Human Activity Recognition in Thermal Infrared Imagery

Ju Han   and   Bir Bhanu

Center for Research in Intelligent Systems
University of California, Riverside, California 92521, USA
{jhan,bhanu}@cris.ucr.edu

Abstract

In this paper, we investigate human repetitive activity properties from thermal infrared imagery, where human motion can be easily detected from the background regardless of lighting conditions and colors of the human surfaces and backgrounds. We employ an efficient spatiotemporal representation for human repetitive activity recognition, which represents human motion sequence in a single image while preserving some temporal information. A statistical approach is used to extract features for activity recognition. Experimental results show that the proposed approach achieves good performance for human repetitive activity recognition.

1   Introduction

Human repetitive activity involves a regularly repeating sequence of motion events such as walking, running and jogging. Most existing human activity recognition approaches detect human motion in visible spectrum. However, it is very likely that some part of human body or clothing has similar color as background colors. In this case, human motion detection usually fails on this part. Moreover, the existence of shadows is a problem in visible spectrum. In addition, sensors in visible spectrum do not work under low lighting conditions such as night or indoor environment without lighting.

To avoid the disadvantages of using sensors in visible spectrum, we investigate the possibility of using the thermal infrared (long wave infrared) sensor for human activity analysis. Unlike a regular camera which records reflected visible light, a long wave $(8 - 12\mu m)$ infrared camera records electromagnetic radiation emitted by objects in a scene as a thermal image whose pixel values represent temperature. In a thermal image that consists of humans in a scene, human silhouettes can be easily extracted from the background regardless of lighting conditions and colors of the human surfaces and backgrounds, because the temperatures of the human body and background are different in most situations [3]. Figure 1 shows an example of hu-

man walking at different time in a day recorded from a thermal infrared sensor: noon (first row), late afternoon (second row) and night (third row). There is no obvious shadow introduced in the thermal infrared images recorded at noon and late afternoon. The thermal images also provide enough contrast between the human and the background at night.

In this paper, we investigate human repetitive activities in thermal infrared imagery. First, human motion is detected and human silhouettes are extracted from the background. Then, we employ an efficient spatio-temporal representation, Gait Energy Image (GEI), for human repetitive activity recognition. Unlike other representations which consider motion as a sequence of templates (poses), GEI represents human motion sequence in a single image while preserving some temporal information. Finally, we use a statistical approach to extract features from GEI for activity recognition.

2   Related Work

2.1   Object Detection in Thermal Infrared Imagery

Images from different kind of sensors generally have different pixel characteristics due to the phenomenological differences between the image formation process of the sensors. In recent years, some approaches have already been proposed to detect and recognize object in thermal infrared imagery, especially in the field of remote sensing and automated target recognition.

Object detection in thermal infrared imagery has been widely used in remote sensing. Holland and Yan [11] propose a method to quantitatively measure ocean surface movement using sequential $10.8\mu m$-band thermal infrared satellite images. Ocean thermal pattern features are selected by detecting and mapping gradients and at the same time discriminating between the water surface, land, and clouds. The pattern features are then tracked by a constrained correlation based feature recognition scheme in a subsequent image. Abuelgasim and Fraser [1] investigate the applicability of NOAA-16/AVHRR (N-16) satellite data for detecting and mapping active wildfires across North American for-
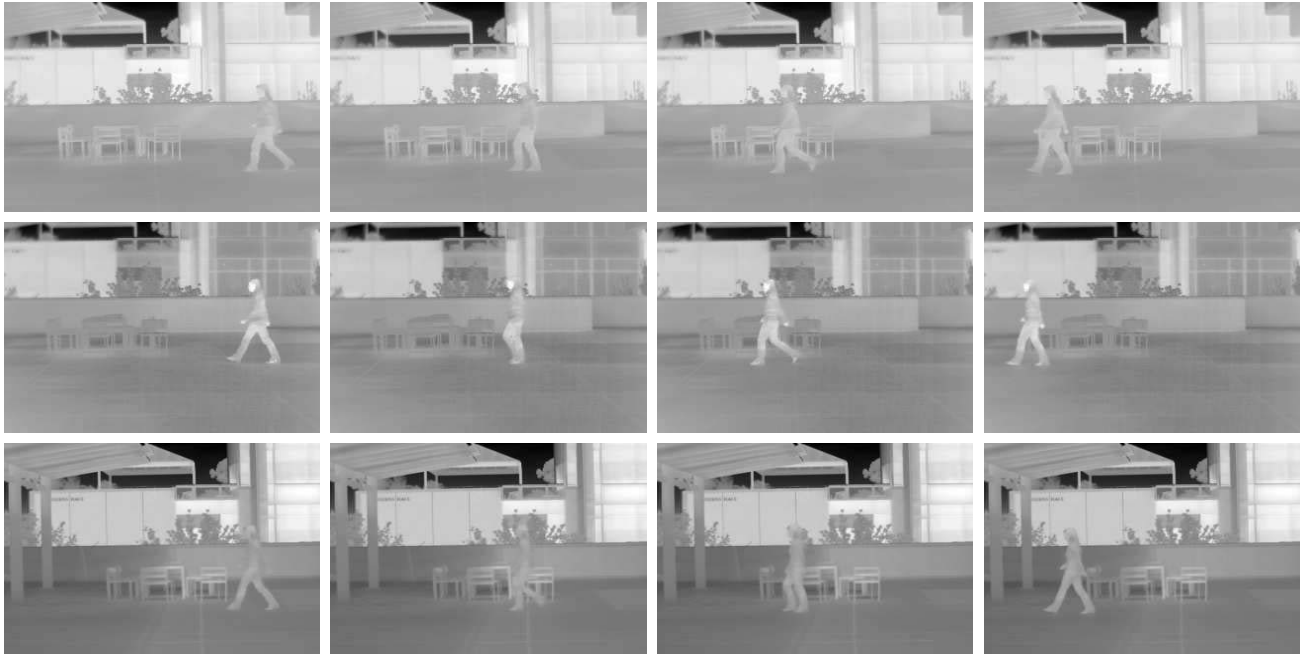
Figure 1. An example of human walking at different time in a day recorded from a thermal infrared sensor: noon (first row), late afternoon (second row) and night (third row).

