Statistical and Entropy Based Human Motion Analysis

Chin-Poo Lee[1], Wei-Lee Woon[2] and Kian-Ming Lim[1]
[1]Faculty of Information Science and Technology, Multimedia University, Malaysia
[e-mail: {cplee, kmlim}@mmu.edu.my]
[2] Information Technology Program, Masdar Institute of Science and Technology, UAE
[e-mail: wwoon@masdar.ac.ae]
*Corresponding author: Chin-Poo Lee

Received July 13, 2010; revised September 3, 2010; accepted October 27, 2010; published December 23, 2010

Abstract

As visual surveillance systems gain wider usage in a variety of fields, it is important that they are capable of interpreting scenes automatically, also known as "human motion analysis" (HMA). However, existing HMA methods are too domain specific and computationally expensive. This paper proposes a general purpose HMA method that is based on the idea that human beings tend to exhibit erratic motion patterns during abnormal situations. Limb movements are characterized using the statistics of angular and linear displacements. In addition, the method is enhanced via the use of the entropy of the Fourier spectrum to measure the randomness of subject's motions. Various experiments have been conducted and the results indicate that the proposed method has very high classification accuracy in identifying anomalous behavior.

Keywords: Entropy, image processing, motion analysis, computer vision, neural networks

DOI: 10.3837/tiis.2010.12.012

1. Introduction

Existing visual surveillance systems operate by capturing, storing and distributing video clips, while the detection of incidents is largely left to human operators. Unfortunately, the dramatic growth in the number of installed surveillance cameras means that the resulting video streams are rarely monitored, if at all. Instead, their main use has been to provide forensic evidence after the occurrence of an incident. This situation emphasizes the importance of human motion analysis (HMA); of particular importance is the automatic analysis of video content so that only scenes which are likely security concerns are presented to human operators. Existing HMA approaches are often domain specific, and work by detecting specific activities which have to be determined in advance; examples include vandalism and elevator crime detection systems [1][2]. In addition, most of these approaches are computationally expensive or simply too complicated.

In view of the issues described above, this paper presents a general purpose method that can support a wide range of smart surveillance applications. The proposed method is based on the observation that videos depicting security-relevant incidents often depict non-rigid or erratic motion patterns. Such a situation might arise when the victim is struggling or trying to attract the attention of passers-by or when a delinquent is committing an act of vandalism. The technique described here seeks to extract motion patterns which are common to the abnormal situations described above.

2. Related Work

There are three major stages in HMA, namely motion segmentation, object tracking and behavior understanding. Each of these stages are inter-related and the accuracy of the preceding stage affects the performance and reliability of subsequent stages. In the following review, we only describe existing work that is related to each stage. Further details could be found in the original articles.

2.1. Motion Segmentation

This process aims to detect and segment moving regions from the rest of an image so that later processes only focus on these regions of interest. Most of the existing segmentation methods use either temporal or spatial information of the images. The motion segmentation methods studied in this work include background subtraction [3], statistical methods [4][5], temporal differencing [6] and optical flow [7][8].

2.2. Object Tracking

Tracking is particularly important to human motion analysis since it helps to prepare the input data for subsequent processing stages. Object tracking over time typically involves establishing articulate relations between image features in subsequent frames with respect to position, velocity, shape, color and texture. Main tracking approaches include model-based tracking [9][10][11], region-based tracking [4], contour-based tracking [7] and feature-based tracking [12][13][14][15].

2.3. Behavior Understanding

Behavior understanding takes place after the features have been tracked from one frame to the next. The extracted features are used to interpret and recognize motion patterns. Various methods have been suggested in previous studies, examples of which include Dynamic Time Warping [16], Hidden Markov Model [17][18][19], template matching [12] and neural network based methods[20][21][22].

3. Human Motion Analysis

Basically, the proposed system performs four main tasks, i.e. image preprocessing, motion segmentation, feature extraction and finally learning and classification. The overall system flow is depicted in **Fig. 1**.

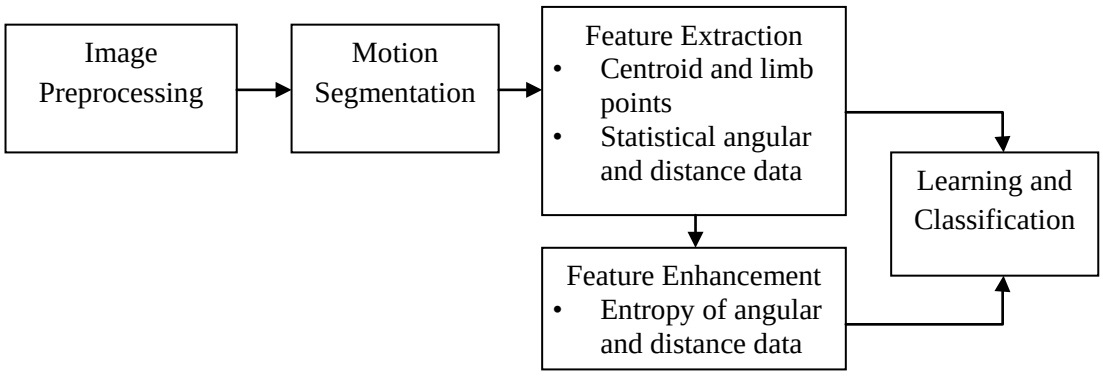


Fig. 1. System flow of the proposed method

3.1 Image Preprocessing

In this task, a video sequence is captured and stored as a series of grayscale image frames; these frames will subsequently be used to obtain motion data of objects present in the scenes. Grayscale images, also known as intensity images, are characterized by a variety of quantization levels; these range from [0, 1] for binary image arrays, to [0, 255] for 8-bit images all the way up to [0, 65535] for 16-bit images. For the purpose of this research project, 8-bit quantization was used to represent the image data. This bit depth is supported by most graphics file formats and the memory requirements are reasonable.

