 3.2.3, there are another options to divide a subject's area, and one of them is to divide the subject's area into each body and shadow area [7]. In this section we separate the body and shadow areas as shown in Fig. 9, by detecting the lowest pixel of shadow 1 and the leftest pixel of shadow 2, and extract features from each area followed by person identification. We carried out five experiments with standard sequences (IRSD-ST), carrying-bag sequence (IRSD-BG), and changing-cloth sequence (IRSD-CL):

(1) Person identification with a whole area (i.e. the area was not divided, Fig. 9 (a))
(2) Person identification with $\boldsymbol{shadow1}$ (Fig. 9 (b))
(3) Person identification with $\boldsymbol{shadow2}$ (Fig. 9 (c))
(4) Person identification with $\boldsymbol{body}$ (Fig. 9 (d))
(5) Person identification with combination of $\boldsymbol{shadow1}$, $\boldsymbol{shadow2}$, and $\boldsymbol{body}$

Here, the voting-based method [1], which is not robust to appearance changes, is applied for evaluation. Figures 13 $\sim$ 15 show results of the above five experiments with standard sequences (IRSD-ST), carrying-bag sequence (IRSD-BG), and changing-cloth sequence (IRSD-CL). We varied the parameter of M from 1 to 80. Figures 13 $\sim$ 15 also show results by the proposed method with $K = 20$. The results show that the proposed method achieved the highest performance. Interestingly, the results of shadow 1 showed better results compared with other areas (i.e. body and shadow 2). This is because the area of shadow 1 tends to receive smaller influence of appearance changes compared with other areas.

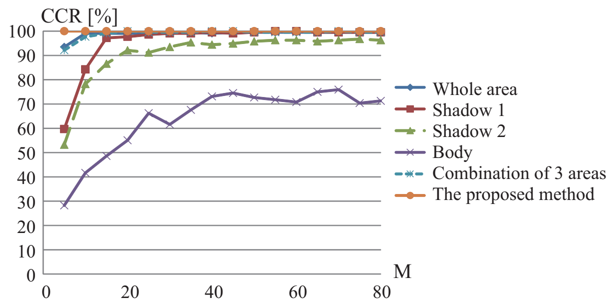


Figure 13. Comparison of correct classification rates of (1) whole area, (2) shadow 1, (3) shadow 2, (4) body, (5) combination of 3 areas, and (6) the proposed method ($K = 20$) (IRSD-ST).

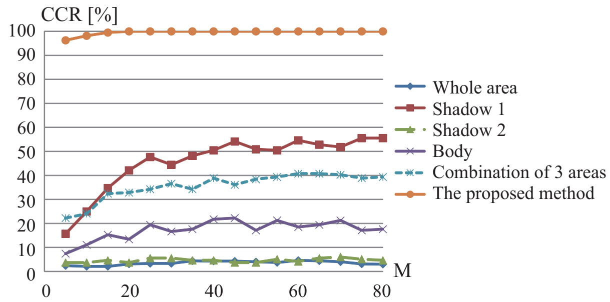


Figure 14. Comparison of correct classification rates of (1) whole area, (2) shadow 1, (3) shadow 2, (4) body, (5) combination of 3 areas, and (6) the proposed method ($K = 20$) (IRSD-BG).

V. Conclusion

We proposed a shadow-based person identification technique robust to appearance changes caused by variations of clothes and carried bags conditions. In addition, we built a new shadow database with a variety of clothes and bags. The system to collect the database consists of multiple infrared lights and a camera, and it has the advantages of an invisible and undetected sensing system. 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. We carried out experiments with the new database, and showed the robustness of the proposed method compared with conventional methods against appearance changes in all, except one clothing condition (down jacket).

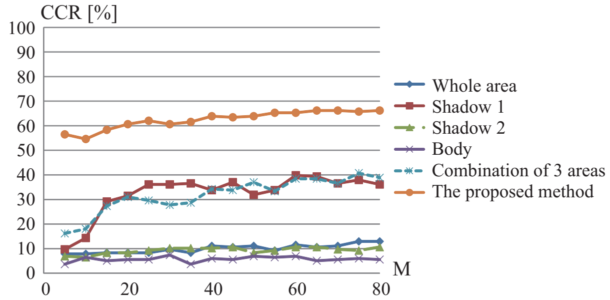


Figure 15. Comparison of correct classification rates of (1) whole area, (2) shadow 1, (3) shadow 2, (4) body, (5) combination of 3 areas, and (6) the proposed method ($K = 20$) (IRSD-CL).

Acknowledgment

This work is supported by Strategic Information and Communications R&D Promotion Program (SCOPE) No.121810005.

References

[1] Authors. Identification of people walking along curved trajectories. *Pattern Recognition Letters, Supplied as additional material prl.pdf*.
[2] K. Bashir, T. Xiang, and S. Gong. Gait recognition without subject cooperation. *Pattern Recognition Letters*, 31(13):2052–2060, 2010.
[3] I. Bouchrika, M. Goffredo, J. Carter, and M. Nixon. On using gait in forensic biometrics. *Journal of Forensic Sciences*, 56(4):882–889, 2011.
[4] J. Flusser, T. Suk, and B. Zitova. Moments and moment invariants in pattern recognition. *Wiley & Sons Ltd.*, 2009.
[5] J. Han and B. Bhanu. Individual recognition using gait energy image. *IEEE Trans. PAMI*, 28(2):316–322, 2006.
[6] Y. Iwashita and R. Kurazume. Person identification from human walking sequences using affine moment invariants. *Proc. IEEE Int. Conf. Robotics and Automation*, pages 436–441, 2009.
[7] Y. Iwashita, A. Stoica, and R. Kurazume. Gait identification using shadow biometrics. *Pattern Recognition Letters*, 33:2148–2155, 2012.
[8] Y. Iwashita, K. Uchino, and R. Kurazume. Gait-based person identification robust to changes in appearance. *Sensors*, 13(6):7884–7901, 2013.
[9] Y. Iwashita, K. Uchino, R. Kurazume, and A. Stoica. Gait identification from invisible shadows. *SPIE Biometric Technology for Human Identification IX*, 2012.
[10] M. Nixon, T. Tan, and R. Chellappa. Human identification based on gait. *Springer*, 2005.
[11] K. Sugiura, Y. Makihara, and Y. Yagi. Gait identification based on multi-view observations using omnidirectional camera. *Asian Conf. on Computer Vision*, 1:452–461, 2007.
[12] S. Yu, D. Tan, and T. Tan. A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition. *Int. Conf. on Pattern Recognition*, 4:441–444, 2006.
[13] E. Zhang, Y. Zhao, and W. Xiong. Active energy image plus 2dlpp for gait recognition. *Signal Processing*, 90(7):2295–2302, 2010.