est ecosystems. Their algorithms exploit both the multi-spectral and thermal information from the AVHRR daily images. Riggan and Hoffman [18] discussed the applications of a thermal imaging radiometer in field trials. The system has successfully demonstrated remote detection of a small spot fire, landscape temperature mapping; and quantitative, unsaturated measurements of flame radiance during large-scale open burning. Martinez et. al [14] develop a three-dimensional thermal model to study the effect of the presence of landmines in the thermal signature of the bare soil. In their approach, landmines are regarded as a thermal barrier in the natural flow of the heat inside the soil, which produces a perturbation of the expected thermal pattern on the surface. Surface-laid and shallowly buried landmines are detected and classified from measured infrared images by studying these perturbations.

Object detection in thermal infrared imagery has also been widely used in automated target recognition in surveillance environments, especially for human target detection. Andreone et. al [2] propose an approach for detecting vehicles in thermal infrared imagery. Initially the attention is focused on portions of the image that contains hot objects only. The result is further investigated exploiting specific vehicle thermal characteristics. Arlowe [3] develop an automatic detection systems based on the thermal contrast and motion of human intruders. The conditions and energy transfer mechanisms that lead to difficult thermal detection are discussed in his work. The heat flux balance equation can be used in an iterative computer program to predict the

surface temperatures of both the background and the target for human intruder detection. Ginesu et. al [8] propose a novel method to detect foreign bodies, which are not detectable using conventional methods, by inspecting food samples using thermographic images. Pavlidis et. al [15] propose a method for detecting suspects engaged in illegal and potentially harmful activities in or around critical military or civilian installations. The use of thermal image analysis are investigated to detect at a distance facial patterns of anxiety, alertness, and/or fearfulness. Yoshitomi et. al [21] develop a face identification approach by thermal image analysis. The front-view face is first normalized in terms of location and size, the temperature distribution is then measured as well as the locally averaged temperature and the shape factors of face. These features are used for supervised classification by neural network.

2.2 Human Activity Recognition

In recent years, various approaches have been proposed for human activity recognition. These approaches generally fall under two major categories: model-based approaches and model-free approaches.

When people observe human motion patterns, they not only observe the global motion properties, but also interpret the structure of the human body and detect the motion patterns of local body parts. The structure of the human body is generally interpreted based on their prior knowledge. Model-based activity recognition approaches focus

on recovering a structural model of human motion, and the motion patterns are then generated from the model parameters for activity recognition. Guo et. al [9] represent the human body structure in the silhouette by a stick figure model. The human motion characterized by a sequence of the stick figure parameters are used as input of a neural network for classification. Fujiyoshi and Lipton [7] analyze the human motion by producing a star skeleton determined by extreme point estimation from the silhouette boundaries extracted. These cues are used to recognize human activities such as walking or running. Sappa et. al [19] develop a technique for human motion recognition based on the study of feature points' trajectories. Peaks and valleys of points' trajectories are first detected to classify human activity using prior knowledge of human body kinematics structure together with the corresponding motion model. In model-based approaches, the accuracy of human model reconstruction strongly depends on the quality of the extracted human silhouette in these model-based approaches. In the presence of noise, the estimated parameters may not be reliable.