3.2 Motion Segmentation

Motion segmentation aims at finding the main motion patterns in the image sequences. In the proposed method, the background subtraction method has been used. However, the reference background is dynamically updated to adapt to changes in the brightness levels of the background.

After the difference images are obtained, thresholding is applied to help reject noise caused by shifting shadows and changes due to weather conditions. Difference values close to zero are likely to have resulted from minor background fluctuations while large difference values indicate that an object is present in the foreground. Specifically, Eq. (1) is used to create a binary map $I_{bi}(x,y)$ where the presence of ones is used to denote moving objects. **Fig. 2** shows the difference image before thresholding and after thresholding.

$$I_{bi}(x,y) = \begin{cases} 1 & \text{if} \quad I_d(x,y) > \tau \\ 0 & \textit{otherwise} \end{cases} \tag{1}$$

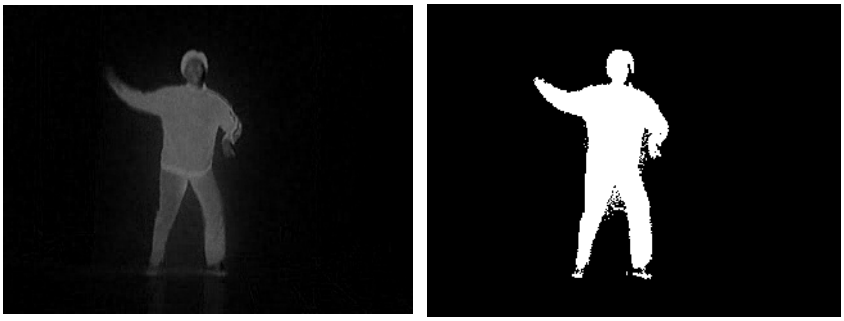


Fig.2. Difference image before thresholding(left) and after thresholding(right).

There will still be some noise remaining in the difference images, though this is often imperceptible. A morphological operation is used to further remove the remaining noise. This uses the majority rule in which a pixel is set to 1 if five or more pixels in its 3-by-3 neighborhood are 1's; conversely, a pixel with less than five 1's in its neighborhood will be set to 0. This rule is selected because it is straightforward and extremely efficient in comparison with other morphological approaches.

Hence, the output of the motion segmentation process would be these difference images where only the moving regions are highlighted. Next, in the feature extraction task, these moving regions will be represented in a form suitable for the description of motion patterns.

3.3 Feature Extraction

This section will be divided into 2 parts. The first part describes the detection of points of interest in the difference images while the second part explains the calculation of the angular and distance information corresponding to these points.

3.3.1 Determination of Centroid and Limb Points

In order to efficiently characterize the motion of the subject, it was decided to focus on the centroid and extremal points of limbs (hands and feet). The centroid of the human body is used as the reference point for all future calculations. This property ensures that the proposed method is independent of the position of the object in the scene.

The first step in this process is to determine the centroid of the moving region. The centroid is defined as the center of mass of a two dimensional planar lamina or a three dimensional solid [23]. The centroid, $\overline{x}$ of a set of n point masses m_i located at positions x_i is given by the following equation.

$$\overline{x} = \frac{\sum_{i=1}^{n} m_i x_i}{\sum_{i=1} m_i} \tag{2}$$

As a proof-of-concept, we assume that there is only one individual in the scene. Thus the largest region will be denoted as the object of interest. **Fig. 3** shows the segmented region corresponding to the individual and the centroid of this region.

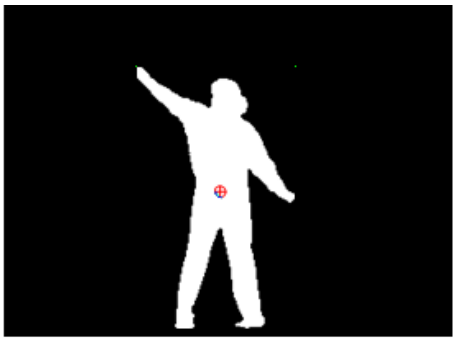


Fig. 3. Example image with centroid marked

After the centroid has been located, the next step is to trace the boundary or contour of the region. We start the boundary tracing from the highest point which is vertically aligned with the centroid. This point is selected because head is considered as the most static points among other points. Then, the detection of limb points will be based on the proportions of the boundary obtained. Some preliminary observations were made to establish the relationship between the boundary and the position of the limb points.

Fig. 4 provides an example of the calculations involved. As can be seen in the figure, the boundary vector contains 856 elements. The head point is defined by the 0-th element, while other limb points are defined by the 107-th, 321-th, 535-th and 749-th elements. The gaps between these elements are calculated. After that, the ratios of these gaps to the length of the boundary are computed.

The ratios for many images were calculated, from which it was determined that the outline of the human body can be approximately divided into eight equal portions. Each of these portions are listed below.

