ICIC Express Letters
Part B: Applications
Volume **8**, Number **2**, February **2017**

 ISSN 2185-2766
pp. **297–303**

IDENTIFICATION OF PEDESTRIAN FROM MOVING IMAGE OF SURVEILLANCE CAMERA USING ACCUMULATION AND FOURIER TRANSFORM

Shigeyoshi Nakajima

Graduate School of Engineering
Osaka City University
3-3-138, Sugimoto Cho, Sumiyoshi Ku, Osaka City 558-8585, Japan
nakajima@info.eng.osaka-cu.ac.jp

Received July 2016; accepted October 2016

Abstract. *Needs for surveillance systems increase in these days because of terrorisms, injuries and others. We propose a new method to determine the same person or a different person in a movie taken by sparsely distributed surveillance cameras. GEI (gait energy image) is one of the methods to determine the same person or different persons in two shots of moving images. GEI is an accumulation image. However, temporal changes vanish in GEI. We propose a method to extract temporal changes using matches of an accumulation image and contour of each frame, and to calculate the distance of results of Fourier transform from the temporal changes. Precision and recall of the proposed method were over 0.9.*

Keywords: GEI, Fourier transform, Surveillance camera, Sparsely distributed cameras, Pedestrian, Identification

1. **Introduction.** There are many surveillance cameras in our day-to-day life indoors and outdoors. Needs for surveillance systems increase in these days because of terrorisms, injuries and others. A person watches some monitors of surveillance cameras in an ordinary surveillance system. However, it costs much human powers. So an automatic surveillance system will be very useful.

V. R. Jadav et al. [1] employed GEI. Their method compensated lost gait contour sequences from other extracted gait contours using GEI. H. Ali et al. [2] extracted features of gaits of humans from analysis of GEI using Radon transform. They used ordinary image date to make GEI data. Our method increases ability of GEI using MSER data of images. GEI is available with a small area of a pedestrian in an image and low resolution of a camera. We tried to improve GEI using MSER [3].

P. Viola and M. Jones [4] proposed a fast method to detect face areas from images of persons. J. Zhou and J. Hoang [5] proposed a method to track a person and detect face regions. D. R. Chourasiya et al. [6] achieved accuracy over 0.95 using GEI. However, their method needs a learning as a preprocessing. A learning process costs much time. However, our proposed method need not a learning process.

GEI is one of the methods to distinguish other persons. However, a shape of GEI is affected by a cloth and a bag. Assuming that the first movie α of a man with a thick coat was taken by a camera, and then he undressed the coat and had it on his hand and his second movie β was taken by another camera, and the GEI of α does not match with the GEI of β. And the same situation could occur if he had a bag. However, if the coat (or the bag) was not so heavy as compared to his weight, a motion of his gravity center in α is almost the same as the motion in β. It is very useful to extract a motion from a movie. However, a region of a human body is affected if a color of a part of a body is similar to a color of a part of background. So extraction of a correct motion is difficult.

On the other hand, temporal changes of human gravity centers are investigated. The motion of human body shows features of gaits. Visualization of invisible data is a target of many researchers [7]. M. Ikebuchi et al. [8,9], Kato et al. [10-12] and N. Akamatsu et al. [13] worked with accelerometer data using DCT (discrete cosine transform) method for diagnoses of recoveries of patients after surgical operations. S. Nakajima et al. proposed a method to visualize a 3D trajectory of a human gravity center [14].

We propose a method to discriminate pedestrian images of the same person and pedestrian images of different persons using combination of accumulated images and Fourier transform. The proposed method need not a learning process different from D. R. Chourasiya et al. [6]. Our method uses frequency distribution which is not affected by change of clothes nor bags.

The rest of this paper is organized as follows. Section 2 outlines the proposed method to measure similarity of moving image. Section 3 presents experimental result and a consideration. Section 4 shows a conclusion.

I and collaborators analyzed human gaits from data of a wearable accelerometer. Acceleration data is time variable.

2. **Method.** Figure 1 shows a block diagram of the proposed method.