Model-free approaches make no attempt to recover a structural model of human motion. Polana and Nelson [16] analyze human repetitive motion activity based on bottom up processing, which does not require the prior identification of specific parts. Motion activity is recognized by matching against a spatiotemporal template of motion features. Rajagopalan and Chellappa [17] develop a higher-order spectral analysis-based approach for detecting people by recognizing repetitive motion activity. The stride length is determined in every frame, and the bispectrum which is the Fourier transform of the triple correlation is used for recognition. Sarkar and Vega [20] discriminate between motion types based on the change in the relational statistics among the detected image features. They use the distribution of the statistics of the relations among the features for recognition. Davis [5] proposes a probabilistic reliable-inference framework to address the issue of rapid-and-reliable detection of human activities using posterior class ratios to verify the saliency of an input before committing to any activity classification..

3 Human Repetitive Activity Representation

Human repetitive activity is a cyclic motion where human motion repeats at a stable frequency. Assuming that the order of poses in a specific repetitive activity is the same among different people, it is possible to compose a spatiotemporal template in a single image instead of an ordered image sequence as usual. The fundamental assumptions made here are: (a) the order of poses in different cycles is the same, i.e., limbs move forward and backward in a similar way among normal people; (b) differences exist in the phase of poses in a motion cycle, the extend of limbs, and the shape of the torso, etc.
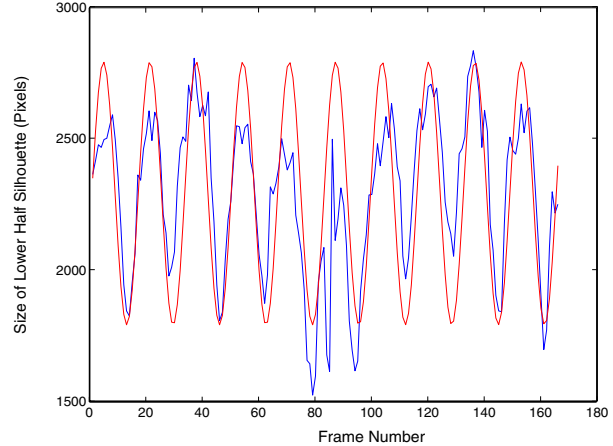

Figure 2. Frequency and phase estimation of human walking.

3.1 Human Silhouette Extraction

The raw silhouettes are extracted by a simple background subtraction method. This method is relatively easy to implement and requires less computational time. Once a Gaussian model of the background is built using sufficiently large number of frames, given a foreground frame, for each pixel, we can estimate whether it belongs to the background or foreground by observing the pixel value compared with the mean and standard deviation values at this pixel.

The raw silhouettes are further processed by size normalization (proportionally resizing each silhouette image so that all silhouettes have the same height) and horizontal alignment (centering the upper half silhouette part with respect to its horizontal centroid). In a so-obtained silhouette sequence, the time series signal of lower half silhouette part size from each frame indicates the motion frequency and phase information. The obtained time series signal consists of few cycles and lots of noise, which lead to sidelobe effect in the Fourier spectrum. To avoid this problem, we estimate the motion frequency and phase by maximum entropy spectrum estimation [13] from the obtained time series signal as shown in Figure 2.

3.2 Gait Energy Image

Given the preprocessed binary human silhouette images $B_t(x, y)$ at time t in a sequence, the grey-level gait energy image (GEI) is defined as follows