1. One portion from top of head to right hand (head).
2. Two portions from right hand to right foot (right hand).
3. Two portions from right foot to left foot (right foot).
4. Two portions from left foot to left hand (left foot).
5. One portion from left hand to top of head (left hand).

Head	Right Hand	Right Foot	Left Foot	Left Hand	Last element
0	107	321	535	749	856
Ratio:	1:8	2:8	2:8	2:8	1:8

Fig. 4. Example of the position of limb points in boundary vector

Fig. 5 depicts the partitioning of the body contour with each part labeled, while the corresponding points of interest are marked in **Fig. 6**. In the determination of limb points, we assume that all the people in the scenes are normal, i.e. with two hands and two legs. Though the method might be a little inaccurate if a non-standard figure is encountered (such as in the case of a handicapped individual), it would still be able to track his general form.

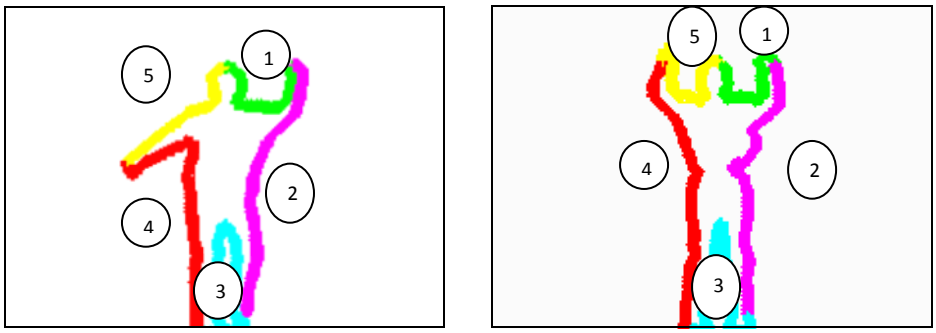


Fig. 5. Body contour divided into 5 portions

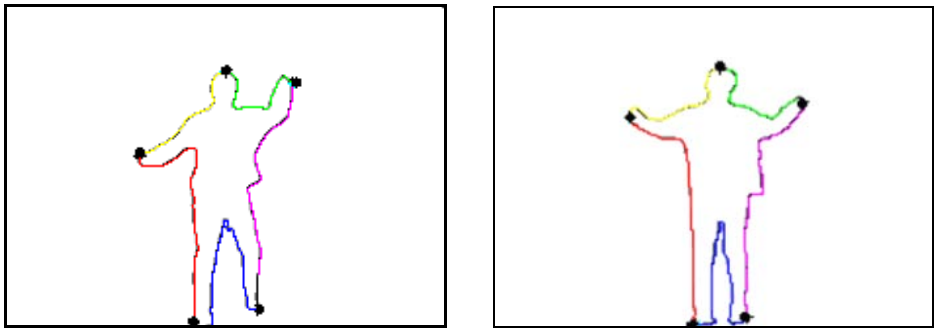


Fig. 6. Detected head and limb points

3.3.2 Angular and Distance Data Calculation

It is impractical to feed the raw angle and distance data into a neural network as there is unlikely to be sufficient information content in the data to justify such a high dimensional input space. Instead, relevant statistical quantities which characterize the angle and distance are extracted before it is fed into the neural network. As a starting point, common statistical measures such as the mean, variance and rate of change have been tested.

3.3.2.1 Angular Data

The relative angle is defined as the angle formed by the lines drawn from the extremal points to the centroid, and the horizontal axis. This is calculated using the equation below, where the relative angle is in the range of [-π, π] radians:

$$\theta_p = \tan^{-1}\frac{y_p - y_c}{x_p - x_c} \tag{2}$$

(x_p,y_p) is the coordinate of the point of interest, and (x_c,y_c) the coordinate of the centroid.

Next, the calculation of each of the basic statistical measures, namely the mean, variance and rate of change of the angle will be described. The mean angle is calculated over the length of a sliding window with n_f frames.

$$\bar{\theta} = \frac{\sum_{i=1}^{n_f}\theta i}{n_f} \tag{3}$$

Fig. 7 shows the motion sequence for the right hand and foot under normal and abnormal circumstances. The figure seems to show that the mean angle of the hand and foot is higher in abnormal situations.

-

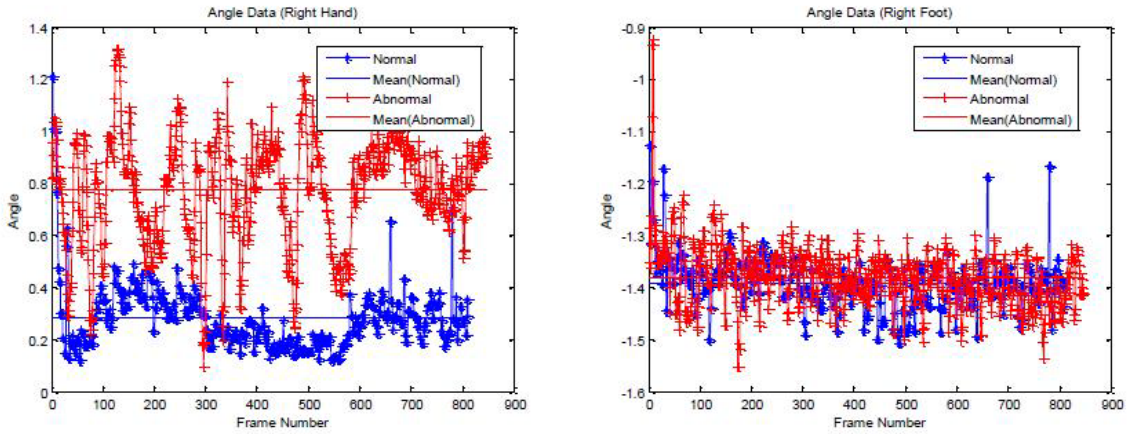


Fig. 7. Mean angle of **(a)** right hand and **(b)** right foot.