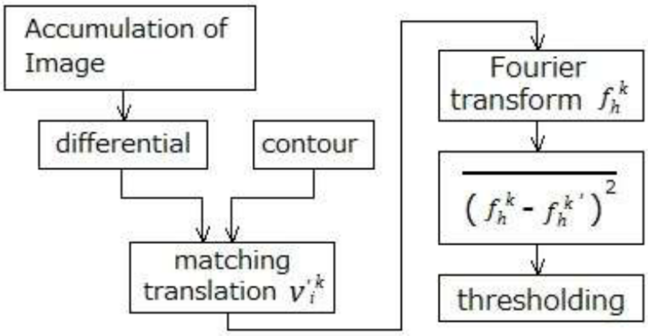

FIGURE 1. Block diagram

2.1. **Accumulation.** The ith image frame $I_i^k(x,y)$ captured from a moving image in which a pedestrian k was taken is accumulated to an image $A^k(x,y)$ as the following Equations (1) and (2).

$$A^k(x,y) = \frac{1}{n}\sum_{i=1}^{n} I_i^k\left(x - x_i^k, y - y_i^k\right) \tag{1}$$

$$v_i^k = \left(x_i^k, y_i^k\right) \tag{2}$$

where n: a number of frames, k: a person number, $I_i^k(x,y)$: a pixel gray level of the ith frames of the kth person, v_i^k: a 2D translation vector. The image $I_i^k(x,y)$ is a monochrome image captured from a moving image of a surveillance camera. Figure 2 shows a sample of $I_i^k(x,y)$. The moving image is one of TUMIITGAIT data bases provided by Technology University Munthen [15].

The variables x and y mean coordinate values. The accumulated image is $I^k(x,y)$. A vector v_i^k is a centroid which is the mean position of positive (white) pixels. Figure 3 shows one of the accumulated images $A^k(x,y)$. GEI was an accumulated image with normalization. However, we do not use normalization because it violates an original contour.

FIGURE 2. Silhouette

FIGURE 3. Accumulation A^k

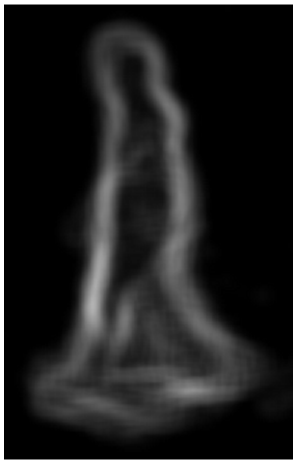

FIGURE 4. Differential image D^k

2.2. **Differential.** Figure 4 shows a differential image $D^k(x, y)$ of the accumulated image $A^k(x, y)$. We used Prewitt filter. $D^k(x, y)$ is a square root of sum of squares of Prewitt values along x-axis and y-axis. There are some thick parts. One of them is a lower front part of a belly. And another part is a front of a thigh. There are high values at a front of a thigh and a front of a belly.

(a) Contour (b) Best $e_i^k(\boldsymbol{v})$ (c) Not best $e_i^k(\boldsymbol{v})$

FIGURE 5. Matching of differential and contour

2.3. **Matching of contour and differential.** We made a contour image $C_i^k(x, y)$ shown in Figure 5(a) of the ith frame of the kth pedestrian shot and matched it onto the differential image $D^k(x, y)$ as shown in Figure 5(b) and Figure 5(c). A 2-D vector $v_i'^k$ is the best translation vector which superimposes a contour onto the differential image. The pixel values of the differential image on a contour are summed $e_i^k(\boldsymbol{v})$ as shown in Equation (3). The value of $e_i^k(v_i'^k)$ is the best value among the values of $e_i^k(\boldsymbol{v})$ for the ith frame of the kth pedestrian. Figure 5(b) is the best match. Figure 5(c) is not the best match.

$$e_i^k(\boldsymbol{v}) = \frac{1}{wh} \sum_{x=1}^{w} \sum_{y=1}^{h} D^k(x, y) C_i^k(x - \boldsymbol{v}.x, y - \boldsymbol{v}.y) \tag{3}$$

There are some artifacts in the silhouette Figure 2. There is a luck part of a silhouette in a chest part of Figure 2. Also there is a convex part in front of a belly. However, illegal parts of contour do not match to the differential image in Figure 5(c). However, the front of thigh matches to the differential image as shown in Figure 5(b).

