An Analysis of Human Silhouettes with Normalised Silhouettes Images and Shape Fourier Descriptors

P.L. Wong[1] and Z. Abal Abas[2]

[1,2]Centre for Advanced Computing Technology (C-ACT), Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia.

Email: [2]zuraidaa@utem.edu.my

Abstract—**This paper aims to investigate the human silhouettes in video frames, which involves normalizing the silhouettes and describing the shape of the region in video frames using Shape Fourier Descriptors. Shape Fourier Descriptor describes the shape of an object by considering its boundaries, which are the shape centroid and calculated by a particular formula through all the video frames after normalized the videos. This shows the changes of the objects with various actions and can be recognized and characterized human or non-human in the video frames. Normalized Silhouette Image is significant before the videos are being processed to describe in descriptors. It focuses on the region based on the object's ratio in images of the shape of the object and silhouette images are centred after action segmentation. This reduces the burden of the process of extract unnecessary part in whole videos. Various human action videos and animal videos are used for the training and testing in this study to make sure the system performed better.**

Index Terms—**Human silhouette, shape Fourier descriptor, video.**

I. INTRODUCTION

NOWADAYS, vision-based in the field of the human movement analysis is growing based on a huge amount of the impressive application such as surveillance machine by visual and machine interface with a human approach to recognise human with the machines. Recognition of the gesture and behaviour of human from movement is needed when human performed some task or activity [1]. The gesture and behaviour of human are naturally presented through moving some part of the body to do a task such as running, walking, hand waving, bending and others. Human motion can involve many of the moving silhouettes which consider as the changes of silhouettes [2]. When a human motion is performed, it is a combination of individual silhouettes of a moving human in frames or images in a video. Hence, the feature extraction is needed to characterize and recognize human motions from videos, which can divide into many frames and images of human motion [3]. This is considered as Human Silhouette Recognition.

Silhouettes, as well as edges and contours are used to fit human body in images. This is important because most of the body poses information remains in its silhouette. A recognition has taken place to identify the human silhouette through some methods [4]. This human silhouette recognition can be used in Closed-circuit television (CCTV) to detect and identify the human motion when there is needed to avoid a crime happen in some situation, suitable to monitor the processes in industrial manufacturing or traffic used, make sure the transport in safe condition and others [5]. CCTV can be defined as a video surveillance which utilizing digital video recorders to

Article history: Manuscript received 31 August 2017; received in revised form 21 September 2017; Accepted 24 September 2017.

record the surroundings of the environment for a period of time whenever it is crowded or uncrowded for the particular area. Therefore, it is commonly used in many areas around the world. The recorded video on CCTV shows the movement of each human or non-human in the environment. Therefore, all the creatures' silhouettes in the video can be extracted by removing the background and noise in the video and track the body part of the human [6].

This paper is focusing on how Human Silhouette Recognition presented by using the Normalized Silhouette Image (NSI) and Shape Fourier Descriptor (SFD) which shown the characterization and recognition of human and non-human in videos. Normalized Silhouette Image is where the images are being normalized based on the object's ratio in images of the shape of the object and silhouette images are centred after action segmentation [7]. This normalized is done before the video is being processed to another stage. The images are not allowed to distort and without pre-processed. Shape Fourier Descriptor describes the shape of an object by considering its boundaries, which is the shape centroid, and calculate by a particular formula. It must be noted here that edge and silhouette as a part of geometrical concept [8]–[12] which play an important role fundamentally.

II. METHODOLOGY

Mathematical methods plays an important role in this work [13]–[17]. The method that involved in this project is shape-based method. It describes the details of silhouettes, which are moving objects such as a point, a blob and a box. It is the important details where the information will get from the system for easy define, label or even recognized the object. Therefore, it is ordinarily being classified as a standard to the recognition of pattern or object in a recognition case. Yet, the formulation of the human body in an image and having distinct from the point of views because many surfaces of body shape can be made, the difficulty of recognizing a moving human silhouette can make the recognition not accurate from other moving objects in an image or region with the method which is using shape-based.

III. RESULTS AND DISCUSSION

The deformation of silhouette is starting from normalized the videos then decomposed the videos into frames. A threshold value is needed when the image of the frames is converted into a binary image. Thresholding is an obvious way to extract the objects from the background which need a threshold value 0.3. It is based on the grey level to an image with the composed of foreground and background. If the value is lower than 0.3, the edge of the region will be counted as background and the area of the background to be extracted is larger. Then, the silhouette of the human is smaller. If the value is greater than 0.3, some background which is closer to the region will be counted as a part of the region and will not be extracted. Then, the silhouette of the human is larger. The manipulation of the value is necessary due to the grey level of the image.