The bias-corrected sample variance is used to obtain an unbiased estimator for variance. It is given by the following equation.

$$s_\theta^{\,2} = \frac{\sum_{i=1}^{n_f} (\theta_i - \bar{\theta})^2}{n_f - 1} \tag{4}$$

The variance of the right hand and right foot are displayed in **Fig. 8**. Again, it should be noted that the angular variance is higher during abnormal situations when compared with normal ones.

Finally, the angular rate of change measures the speed of circular displacements in the motion. The rate of change, rc_θ in a sequence of n_f frames, is given by Eq. (5).

$$rc_\theta = \frac{Total\ magnitude\ of\ \ circular\ displacements\ in\ n_f\ \ frames}{n_f} \tag{5}$$

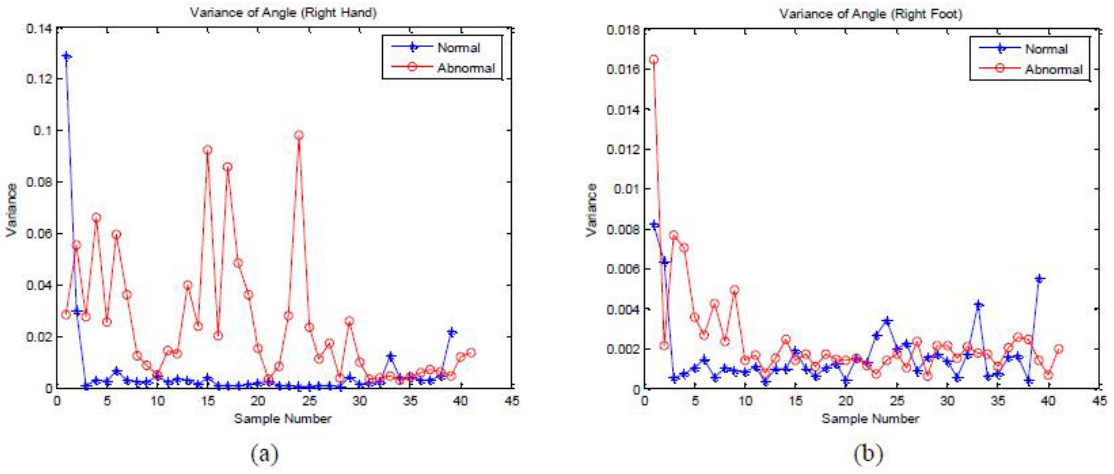


Fig. 8. Variance of angle of **(a)** right hand and **(b)** right foot

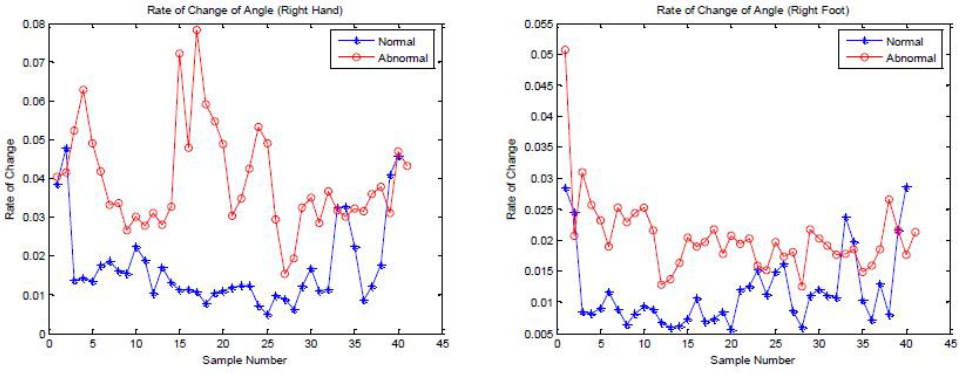


Fig. 9. The angular rate of change of right hand and right foot

As shown in **Fig. 9**, irregular movements tend to result in higher and more diverse angular rates of change, while normal motions have lower and more stable rates of change.

3.3.2.2 The Distance Measurement

The angular measurement is more focused on the circular displacement of the limbs while the distance measurement places more emphasis on the linear displacement of the limbs. The distances between the extremal points and the centroid indicate the extents to which the hands and legs are extended in motion.

Again, let (x_p,y_p) be the coordinate of the point of interest, (x_c,y_c) be the coordinate of the centroid, the distance of limb point to centroid, d_p is given by the following:

$$d_p = \sqrt{(x_p - x_c)^2 + (y_p - y_c)^2} \tag{6}$$

The mean distance is the average distance between the end points of limbs and the centroid of the human body. The mean distance is calculated by averaging the distance, d_i over the sliding window containing n_f frames, and is shown in Eq. (7). The motion sequences for the right hand and foot under normal and abnormal situations are shown in **Fig. 10-(a) and (b)** respectively.