Figure 6 shows vertical elements of $v_i'^k$ (i.e., up-down motion) and Figure 7 shows horizontal elements of $v_i'^k$. Both were results after linear revision. We expected that $v_i'^k$ is different from $v_i'^{k'}$ if pedestrians k and k' are other persons but is similar if k and k' are the same person.

2.4. **Fourier transform.** We employed Fourier transform to calculate similarity of $v_i'^k$. Figure 8 shows f_h^k. f_h^k is Fourier transform of $v_i'^{k'}$. The variable h is an integer multiplier. The value 1 of h means a stride frequency. A stride consists of a left step and a right step. There are 2 strides in Figure 6. The value 2 of h means a step frequency which is twice of a stride frequency. There are 4 steps in Figure 6. In Figure 8 data H_RE means a real part of a horizontal motion. IM means an imaginary part. V means a vertical motion. Figure 9 shows a frequency distribution of a different pedestrian k'. The value of h means a multiple number of frequency when harmonics of the stride frequency is 1. So f_h^k is a 4D vector. And $1 \leqq h \leqq 6$ then there are 6 vectors for a stride. We calculate $m_{kk'}$ mean of squares of norms of difference vector in Equation (4). We expect that the value $m_{kk'}$

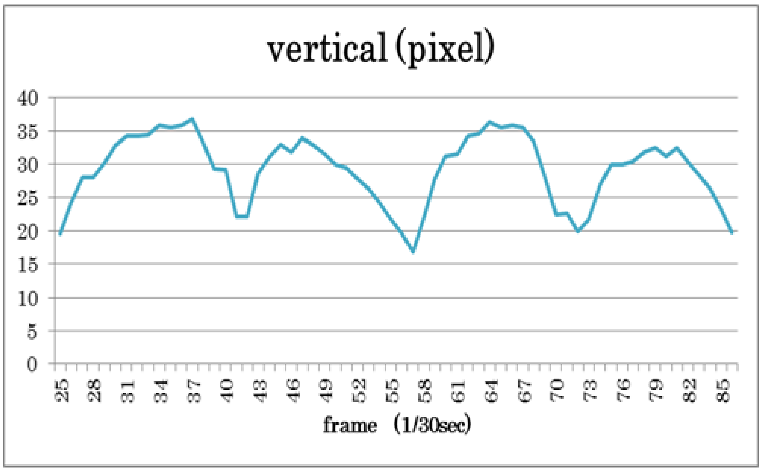

FIGURE 6. Vetrical element of $v_i'^k$

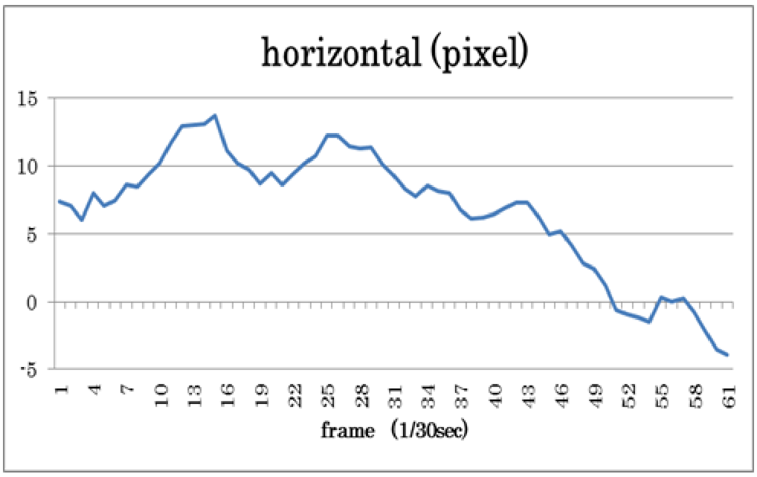

FIGURE 7. Horizontal element of $v_i'^k$

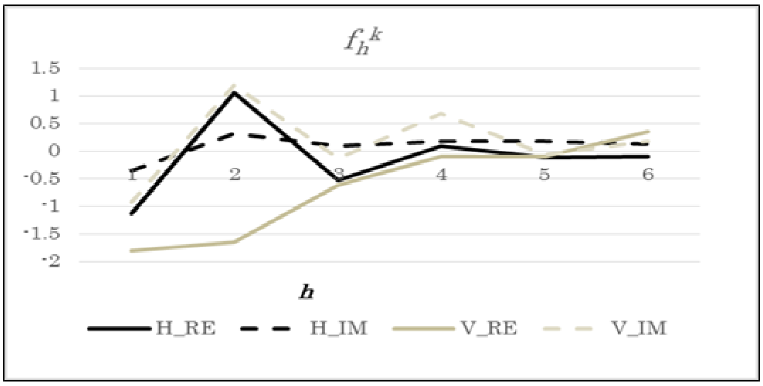

FIGURE 8. f_h^k: Fourier transform of $v_i'^k$

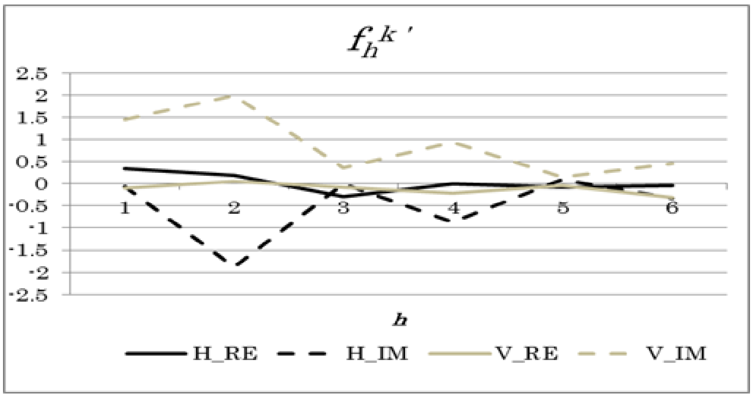

FIGURE 9. $f_h^{k'}$: Fourier transform of $v_i'^{k'}$

is small if the same person but large if other persons.

$$m_{kk'} = \frac{1}{6}\sum_{h=1}^{6}\left|\left|f_h^k - f_h^{k'}\right|\right|^2 \tag{4}$$