$$G(x, y) = \frac{1}{N} \sum_{t=1}^{N} B_t(x, y) \qquad (1)$$

where N is the number of frames in the complete cycle(s) of a silhouette sequence, t is the frame number in the sequence
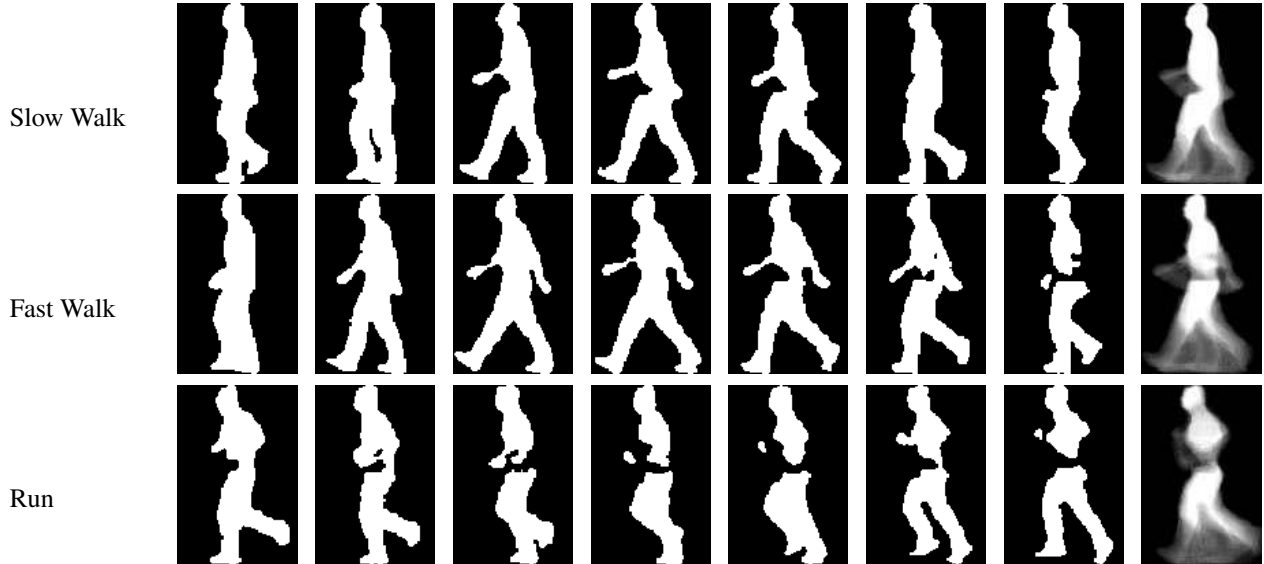
Figure 3. Examples of normalized and aligned silhouette frames in different human motion sequences. The rightmost image in each row is the corresponding gait energy image (GEI).

(moment of time), x and y are values in the 2D image coordinate. Figure 3 shows the sample silhouette images in a motion cycle of different activities of the same person, and the right most image is the corresponding GEI. As expected, GEI reflects major shapes of silhouettes and their changes over the motion cycle. We refer to it as gait energy image because: (a) each silhouette image is the space-normalized energy image of human walking at this moment; (b) GEI is the time-normalized accumulative energy image of human walking in the complete cycle(s); (c) a pixel with higher intensity value in GEI means that human walking occurs more frequently at this position (i.e., with higher energy).

Bobick and Davis [4] propose motion-energy image (MEI) and motion-history image (MHI) for human movement type representation and recognition.

MEI is a binary image which represents where motion has occured in an image sequence:

$$E_\tau(x,y,t) = \cup_{i=0}^{\tau-1} D(x,y,t-i), \qquad (2)$$

where $D(x,y,t)$ is a binary sequence indicating regions of motion, τ is the duration of time, t is the moment of time, x and y are values of 2D image coordinate.

MHI is a grey-level image which represents how motion in the image is moving:

$$H_\tau(x,y,t) = \begin{cases} \tau, & \text{if } D(x,y,t) = 1; \\ \max\{0, H_\tau(x,y,t-1) - 1\}, & \text{otherwise.} \end{cases} \qquad (3)$$

Both MEI and MHI are vector-images where the vector value at each pixel is a function of the motion properties at this location in an image sequence. As compared to MEI

and MHI, GEI targets specific repetitive activity representation, which has been successfully used in human identification by gait [10].

3.3 GEI Properties

In comparison with the activity representation by binary silhouette sequence, GEI representation saves both storage space and computation time for recognition and is less sensitive to silhouette noise in individual frames. Consider a noisy silhouette image $B_t(x,y)$ that is formed by the addition of noise $\eta_t(x,y)$ to an original silhouette image $f_t(x,y)$, that is, $B_t(x,y) = f_t(x,y) + \eta_t(x,y)$, where we assume that at every pair of coordinates (x,y) the noise at different moments t is uncorrelated and identically distributed. Under these constraints, we further assume that $\eta_t(x,y)$ satisfies the following distribution:

$$\eta_t(x,y) = \begin{cases} \eta_{1t}(x,y), & \text{if } f_t(x,y) = 1 \\ \eta_{2t}(x,y), & \text{if } f_t(x,y) = 0 \end{cases} \qquad (4)$$

where

$$\eta_{1t}(x,y): P\{\eta_t(x,y) = -1\} = p, \ P\{\eta_t(x,y) = 0\} = 1-p$$
$$\eta_{2t}(x,y): P\{\eta_t(x,y) = 1\} = p, \ P\{\eta_t(x,y) = 0\} = 1-p.$$

We have

$$E\{\eta_t(x,y)\} = \begin{cases} -p, & \text{if } f_t(x,y) = 1 \\ p, & \text{if } f_t(x,y) = 0 \end{cases} \qquad (5)$$

and

$$\sigma^2_{\eta_t(x,y)} = \sigma^2_{\eta_{1t}(x,y)} = \sigma^2_{\eta_{2t}(x,y)} = p(1-p). \qquad (6)$$