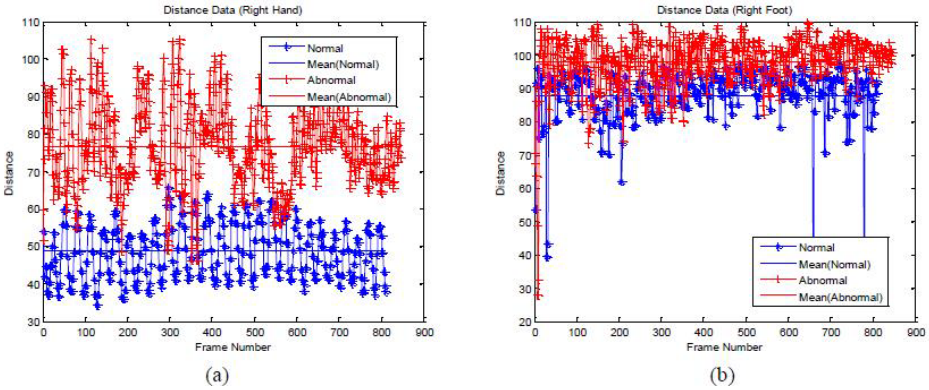
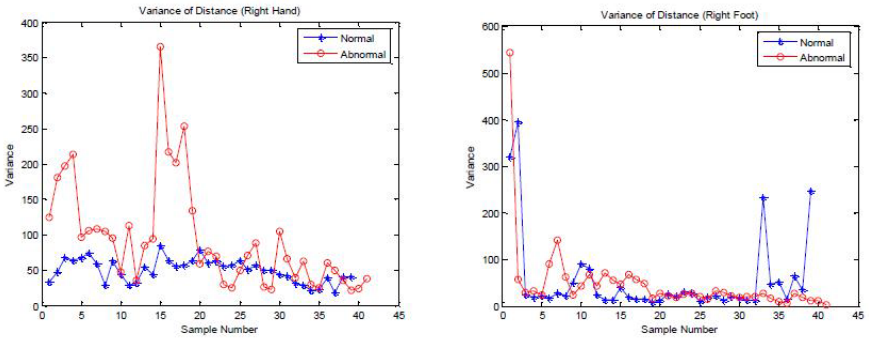


Fig. 10. Mean distance of **(a)** right hand and **(b)** right foot

$$\bar{d} = \frac{\sum_{i=1}^{n_f} d_i}{n_f} \tag{7}$$

The bias-corrected sample variance of distance, s_d^2 is computed using Eq. (8). **Fig. 11** shows the variance of the right hand and right foot respectively. It can be seen from these figures that in most cases, the degree of dispersion of distance during anomalous situations is higher than during normal situations.

$$s_d^{\,2} = \frac{\sum_{i=1}^{n_f} (d_i - \bar{d})^2}{n_f - 1} \tag{8}$$



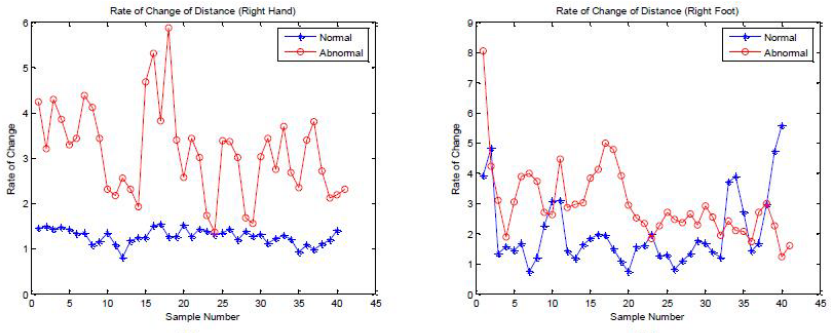
(a) (b)

Fig. 11. Variance of distance of **(a)** right hand and **(b)** right foot

The rate of change of distance is a measure of the speed with which the extremal points are moving. The rate of change, rc_d in a sequence of n_f frames, is given by the following equation.

$$rc_d = \frac{Total\ magnitude\ of\ linear\ displacements\ in\ n_f\ \ frames}{n_f} \tag{9}$$

Fig. 12 illustrates the rate of change for right hand and right foot in erratic and non-erratic motion. It can be clearly observed that the rate of change for irregular movements is more inconsistent compared to regular movements.



(a) (b)

Fig. 12. The rate of change of distance for **(a)** right hand and **(b)** right foot

3.4 Learning and Classification

In this task, a multi-layer perceptron (MLP) with a single hidden layer and sigmoidal activation functions is used. It is well known that a neural network with a sufficient number of hidden units is a universal approximator capable of fitting arbitrary functions. In addition, the use of sigmoidal hidden units means that small changes in the weights will usually result in large changes in the outputs, making it easier to determine if the weight change was a good or a bad one [24].

The MLP is optimized by minimizing the global error function with respect to the network weights. It is an iterative process in which an approximation to the error function in the neighborhood of the current point in the weight space is minimized [25]. A wide range of algorithms are available for performing the optimization. In this research, the scaled conjugate gradients method is used.