2.5. **Thresholding.** We expect that the graph of Figure 8 and Figure 9 is similar if pedestrians of both shots are the same person. They are not similar if pedestrians are different persons. So $m_{kk'}$ is small if pedestrians are the same person. And $m_{kk'}$ is large if pedestrians are different persons. We determine threshold level t. We determine that if $m_{kk'} \leqq t$, the pair kk' are positive (estimated as the same person) and if $m_{kk'} > t$, the pair kk' are negative (estimated as different persons).

3. **Experimental Result.** There are 35 video sequences of "normal idle" folder of TU-MIITGAIT. We selected 20 sequences and captured 2 shots for each sequence. We made 20 pairs of shots of the same persons (taken at other times of the same sequence). Also we made 20 pairs of shots of different persons (i.e., different sequences). We calculated 40 values of $m_{kk'}$ of those 20+20 pairs. We selected a threshold value t as 0.21 for Table 1.

TABLE 1. Experimental result

	$m_{kk'} \leqq t$	$m_{kk'} > t$
Same	18 (TP)	2 (FN)
Different	1 (FP)	19 (TN)

18 pairs of the same persons are under the threshold value t (TP: true positive) in Table 1. 2 pairs of the same persons are over t (FN: false negative). 1 pair of different persons is under t (FP: false positive). 19 pairs of different persons are over t (TN: true negative). The precision value and recall value with the threshold value t $(= 0.21)$ are shown in Equations (5) and (6).

$$precision = TP/(TP + FP) = 18/19 \fallingdotseq 0.947 \tag{5}$$

$$recall = TP/(TP + FN) = 18/20 = 0.9 \tag{6}$$

Figure 10 shows ROC (receiver operating characteristic) curve of this experiment. There are sensitivities (TP/(TP+FN)) and 1-specificities (TN/(TN+FP)) with various threshold values. AUC (area under curve) value of Figure 10 was 0.97.

