Journal of Engineering and Applied Sciences 11 (3): 446-451, 2016
ISSN: 1816-949X

Aggressive Behaviour Analytics: A Brief Review

Nur Syazwani Samanu, Mohd Asyraf Zulkifley Nor Afiqah Ramdun and Aini Hussain
Department of Electrical, Electronic and Systems Engineering,
Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia,
UKM, 43600 Bangi, Selangor, Malaysia

Abstract: This study provides a brief review on human behaviour analytics system used in video surveillance applications. The topic of aggressive human behaviour detection system is chosen because of its rising popularity in the field of computer vision, specifically for video surveillance applications. Most of the existing systems were designed for closed room application such as detention room or prison to prevent any anxious situations that may lead to violence activities. Furthermore, the advancement of video surveillance system has allowed safety features to be embedded so that human activities both in public and private places can be monitored and analysed. Hence, many researchers have developed an intelligent system to study human behaviours to differentiate between normal and abnormal behaviours with various computer vision methodologies. Therefore, the main objective of this study is to explore the related researches that have been done in automatic aggressive behaviour detection system to observe a small crowd behaviours for real-time video surveillance application. The review conludes that there are many commonly and favourable methods for study about human behaviours by using video applications and describes the purpose of highlighted methods of each researches.

Key words: Aggressive behaviour, behavioural analytics, optical flow, video surveillance system, video

INTRODUCTION

Nowadays, information taken from video has become the most popular medium used in education, security and entertainment. With the advancement of network and communication technology, users can easily upload and download videos to their social networking platforms such as Facebook, Twitter and YouTube. Since, the previously mentioned platforms are all electronic in nature, they have the capability to perform video recording, copying, playback, broadcasting and display of moving visual data (Murthy *et al*., 2013). By definition, a video is a grouped of sequential images that functions is to produce moving images with addition of audio component.

Furthermore, researchers have proven that video provides more information and conveys more meaningful visual impact to the users. It is more convincing communication mode compared to other types because of the influence of human brain on emotional input to make decisions or actions. This is because of the fact that video caters to both auditory and visual message to the brain that may include facial expression, music and body language. Moreover, an infographic from researcher showed that more people remember what they have seen and done compared to just what they have heard and read (Blasch, 2009).

Besides, video provides a proof of an event that include the details of the movement that can be viewed in various ways such as slow motion or stop action which cannot be done by normal human observation alone and they can examine the events repeatedly.

Therefore, intelligent video surveillance is becoming an essential system in order to monitor and detect suspicious situations in real-time. Most of existing video surveillance systems are based on analogue cameras that are connected throughout coaxial cable network to Digital Recorders (DVRs) or tape recorders located in the server room. Hence, the captured video from these analogue cameras are usually poor in quality with little to no signal processing is applied. This may cause critical moment in the scene of interest is easily missed and put a high burden on the safety of the observer (Appiah *et al*., 2009).

Due to the technology capabilities nowadays, Internet Protocol (IP) video surveillance system is becoming the new trend since it is more powerful, flexible

Corresponding Author: Nur Syazwani Samanu, Department of Electrical, Electronic and Systems Engineering, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia, UKM, 43600, Bangi, Selangor, Malaysia

and can be applied to capture wider area. IP camera is connected via internet or local network which is capable of supplying high resolution image for behavioural analytics system to notice any abnormal event in the video scenes (Ju *et al*., 2015). Thus, it will decrease the workload of safety personnel and law agencies since it enables them to concentrate more on questionable or suspicious scenarios (Kim *et al*., 2011).

Physical action is an intricate, multi-faceted and broad ranging that covers various behaviours. It may cover daily activities of an individual such as housework or stair climbing, occupational related activities like packing and lifting, transportation activity such as walking or cycling and daily exercise or hobbies (Stratton *et al*., 2012). However, someone may act roughly or aggressively when something happened not according to his desire where the person might hit, push or throw things to express the feelings (Nirjon *et al*., 2014).

Aggressive behaviour is a bad behaviour that damages the society and leads to violation of social rules and rise the conflicts between individuals. The situation may get worse if the aggressor acts unpleasantly that may lead to fighting and fatal accident (Murata *et al*., 2012). Therefore, the main concern of this study is to review on intelligent detection system of aggressive human behaviour where the actions to be detected have been well-defined so that unexpected events can be quickly avoided.