The scaled conjugate gradients algorithm was introduced as a way of avoiding the line search procedure of conventional conjugate gradients. In particular, every line minimization involves several error function evaluations, each of which is computationally expensive. The method of scaled conjugate gradients applies an optimization technique called model trust region method. Instead of line by line searching, this method performs the search using a variable step length, where this step length is only trusted in a small region around the current search point and its size is allowed to vary from one iteration to the next [26]. This method has been found to converge faster than the back propagation method and is computationally inexpensive. It offers a significant improvement in terms of speed compared to conventional conjugate gradient algorithms. More details of the algorithm can be found in [27].

4. System Implementation

The proposed system was implemented using Matlab 7.0 in combination with the Netlab neural network toolbox. To simplify the experiments, all test video clips were taken against a static white background, the webcam was set at an appropriate distance where the entire human body fits into the scene and all recordings starts with nobody in the scene.

The videos have a sampling rate of 15 frames per second and a resolution of 240 × 320 pixels. Some initial experimental trials were conducted to determine suitable parameter values, where the threshold value of 45 was found to be optimal and the image sequences are subdivided into sliding windows containing 60 frames and with 2-second overlaps.

As the system is only required to distinguish abnormal from normal behavior, only one output node is required in the neural network. A ‘1’ is used to denote normal behavior while a ‘0’ denotes abnormal behavior. Output values that are equal or greater than 0.5 are considered as 1’s, and vice versa.

5. Experimental Results

5.1 Results for Specific Scenarios

Each of the walking, running and struggling scenarios are captured separately in video sequences lasting 60 each. The classification accuracy is given by Eq. (10) and the overall classification accuracy is defined in Eq. (11). The results obtained from the experiments are given in **Table 1**.

$$Classification\ Accuracy = \frac{number\ of\ correct\ classification}{number\ of\ sample} \times 100\% \tag{10}$$

$$Overall\ Accuracy = \frac{total\ number\ of\ correct\ classification}{total\ number\ of\ sample} \times 100\% \tag{11}$$

Table 1. Results for specific scenarios.

Type of Scenario	No. of Sample	Classification		Classification Accuracy (%)	Overall Accuracy (%)
		Normal	Abnormal		
Walking	28	28	0	100	96.4
Running	28	25	3	89.3	
Struggling	27	0	27	100	

From the table, it was found that the overall classification accuracy for all the scenarios is good although there are some false alarms in the case of the running scenarios. However, achieving higher classification accuracy at the cost of a slightly elevated false alarm rate is tolerable in automated visual surveillance systems of this kind, since human operators will still be required to manually evaluate situations before any actions are taken.

5.2 Overall Classification Results

A total of 18 video sequences of approximately 30 seconds length were tested. The videos included various situations where the subjects were walking and running at different speeds as well as struggling and waving their hands to attract attention. The classification results are displayed in **Table 2**.

Table 2. Overall classification results.

Type of Behavior	No. of Sample	Classification		Classification Accuracy (%)	Overall Accuracy (%)
		Normal	Abnormal		
Normal	110	101	9	91.8	90.6
Abnormal	114	12	102	89.5	

The false negative rate, which is more than 10%, indicates that basic statistical quantities might not be sufficient to distinguish between more complicated movements. To further improve performance, a measure of complexity based on the information theoretic concept of entropy is presented.

6. Measure of Complexity: Entropy

In the previous sections, some basic statistical quantities and the classification accuracy obtained using those quantities were discussed. However, the experimental results show that approximately 10% of the scenarios with abnormal behavior went undetected. In addition, it is observed that the existing statistics are incapable of creating a clear separation between certain classes of normal and abnormal behavior. Examples of this are shown in **Fig. 13**, where it is noticeable in both figures that the statistics for normal and abnormal behavior are hardly distinguishable. Hence, it would be useful to adopt an additional metric to further

improve the classification performance of the system. Towards this end, we propose the use of a further statistic, entropy, to help address the shortcomings described above.

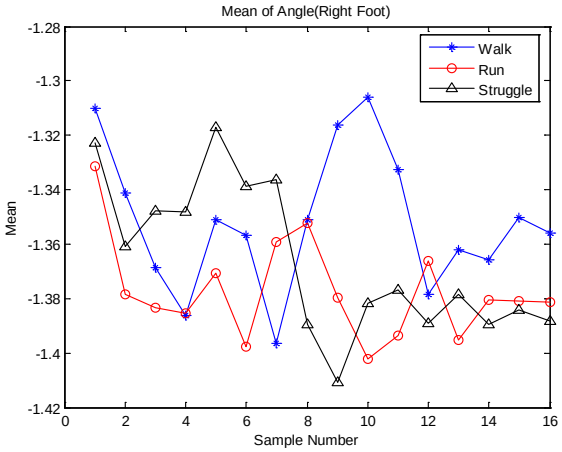


Fig. 13. The mean of angle for right foot

Generally, the entropy is a measure of how random a data set is [28]. To be more specific, consider the angular displacements of the limbs of a walking subject, encoded as a frequency spectrum. Since the motion of limbs is quite rhythmic, the frequency spectra of the angular displacements are likely to be dominated by a few large coefficients. This reflects a smaller amount of randomness or "complexity". Conversely, consider the case where the subject is struggling or attempting to attract attention. Since we cannot predict what the next value of angular displacement will be, it is to some degree, "random". Entropy is one possible way by which this randomness can be measured.

6.1 The Entropy of Angle and Distance

The computation of the entropy of both angle and distance, H is performed using the equation below.

$$H = \sum_{k=1}^{N} \ln[FT(k)] \tag{12}$$

where FT is the Fourier transform with N coefficients.