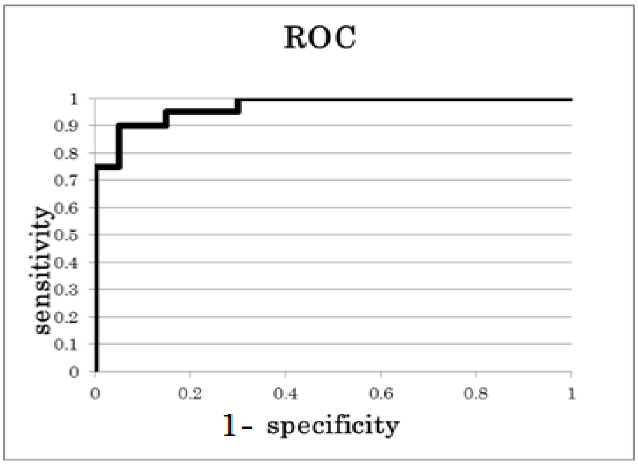

FIGURE 10. ROC curve

4. **Conclusion.** We proposed a method to calculate distance of movies using accumulation and Fourier transform and discriminate movies of the same person and movies of different persons with high accuracy. The proposed method is better than GEI with a popular binarization method. However, it is not enough for a practical surveillance scene. We will apply the proposed method for more data of TUMIITGAIT and other data and improve the accuracy of the method.

REFERENCES

[1] V. R. Jadav, G. Lachiram and N. Swathi, Gait recognition system for human identification using temporal preserving information, *International Journal of Computer Science and Electronics Engineering UK*, vol.4, no.3, pp.1-11, 2014.

[2] H. Ali, J. Dargham, C. Ali and E. G. Moung, Gait recognition using gait energy image, *International Journal of Signal Processing, Image Processing and Pattern Recognition*, vol.4, no.3, pp.141-152, 2011.

[3] J. Matas, O. Chum, M. Urban and T. Pajdla, Robust wide baseline stereo from maximally stable extremal regions, *Proc. of British Machine Vision Conference*, Cardiff, UK, pp.384-396, 2002.

[4] P. Viola and M. Jones, Rapid object detection using a boosted cascade of simple features, *Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Cambridge, MA, USA, vol.1, pp.I-511-I-518, 2001.

[5] J. Zhou and J. Hoang, Real time robust human detection and tracking system, *Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Diego, USA, p.149, 2005.

[6] D. R. Chourasiya, V. Richhariya and P. V. Rajpoot, Human recognition through gait energy image and principal component analysis using TUMIITKP database, *International Journal of Computer Science and Information Technology & Security*, vol.3, no.6, pp.409-416, 2013.

[7] S. Ishimitsu, K. Sakamoto, T. Yoshimi, Y. Fujimoto and K. Kawasaki, Study on the visualization of the impression of button sounds, *International Conference on Innovative Computing Information and Control*, vol.5, no.11(B), pp.4189-4203, 2009.
[8] M. Ikebuchi et al., Gait analysis by discrete cosine transform (DCT) in patients with osteoarthiritis of the hip using a wearable accelerometer, *Japanese Journal of Clinical Biomechanics*, vol.29, pp.355-359, 2008.
[9] M. Ikebuchi et al., Gait analysis by discrete cosine transform (DCT) in patients of hip osteoarthritis using wearable accelerometer, *Hip Joint*, vol.35, pp.726-731, 2009.
[10] R. Kato et al., Gait analysis by discrete cosine transform (DCT) in patients after total hip arthroplasty using a wearable accelerometer, *Japanese Journal of Clinical Biomechanics*, vol.29, pp.325-329, 2008.
[11] R. Kato et al., Prospect of gait after total hip arthroplasty in the early stage using a wearable accelerometer, *Hip Joint*, vol.35, pp.72-75, 2009.
[12] R. Kato et al., The change of hip abductor muscle strength after total hip arthroplasty, *Hip Joint*, vol.35, pp.115-118, 2009.
[13] N. Akamatsu et al., Assessment of gaiting body sway in patients with neck disease pre and post-operation, *Japanese Journal of Clinical Biomechanics*, vol.30, pp.116-166, 2009.
[14] S. Nakajima, M. Ikebuchi and T. Toriu, Band pass filtering at harmonics for visualization of 3D trajectory of human periodical walking motion from accelerometer, *ICIC Express Letters, Part B: Applications*, vol.2, no.3, pp.553-558, 2011.
[15] *http://www.mmk.ei.tum.de/∼hom/tumiit/tumiitgait.html*.