Given a walking cycle with N frames where $f_t(x,y) = 1$ at a pixel (x,y) only in M frames, we have

$$G(x,y) = \frac{1}{N}\sum_{t=1}^{N} B_t(x,y)$$
$$= \frac{1}{N}\sum_{t=1}^{N} f_t(x,y) + \frac{1}{N}\sum_{t=1}^{N} \eta_t(x,y)$$
$$= \frac{M}{N} + \bar{\eta}(x,y).$$

Therefore, the noise in GEI is

$$\bar{\eta}(x,y) = \frac{1}{N}\sum_{t=1}^{N} \eta_t(x,y)$$
$$= \frac{1}{N}\left[\sum_{t=1}^{M} \eta_{1t}(x,y) + \sum_{t=M+1}^{N} \eta_{2t}(x,y)\right].$$

We have

$$E\{\bar{\eta}(x,y)\} = \frac{1}{N}[\sum_{t=1}^{M} E\{\eta_{1t}(x,y)\}$$
$$+ \sum_{t=M+1}^{N} E\{\eta_{2t}(x,y)\}]$$
$$= \frac{1}{N}[M(-p) + (N-M)p]$$
$$= \frac{N-2M}{N}p$$

and

$$\sigma^2_{\bar{\eta}(x,y)} = E\{[\bar{\eta}(x,y) - E\{\bar{\eta}(x,y)\}]^2\}$$
$$= \frac{1}{N^2}E\{[\sum_{t=1}^{M}[\eta_{1t}(x,y) - E\{\eta_{1t}(x,y)\}]$$
$$+ \sum_{t=M+1}^{N}[\eta_{2t}(x,y) - E\{\eta_{2t}(x,y)\}]]^2\}$$
$$= \frac{1}{N^2}[M\sigma^2_{\eta_{1t}(x,y)} + (N-M)\sigma^2_{\eta_{2t}(x,y)}]$$
$$= \frac{1}{N}\sigma^2_{\eta_t(x,y)}.$$

Therefore, the mean of the noise in GEI varies between $-p$ and p depending on M while its variability ($\sigma^2_{\bar{\eta}(x,y)}$) decreases. If $M = N$ at (x,y) (all $f_t(x,y) = 1$), $E\{\bar{\eta}(x,y)\}$ becomes $-p$; if $M = 0$ at (x,y) (all $f_t(x,y) = 0$), $E\{\bar{\eta}(x,y)\}$ becomes p. At the location (x,y), the mean of the noise in GEI is the same as that in the individual silhouette image, but the noise variance reduces so that the probability of outliers is reduced. If M varies between 0 and N at (x,y), $E\{\bar{\eta}(x,y)\}$ also varies between p and $-p$.

Therefore, both the mean and the variance of the noise in GEI are reduced compared to the individual silhouette image at these locations. At the extreme, the noise in GEI has zero mean and reduced variance where $M = N/2$. As a result, GEI is less sensitive to silhouette noise in individual frames.

4   Repetitive Activity Recognition

In this section, we describe the proposed repetitive activity recognition approach using gait energy image. In the training procedure, GEI templates are generated from the original silhouette sequences. A component and discriminant analysis is them performed on the training templates for feature extraction. Human activity recognition is based on the extracted features.

Given a series of training GEI templates for each activity, the problem of their excessive dimensionality occurs. To reduce their dimensionality, there are two classical approaches of finding effective linear transformations by combing features - Principal Component Analysis (PCA) and Multiple Discriminant Analysis (MDA). As described in [6], PCA seeks a projection that best represents the data in a least square sense, while MDA seeks a projection that best separates the data in a least-square sense. Huang et al. [12] combine PCA and MDA to achieve the best data representation and the best class separability simultaneously. In this paper, the learning procedure follows this combination approach.

Given n d-dimensional training templates $\{\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n\}$, PCA minimizes the criterion function

$$J_{d'} = \sum_{k=1}^{n} ||(\mathbf{m} + \sum_{i=1}^{d'} a_{ki}\mathbf{e}_i) - \mathbf{x}_k||^2, \qquad (7)$$

where $d' < d$, $\mathbf{m} = \frac{1}{n}\sum_{k=1}^{n} \mathbf{x}_k$, and $\{\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_{d'}\}$ are a set of unit vectors. $J_{d'}$ is minimized when $\mathbf{e}_1$, $\mathbf{e}_2$, ..., and $\mathbf{e}_{d'}$ are the d' eigenvectors of the scatter matrix S having the largest eigenvalues, where

$$S = \sum_{k=1}^{n}(\mathbf{x}_k - \mathbf{m})(\mathbf{x}_k - \mathbf{m})^T. \qquad (8)$$

The d'-dimensional principal component vector $\mathbf{y}_k$ is obtained from the d-dimensional GEI template $\mathbf{x}_k$ by multiplying the transformation matrix $[\mathbf{e}_1, ..., \mathbf{e}_{d'}]$:

$$\mathbf{y}_k = [a_1, ..., a_{d'}]^T = [\mathbf{e}_1, ..., \mathbf{e}_{d'}]^T \mathbf{x}_k, \quad k = 1, ..., n. \quad (9)$$

where n is the number of the expanded GEI templates from all people in the training dataset.