Fig. 14 illustrates the entropy for the angle of the right hand and left foot. It is noticeable that the larger the amount of disorder in the movement, the higher the corresponding value of entropy.

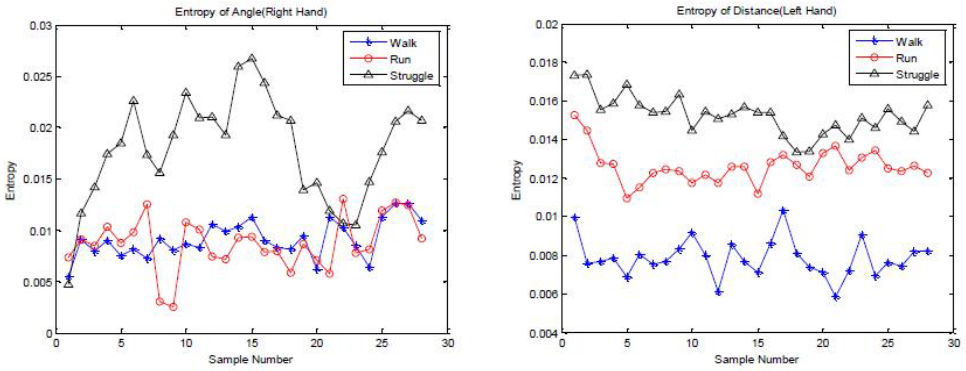


Fig. 14. The entropy of **(a)** angular displacement and **(b)** linear displacement

6.2 Entropy Based Experimental Results

This experiment is repeated on the same 18 video clips which were used previously. **Table 3** shows the classification results obtained from the entropy based method.

Table 3. Classification results for method with entropy.

Type of Behavior	No. of Sample	Classification		Classification Accuracy (%)	Overall Accuracy (%)
		Normal	Abnormal		
Normal	110	110	0	100	99.1
Abnormal	114	2	112	98.2	

The overall classification accuracy has increased from 90.6% to 99.1%. Thus, it can be concluded that entropy is a very good metric to reflect the randomness and complexity of the movements, which is often greater in the motions during abnormal situations.

7. Conclusion

Based on the results presented here, the classification results obtained using only the basic statistics are already encouraging; however, they also highlight the limitations of existing statistical measures in certain situations. Hence, a measure of randomness was used to improve the classification accuracy of the method.

Entropy, often used in information theory as a measure of randomness, is shown to be a potential descriptor of motion patterns. By incorporating the entropies of angle and distance data as additional metrics for characterizing anomalous behavior, it was possible to significantly improve the classification accuracy of the overall system.

References

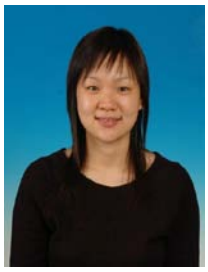
[1] C. Sacchi, C. Regazzoni and G. Vernazza, "A neural network-based image processing system for detection of vandal acts in unmanned railway environments," in *Proc. of the 11th International Conf. on Image Analysis and Processing*, pp. 529-534, Sept. 2001. Article (CrossRef Link)

[2] H. Shao, L. Li, P. Xiao and M. K. H. Leung, "Eleview: an active elevator video surveillance system," *Workshop on Human Motion*, pp. 67-72, Nov. 2000. Article (CrossRef Link)

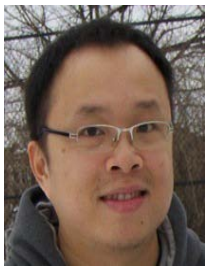
[3] I. Haritaoglu, D. Harwood and L. S. Davis, "W4: real-time surveillance of people and their