BEHAVIOURAL ANALYTICS SYSTEM

Video can be defined as a combination of sequential images that forms moving pictures with added audio element that matched the pictures being shown on the video display. Recently, the demand for video applications in real-time system has grown rapidly with the advancement of network and multimedia technology (Zulkifley *et al*., 2016). Some examples of the application are video surveillance system, video tutorial and video conferencing (Zhen and Su, 2010). Hence, only real-time video surveillance systems are considered in this study that focuses on gaining information about human behaviours for security purpose so that the workload of security personnel will be lighten (Schneiders *et al*., 2015).

Moreover, data from video surveillance system has been successfully used for safety monitoring either at public or private places. The system is normally installed at banks, office buildings, prison, shopping malls, airports and other places (Pang, 2014; Hou *et al*., 2013). Besides, this system can capture images that sometimes even the security personnel cannot see with the naked eye such as high speed movement or stop motion. In addition, video surveillance can be utilised to supervise human activities such as to monitor menace behaviours and to prevent terror and crime. In some cases, video information has been used to identify people with suspicious behaviour so that alert signal to unexpected event can be transmitted to maintain social harmony with just one camera (Ezzahout and Thami, 2013).

Table 1: Summary of the advantages and challenges of using video surveillance system

Advantages	Challenges
Use IP network to provide real-time video	Analogue video
To monitor suspicious situations	Poor quality and fuzzy
To prevent crime	Used many cameras
High-definition video	False alarm
Widely installed	
Low burden of safety worker	

However, most of the existing surveillance systems are based on analogue camera that are placed in various different locations and connected to control room for recording and monitoring (Moon *et al*., 2009). Besides, this analogue security system is less effective since the resolution is comparatively low that hinders or limits data security, storage and recovery and other automated processing. Thus, the technology evolution of video surveillance system is moving towards digital era which is based on IP networking (Lim *et al*., 2006).

An intelligent detection system using IP camera is effective since it can be connected to different computing points or mobile devices through IP network. Moreover, IP network is more secured where data security is more guaranteed which is highly recommended in social security applications through encryption and confirmation methods. Furthermore, most of the systems come with High-Definition (HD) videos up to 720×1080 with less blur which is more suitable for video analytics application (Chan *et al*., 2008). Therefore, criminal activities can be detected and apprehended more easily apart from preventing any possible crimes due to technology advancement in monitoring public and private places. Table 1 shows the summary of the advantages and challenges of using video surveillance system.

Aggressive human behaviour: Human behaviour can be described as the physical conduct and action. Usually, person's personality and characteristics are consistent but it may change due to several factors; knowledge, expectations and attitudes, environmental influences; social norms, influence by others and behavioural factors, self-efficacy and skills as:

Cognitive factors:

- Knowledge
- Expectations
- Attitudes

Environmental factors:
- Social norms
- Access in community
- Influence on others (ability to change own environment)

Determines human behaviour
Behavioural factors:
- Skills
- Practice
- Self-efficacy

Dys functional social action has been associated with the person who suffers from depression, frustration and post-traumatic stress disorder. These situations may lead to aggressive behaviour that can be closely linked to several factors such as family structure research and school environment, relationship with others, health condition, life experience, individual characteristics and psychiatric issues.

Besides, a person with aggressive behaviour tends to be impulsive, restless, violence and irritable person. Furthermore, they may use physical force that might injure other persons and property. Studies in Theodoridis and Hu (2013), Chen *et al.* (2008) and Ouanane and Serir (2013) have addressed aggressive human behaviour into a few actions. For example, boxing, running, jogging, hand clapping and waving, pushing, kicking, slapping and hammering where the last four actions are defined as aggressive behaviour.

In recent years, there are many researches have been done on human behaviour detection by using video information. The rising interest in human behaviour analysis is due to increased demand in surveillance of people activities, group activity detection, restricted-area access detection, behaviours recognition and trajectory scenario.