Although PCA finds components that are useful for representing data, there is no reason to assume that these components must be useful for discriminating between data in

⊕ IEEE
COMPUTER
SOCIETY

| Noon | Late Afternoon | Night |

Figure 4. Thermal images of the background at different time in a day recorded from the thermal infrared sensor: noon, late afternoon and night, each of which is normalized by the temperature range individually.

different classes because PCA does not consider the class label of training templates. Multiple discriminant analysis (MDA) seeks a projection that are efficient for discrimination. Suppose that the n d'-dimensional transformed training templates $\{\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_n\}$ belong to c classes. MDA seeks a transformation matrix W that maximizes the ratio of the between-class scatter S_B to the within-class scatter S_W:

$$J(W) = \frac{|\tilde{S}_B|}{|\tilde{S}_W|} = \frac{|W^T S_B W|}{|W^T S_W W|}. \tag{10}$$

The within-class scatter S_B is defined as

$$S_W = \sum_{i=1}^{c} S_i, \tag{11}$$

where

$$S_i = \sum_{\mathbf{y} \in \mathcal{D}_i} (\mathbf{y} - \mathbf{m}_i)(\mathbf{y} - \mathbf{m}_i)^T \tag{12}$$

and

$$\mathbf{m}_i = \frac{1}{n_i} \sum_{\mathbf{y} \in \mathcal{D}_i} \mathbf{y}, \tag{13}$$

where $\mathcal{D}_i$ is the training template set that belongs to the ith class and n_i is the number of templates in $\mathcal{D}_i$. The within-class scatter S_B is defined as

$$S_B = \sum_{i=1}^{c} n_i (\mathbf{m}_i - \mathbf{m})(\mathbf{m}_i - \mathbf{m})^T, \tag{14}$$

where

$$\mathbf{m} = \frac{1}{n} \sum_{\mathbf{y} \in \mathcal{D}} \mathbf{y}, \tag{15}$$

and $\mathcal{D}$ is the whole training template set. $J(W)$ is maximized when the columns of W are the generalized eigenvectors that correspond to the largest eigenvalues in

$$S_B \mathbf{w}_i = \lambda_i S_W \mathbf{w}_i. \tag{16}$$

There are no more than $c-1$ nonzero eigenvalues, and the corresponding eigenvectors $\mathbf{v}_1,...,\mathbf{v}_{c-1}$ form transformation matrix. The $(c-1)$-dimensional multiple discriminant vector $\mathbf{z}_k$ is obtained from the d'-dimensional principal component vector $\mathbf{y}_k$ by multiplying the transformation matrix $[\mathbf{v}_1, ..., \mathbf{v}_{c-1}]$:

$$\mathbf{z}_k = [\mathbf{v}_1, ..., \mathbf{v}_{c-1}]^T \mathbf{y}_k, \quad k = 1, ..., n. \tag{17}$$

The obtained multiple discriminant vectors compose the feature database for activity recognition.

Assuming that the obtained feature vectors in each class are Gaussian distributed with the same covariance matrix, Bayesian classifier becomes minimum Euclidean distance classifier that is used for activity recognition.

5 Experimental Results

We have recorded real thermal image data of human activities by a FLIR SC2000 long-wave infrared camera in an outdoor environment. The image size is 240×320. The field-of-view of the camera is fixed during a human walking. Repetitive activities of five people are recorded at different time: noon (four people), late after noon (three people) and night (two people). The corresponding backgrounds are shown in Figure 4. Each background is normalized by the temperature range individually.