The value setting for the system is presented by concluded the training video values. TABLE I shows the human silhouette value is always greater than 0.01 and the non-human silhouette value is always smaller than 0.01. Hence, it can be concluded that whenever the value is greater than 0.01, it implies that the video frame has a human silhouette. The minimum value of the mean of Fourier Descriptor that is obtained from the human silhouette is 0. It means that the human silhouette is similar with the training videos and the difference of the mean of Fourier Descriptors is excludes negative value. If the difference obtains a negative value, it will be processed by a modulus. In addition, the possibility of silhouette is a human is within the range of 0-0.0125. It is because the minimum value of the mean is 0 and the maximum value of the mean is 0.0125. Moreover, to gather the information from the below bar chart, the possible range is within 0-0.0125 from the bending motion, 0-0.04 from the jumping motion and 0-0.0209 is from the jacking motion. In addition, the jack motion is showing an increasing of frequency after the

value of 0.05. So, some fixed value is needed to be the condition value. Therefore, the mean value of 0, 0.0125, 0.0209, 0.04 is 0.01835 ≈ 0.02 is calculated. As a result, the difference of the mean of Fourier Descriptors is within 0-0.02 or above 0.05 and the test is greater than 0.01, which emphasize the range of specific region, is a human silhouette.

There are 486 results shown when overlapping comparison between 9 bending videos, 9 jumping videos and 9 jacking videos. The result shows 90.74% (441/486) accuracy of a human silhouette under the following condition. There are 27 non-human videos are comparable with 3 sets of 9 videos and it is 100% in 729 results are not being recognized as a human under below condition.

$$tdiff >= 0 \;\&\&\; tdiff <= 0.02 \;||\; tdiff > 0.05 \;\&\&\; test > 0.01 \quad (1)$$

TABLE I. MEAN OF FOURIER DESCRIPTORS OF TRAINING AND TESTING VIDEOS

	bend	pjump	jack	test
001	0.0187	0.1296	0.0806	0.0027
002	0.0182	0.0155	0.0174	0.0021
003	0.0117	0.0117	0.0118	0.0022
004	0.0158	0.0828	0.0320	0.0089
005	0.0162	0.1818	0.1018	0.0020
006	0.0242	0.0137	0.1073	0.0016
007	0.0155	0.0401	0.0115	0.0030
008	0.0140	0.0131	0.0124	0.0019
009	0.0171	0.0313	0.0113	0.0017

Fig.1. shows the comparison between 9 videos of human with a bending motion. The highest value of the mean of Fourier Descriptor is 0.0125, which fall in the class 0.0121-0.0140, and the lowest value of the mean of Fourier Descriptor is 0. The frequency shows that 33 numbers of the difference is falling under the class between 0-0.0020. 18 numbers of the difference are in the range of 0.0021-0.0040. It indicates that most of the human silhouette range is within 0-0.0040. In addition, the frequency is decreasing as the class increasing when human pose is a bending. It shows in the pie chart above which 100% of the human silhouette is successfully recognized.

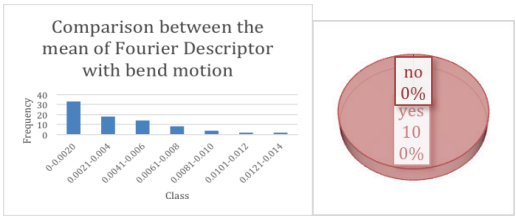

Fig. 1. Difference of mean FD between videos

Fig. 2. shows the comparison between 9 videos of human jumping motion. The highest value is 0.1701 which falling under the range of 0.161-0.2 and the lowest value is 0. The frequency show that most of the value is falling under the class 0-0.04 is 39. 31/39 of the value is below 0.02 within the range of 0-0.04. The pie chart shows the 85% is human silhouette and 15% is not human silhouette.

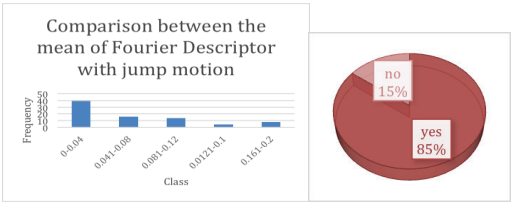

Fig. 2. Comparison between the mean of Fourier Descriptor with jump motion

Fig. 3. shows the comparison between 9 videos of human jacking motion. The highest value is 0.0960 in the class 0.081-0.1 and the lowest value is 0. The frequency shows that most of the value if falling under class 0-0.0209 is 41. The other range between 0.05 and above is considered as human silhouette. The pie chart shows the 85% is human silhouette and 15% is not human silhouette.

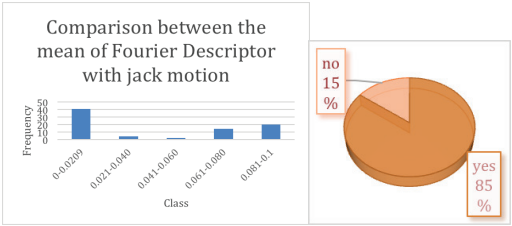

Fig. 3. Comparison between the mean of Fourier Descriptor with jack motion

IV. CONCLUSION

In conclusion, this paper investigates the recognition of a human silhouette with Normalized Silhouette Image and Shape Fourier Descriptors. The human silhouette with motion play an important role to show the comparison of the difference between the motions and produced a satisfied result towards the testing videos. A better performance is shown by calculating the mean of descriptors frame by frame in every videos, which produce information for further analysis.