activities," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 809-830, Aug. 2000. Article (CrossRef Link)
[4] C. R. Wren, A. Azarbayejani, T. Darrell and A. Pentland, "Pfinder: real-time tracking of the human body," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 780-785, July 1997. Article (CrossRef Link)
[5] C. Stauffer and W. E. L. Grimson, "Adaptive background mixture models for real-time tracking," *IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, vol. 2, pp. 246-252, June, 1999. Article (CrossRef Link)
[6] R. T. Collins et al, "A system for video surveillance and monitoring," *VSAM Final Report, CMU-RI-TR-00-12, Technical Report*, Carnegie Mellon University, 2000. Article (CrossRef Link)
[7] D. Meyer and J. Denzler, "Model based extraction of articulated objects in image sequences for gait analysis," *International Conf. of Image Processing*, vol. 3, pp. 78-81, Oct. 1997. Article (CrossRef Link)
[8] H. A. Rowley and J. M. Rehg, "Analyzing articulated motion using expectation-maximization," *IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, pp. 935-941, June, 1997. Article (CrossRef Link)
[9] M. Silaghi, R. Plankers, R. Boulic, P. Fua and D. Thalmann, "Local and global skeleton fitting techniques for optical motion capture," in *Proc. of the International Workshop on Modeling and Motion Capture Techniques for Virtual Environments*, pp. 26-40, Nov. 1998. Article (CrossRef Link)
[10] S. X. Ju, M. J. Black and Y. Yacoob, "Cardboard people: a parameterized model of articulated image motion," *2nd International Conf. on Automatic Face and Gesture Recognition*, pp. 28-44, Oct. 1996. Article (CrossRef Link)
[11] J. M. Rehg and T. Kanade, "Model-based tracking of self-occluding articulated objects," *Fifth International Conf. on Computer Vision*, pp. 612-617, June, 1995. Article (CrossRef Link)
[12] R. Polana and R. Nelson, "Low level recognition of human motion (Or how to get your man without finding his body parts)," in *Proc. of IEEE Computer Society Workshop on Motion of Non rigid and Articulate Objects*, Austin, TX, pp.77-82, Oct., 1994. Article (CrossRef Link)
[13] J. Segen and S. G. Pingali, "A camera-based system for tracking people in real time," *13th International Conf. on Pattern Recognition*, pp.63-67, Aug. 1996. Article (CrossRef Link)
[14] C. Bauckhage, J. K. Tsotsos and F. E. Bunn, "Detecting abnormal gait," in *Proc. of the 2nd Canadian Conf. on Computer and Robot Vision*, pp. 282-288, May, 2005. Article (CrossRef Link)
[15] P. A. Beardsley, A. Zisserman and D. W. Murray, "Sequential updating of projective and affine structure from motion," *International Journal of Computer Vision*, vol. 23, no. 3, pp. 235-259, 1997. Article (CrossRef Link)
[16] A. F. Bobick and A. D. Wilson, "A state-based technique for the summarization and recognition of gesture," *Fifth International Conf. on Computer Vision*, pp. 382-388, June, 1995. Article (CrossRef Link)
[17] T. Starner and A. Pentland, "Real-time american sign language recognition from video using hidden Markov models," *International Symposium on Computer Vision*, pp. 265-270, Nov. 1995. Article (CrossRef Link)
[18] C. Vogler and D. Metaxas, "ASL recognition based on a coupling between HMMs and 3D motion analysis," *Sixth International Conf. on Computer Vision*, pp. 363-369, Jan. 1998. Article (CrossRef Link)
[19] M. Brand, N. Oliver and A. Pentland, "Coupled hidden Markov models for complex action recognition," *IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, pp. 994-999, June 1997. Article (CrossRef Link)
[20] H. Debar, M. Becker and D. Siboni, "A neural network component for an intrusion detection system," in *Proc. of Research in Security and Privacy, 1992 IEEE Computer Society Symposium on*, pp. 240-250, 4-6 May, 1992. Article (CrossRef Link)
[21] C. Sacchi, C. Regazzoni and G. Vernazza, "A neural network-based image processing system for detection of vandal acts in unmanned railway environments," in *Proc. of the 11th International Conf. on Image Analysis and Processing*, pp.529-534, Sept. 2001. Article (CrossRef Link)

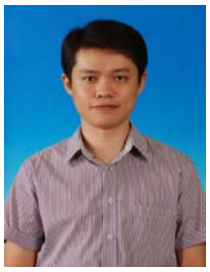
[22] J. Orozco and C. A. R. Garcia, "Detecting pathologies from infant cry applying scaled conjugate gradient neural networks," *11th European Symposium on Artificial Neural Networks*, Bruges, Belgium, pp. 349-354, Apr. 2003. Article (CrossRef Link)
[23] E. W. Weisstein, « Geometric Centroid," Retrieved Jan. 2, 2006, from http://mathworld.wolfram.com /GeometricCentroid.html, Jan., 2006. Article (CrossRef Link)
[24] W. S. Sarle, "Why Use Activation Functions," Retrieved May 11, 2006, from http://www.faqs.org/faqs/ai-faq/neural-nets/part2/section-10.html, June, 2004. Article (CrossRef Link)
[25] J. Orozco and C. A. R. Garcia, "Detecting pathologies from infant cry applying scaled conjugate gradient neural networks," *11th European Symposium on Artificial Neural Networks*, Bruges, Belgium, pp. 349-354, Apr. 2003. Article (CrossRef Link)
[26] C. M. Bishop, "Neural Networks for Pattern Recognition," New York: Oxford University Press, 1995. Article (CrossRef Link)
[27] M. F. Møller, "A scaled conjugate gradient algorithm for fast supervised learning," *Neural Networks*, vol. 6, no. 4, pp. 525–533, 1993. Article (CrossRef Link)
[28] A. Wagner and B. Plattner, "Entropy based worm and anomaly detection in fast IP networks," in *Proc. of the 14th IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprise*, pp. 172-177, June 2005. Article (CrossRef Link)



Chin-Poo Lee is a lecturer at Faculty of Information Science and Technology, Multimedia University, Malaysia. She received her Bachelor in Computer Science and MSc. in Information Technology, Multimedia University, Malaysia in 2004 and 2006, respectively. Her research interests include motion analysis, gait recognition and affective computing.



Dr. Wei Lee Woon received a B.Eng in Electronic Engineering with first class honors from UMIST, UK (now merged with the University of Manchester) in 1997, and his Ph.D in 2002 from the Neural Computing Research Group at Aston University, Birmingham, UK. Upon graduation he joined the Malaysia University of Science and Technology (MUST) as an Assistant Professor, where he served until 2007. He subsequently joined the faculty of the Masdar Institute of Science and Technology (Abu Dhabi, UAE).



Kian-Ming Lim is currently pursuing his MSc. in Information Technology, Multimedia University, Malaysia. At the same time, he is also working as assistant lecturer in Multimedia University, Malaysia. His research interests are in machine learning, classification and image processing.