Each person was asked to slow walk, fast walk and run forward and backward along the fronto-parallel direction at each time. Therefore, there are totally a set of six sequences for each person at each time: two slow walking, two fast walking and two running. Four people are recorded at noon, three people are recorded at late afternoon, and 2 people are recorded at night. Three data sets recorded at noon are used for training, and other data sets are used for testing. Figure 5 shows GEI Examples of the 9 data sets (54 human motion sequences) used in our experiments. An observation from this figure is that the silhouette extraction performance at late afternoon is better than that at noon and night. This means that the temperature contrast between the human object and the background is larger at late afternoon. The mo-
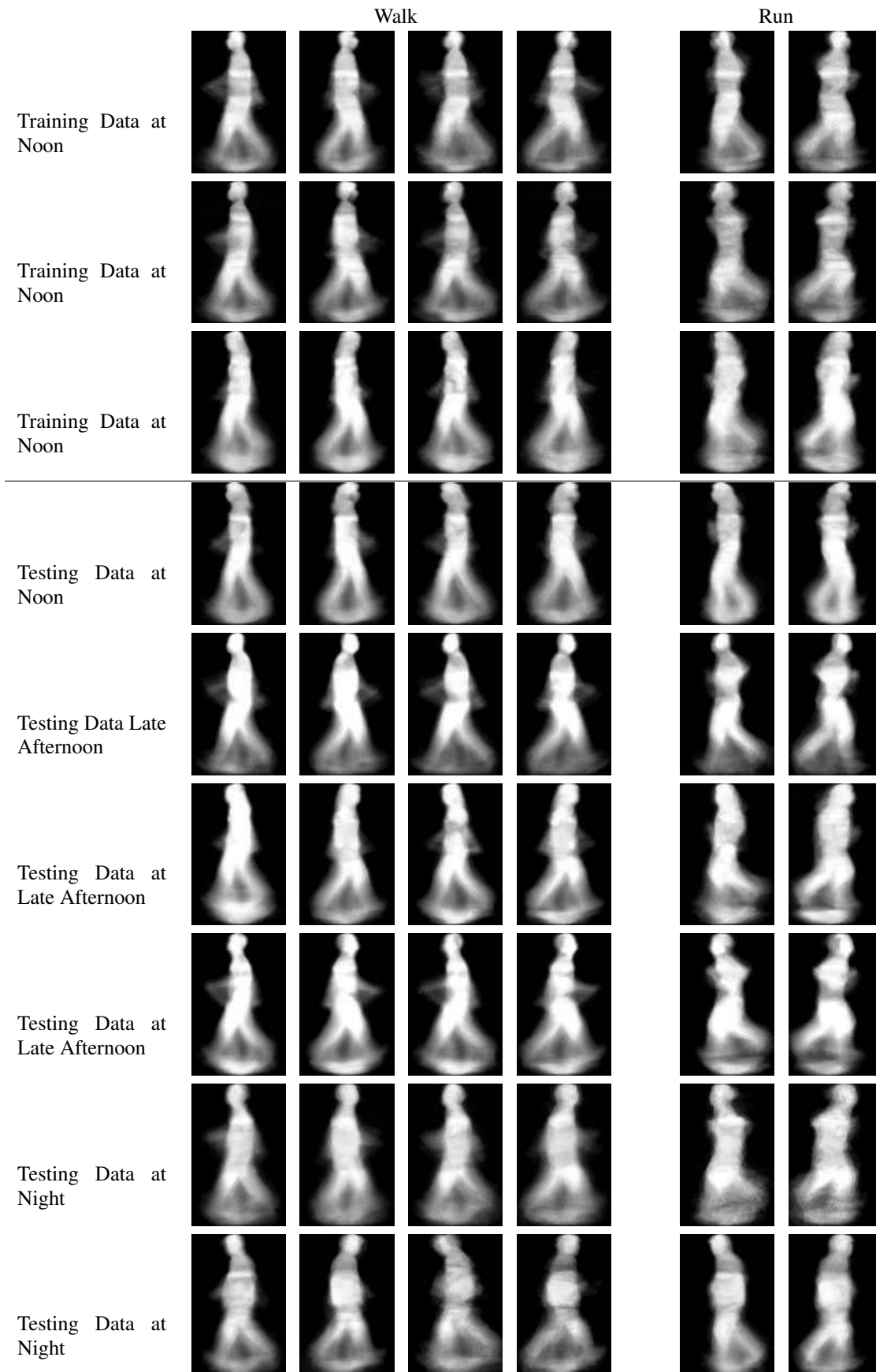
Figure 5. GEI Examples of the 54 human motion sequences used in our experiments.

tion of the trees in the background also contribute to the silhouette extraction performance in some frames.

The goal of our activity recognition here is to discriminate human walking or running regardless of their speed (slow or fast walking). In the recorded data, the speed in some fast walking sequences are equivalent or faster than that in some running sequences. Therefore, the speed is not appropriate for recognition walking or running. We employ the approach of combining PCA and MDA for feature extraction.

The size of the GEI template used in this paper is 128×88. Each GEI template is first converted to a column vector with $d = 11264$ dimensions as the input to PCA. The high dimension is then reduced to $d' = 4$ by PCA as shown in Equation (9). The 4-dimensional vectors are used as the input to the MDA, and the final feature vectors are of $c - 1 = 1$ dimension as shown in Equation (17), where $c = 2$ is the number of classes. The recognition performance on training data and testing data are all 100%. This demonstrates that the proposed approach achieves good performance for human repetitive activity recognition in the limited data set. Its performance will be further explored in a larger data set which is under construction.