ACKNOWLEDGMENT

Sincere thanks to Universiti Teknikal Malaysia Melaka (UTeM) for granting permission to publish the report.

REFERENCES

[1] C. L. Devasena, R. Revathí, and M. Hemalatha, "Video Surveillance Systems – A Survey," Int. J. Comput. Sci. Issues, vol. 8, no. 4, pp. 635–642, 2011.

[2] M. Sivarathinabala and S. Abirami, "An Intelligent Video Surveillance Framework for Remote Monitoring," Int. J. Eng. Sci. Innov. Technol., vol. 2, no. 2, pp. 297–301, 2013.

[3] A. H. Nasution and S. Emmanuel, "Intelligent Video Surveillance for Monitoring Elderly in Home Environments," Multimedia Signal Processing, 2007. MMSP 2007. IEEE 9th Workshop on. pp. 203–206, 2007.

[4] Z. A. Abas et al., "THE IMPLEMENTATION OF PERSONAL MONITORING SYSTEM USING HISTOGRAM OF GRADIENT, HAAR CLASSIFIER AND EIGENFACES FOR HUMAN DETECTION AND RECOGNITION," ARPN J. Eng. Appl. Sci., vol. 10, no. 19, pp. 8799–8806, 2015.

[5] Z. A. Abas, L. Ee-Theng, A. F. N. A. Rahman, Z. Z. Abidin, and A. S. Shibghatullah, "Enhanced scheduling traffic light model using Discrete Event Simulation for improved signal timing analysis," ARPN J. Eng. Appl. Sci., vol. 10, no. 18, pp. 8135–8140, 2015.

[6] Z. Zhou et al., "A real-time system for in-home activity monitoring of elders," Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE. pp. 6115–6118, 2009.

[7] S. F. Abdullah, A. F. N. A. Rahman, and Z. A. Abas, "CLASSIFICATION OF GENDER BY USING FINGERPRINT RIDGE DENSITY IN NORTHERN PART OF MALAYSIA," ARPN J. Eng. Appl. Sci., vol. 10, no. 22, pp. 10722–10726, 2015.

[8] Z. Abal Abas, S. Salleh, and Z. Manan, "Extended Advancing Front Technique for the Initial Triangular Mesh Construction on a Single Coil for Radiative Heat Transfer," Arab. J. Sci. Eng., vol. 38, no. 9, pp. 2245–2262, 2013.

[9] Z. A. Abas and S. Salleh, "Enhanced Advancing Front Technique with Extension Cases for Initial Triangular Mesh Generation," in Proceedings of the World Congress on Engineering 2011, 2011.

[10] A. S. H. Basari, Z. A. Abas, and Shaharuddin Salleh, "Seven Cases Unstructured Triangulation Technique for Simplified Version of Conceptual Model of Ethylene Furnace for Radiative Heat Transfer Approximation," Int. J. Comput. Appl., vol. 44, 2012.

[11] S. Salleh and Z. Abal Abas, Simulation for applied graph theory using Visual C++. Boca Raton: CRC Press Taylor & Francis Group, 2016.

[12] Z. A. Abas et al., "An insight into the science of unstructured meshes in computer numerical simulation," Appl. Math. Sci., vol. 8, no. 171, pp. 8493–8516, 2014.

[13] N. F. Mansor, Z. Abal Abas, A. F. N. Abdul Rahman, A. S. Shibghatullah, and S. Safiah, "An analysis of the parameter modifications in varieties of harmony search algorithm," Int. Rev. Comput. Softw., vol. 9, no. 10, pp. 1736–1749, 2014.

[14] Z. A. Shaffiei, Z. A. Abas, A. F. N. A. Rahman, A. S. Shibghatullah, and B. Hussin, "OPTIMIZATION IN DRIVER's SCHEDULING FOR UNIVERSITY SHUTTLE BUS USING HARMONY SEARCH," Int. Symp. Res. Innov. Sustain. 2014, vol. 2014, no. October, pp. 15–16, 2014.

[15] Z. Abas, Z. Shaffiei, A. F. N. A. Rahman, and A. S. Shibghatullah, "Using Harmony Search for Optimising University Shuttle Bus Driver Scheduling for Better Operational Management," in ITMAR 2014, 2014, vol. 1, pp. 614–621.

[16] M. I. Jasmi, A. F. N. A. Rahman, Z. A. Abas, and A. S. Shibghatullah, "Optimized Coating Design of Energy Saving Glass Using Binary Harmony Search for Better Transmission Signal," Int. J. Comput. Sci. Inf. Secur., vol. 14, no. 8, pp. 436–443, 2016.

[17] Z. A. Shaffiei, Z. A. Abas, A. S. Shibghatullah, and A. F. N. A. Rahman, "An Optimized Intelligent Automation for University Shuttle Bus Driver Scheduling Using Mutual Swapping and Harmony Search," vol. 14, no. 8, pp. 875–884, 2016.