6 Conclusions

In this paper, we use a spatio-temporal gait representation, called the Gait Energy Image (GEI), for human repetitive activity recognition. Human motion is detected in thermal infrared imagery, which provide good contrast between human objects and backgrounds regardless of lighting conditions and colors of the human surfaces and backgrounds. Unlike other motion representations which consider gait as a sequence of templates (poses), GEI represents human motion sequence in a single image while preserving temporal information. A statistical approach is used to extract features from GEI for activity recognition. Preliminary experimental results show that the proposed approach achieves good performance for human repetitive activity recognition.

References

[1] A. Abuelgasim and R. Fraser. Day and night-time active fire detection over north america using noaa-16 avhrr data. *Proc. IEEE International Geoscience and Remote Sensing Symposium*, 3:1489–1491, June 2002.

[2] L. Andreone, P. Antonello, M. Bertozzi, A., Broggi, A. Fascioli, and D. Ranzato. Vehicle detection and localization in infra-red images. *Proc. IEEE International Conference on Intelligent Transportation Systems*, pages 141–146, 2002.

[3] H. Arlowe. Thermal detection contrast of human targets. *Proc. IEEE International Carnahan Conference on Security Technology*, pages 27–33, 1992.

[4] A. Bobick and J. Davis. The recognition of human movement using temporal templates. *IEEE Trans. PAMI*, 23(3):257–267, March 2001.

[5] J. Davis. Sequential reliable-inference for rapid detection of human actions. *Proc. Conference on Computer Vision and Pattern Recognition*, pages 111–118, June 2004.

[6] R. Duda, P. Hart, and D. Stork. *Pattern Classification*. John Willy & Sons, 2000.

[7] H. Fujiyoshi and A. Lipson. Real-time human motion analysis by image skeletonization. *Proc. 4th IEEE Workshop on Applications of Computer Vision*, pages 15–21, 1998.

[8] G. Ginesu, D. Giusto, V. Margner, and P. Meinlschmidt. Detection of foreign bodies in food by thermal image processing. *IEEE Trans. Industrial Electronics*, 51(2):480–490, April 2004.

[9] Y. Guo and S. Tsuji. Understanding human motion patterns. *Proc. International Conference on Pattern Recognition*, 2:325–329, 1994.

[10] J. Han and B. Bhanu. Statistical feature fusion for gait-based human recognition. *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2:842–847, 2004.

[11] J. Holland and X.-H. Yan. Ocean thermal feature recognition, discrimination, and tracking using infrared satellite imagery. *IEEE Trans. Geoscience and Remote Sensing*, 30(5):1046–1053, Sept. 1992.

[12] P. Huang, C. Harris, and M. Nixon. Recognizing humans by gait via parameteric canonical space. *Artificial Intelligence in Engineering*, 13:359–366, 1999.

[13] J. Little and J. Boyd. Recognizing people by their gait: the shape of motion. *Videre: Journal of Computer Vision Research*, 1(2):1–32, 1998.

[14] P. Martinez, L. van Kempen, H. Sahli, and D. Ferrer. Improved thermal analysis of buried landmines. *IEEE Trans. Geoscience and Remote Sensing*, 42(9):1965–1975, Sept. 2004.

[15] I. Pavlidis, J. Levine, and P. Baukol. Thermal imaging for anxiety detection. *Proc. IEEE Workshop on Computer Vision Beyond the Visible Spectrum: Methods and Applications*, pages 104–109, June 2000.

[16] R. Polana and R. Nelson. Low level recognition of human motion (or how to get your man without finding his body parts). *Proc. IEEE Workshop on Motion of Non-Rigid and Articulated Objects*, pages 77–82, 1994.

[17] A. Rajagopalan and R. Chellappa. Higher-order spectral analysis of human motion. *Proc. International Conference on Image Processing*, 3:230–233, Sept. 2000.

[18] P. Riggan and J. Hoffman. Field applications of a multispectral, thermal imaging radiometer. *Proc. IEEE Aerospace Conference*, 3:443–449, March 1999.

[19] A. Sappa, N. Aifanti, S. Malassiotis, and M. Strintzis. Unsupervised motion classification by means of efficient feature selection and tracking. *Proc. International Symposium on 3D Data Processing, Visualization and Transmission*, pages 912–917, Sept. 2000.

[20] S. Sarkar and I. Vega. Discrimination of motion based on traces in the space of probability functions over feature relations. *Proc. IEEE International Computer Society Conference Computer Vision and Pattern Recognition*, 1:976–983, Dec. 2001.

[21] Y. Yoshitomi, T. Miyaura, S. Tomita, and S. Kimura. Face identification using thermal image processing. *Proc. IEEE International Workshop on Robot and Human Communication*, pages 374–379, 1997.

IEEE
COMPUTER
SOCIETY