Microsoft kinect sensor: Pang (2014) implemented an algorithm using the Microsoft Kinect sensor for aggressive human behaviour detection since it offers a synchronous RGBD (Red, Green, Blue and Depth) images that can be applied for object tracking and recognition, hand gesture analysis and human activity analysis. This Kinect sensor comprises of RGB camera, 4 microphone arrays and depth sensor. It is less sensitive to illumination variation. They have analysed audio-video data captured from Microsoft Kinect by using Microsoft Visual C++ 2010 Express in order to implement a better detection of aggressive human behaviour in various kinds of situation.

They used Cartesian coordinate point-plane distance formula in order to compute human body joints position, easily detected by Microsoft Kinect sensor. Their algorithm is capable to distinguish a human falling down action, 5 basic human behaviours that include walking, crouching, running, jumping and sitting and 2 aggressive human behaviours such as kicking and punching and also some random behaviours. For instance, standing rapidly from crouching position, waving hand and hand shake.

Their system was tested with video data of 30 persons of dissimilar gender, height and width. The experimental setup was carried out within a range of 1-4 m from the sensor with varying angles from 180°-180° where the optimal distance of Kinect sensor is between the ranges of 3-4 m. Hence, the system managed to obtain 95.83% of average accuracy for punching posture and almost 98% of average accuracy for normal behaviour and kicking posture.

From this study, the reachers considered to use Kinect sensor because of the sensor ability in getting the joints angle information with better accuracy of experimental results. Besides, Kinect sensor is able to develop with C++, C or visual basic by using Microsoft Visual Studio. So that, the camera of Kinect sensor can be controlled by both software and hardware. Moreover, this modern camera can recognize moving objects and distinguish human body parts and their individual faces.

Trajectory feature and regional optical flow: Generally, human behaviours can be divided into two categories which are normal such as wandering and marching and abnormal behaviour such as acts of violence and fight. Hou *et al.* (2013) proposed an abnormal behaviour recognition based on trajectory feature and regional optical flow characteristics. A modified Gaussian mixture model is used for background modelling which aims to extract the moving foreground like human movement using background subtraction method. They have applied mathematical morphology for pre-processing that includes expansion, erosion, closing and opening operations.

Next, foreground region was labelled by using 8-connected neighbourhood area to obtain the regional centre of the trajectory. For the next extraction and discrimination steps, the connected areas on the boundaries are removed for pixels less than the selected threshold. Then, the least enclosing rectangle was extracted to calculate the centre point of motion region which was used for further analysis to recognize between normal and abnormal behaviour characteristics.

The Lucas-Kanade algorithm was implemented to extract the optical flow information within the moving region. The regional flow features were identified by using histogram with the weighted amplitude direction in order to differentiate the subject state. The system was tested

with 30 video sequences that include 12232 frames by using MATLAB platform. The experimental results proved that their technique is able to distinguish accurately between common abnormal and normal behaviours with false alarm rate of 5.06% and success rate of recognition of abnormal behaviour of 87.28%.

The main idea of the study is to differentiate between normal and abnormal behaviours by using trajectory characteristics and regional flows. The trajectory will have closed curve when the person is normally walking towards certain direction compared to the trajectory of person who is fighting. Furthermore, the range of motion type for abnormal behaviour is wider where the movement velocity is faster. On the other hand, the range of motion type for normal behaviour is small with low average speed.

Bag-of-feature approach and skeleton graph: In the research of Ouanane and Serir (2013) two approaches were evaluated in order to recognize aggressive human behaviour from video information which are bag of features approach for shape representation and the used of skeletonization-based method for feature extraction. There are three main modules in their algorithm which are background subtraction, feature extraction and classification. The movement such as jogging, running and walking indicate the non-aggressive behaviour while boxing represents the aggressive behaviour.

The first module is background subtraction that used sequential kernel density approximation which aims to remove the silhouette from image sequence. Next, the second module is feature extraction which is used to generate an appropriate feature vectors. In this module, they used two approaches which are bag-of-feature for shape representation and the second approach is based on Skeleton graph model.

The bag-of-feature approach is conducted by a set of interest operators and shape parameters including blob size, circularity, dispersion, orientation aspect ratio and centre of gravity. Next, the Skeleton graph model is used to extract region-based shape feature of an object by combining distance map and thinning algorithm.

After that, offline k-means algorithm is utilised to create an optimum codebook for each feature vector. The size of the codebook and the size of feature vectors for both shape representations are utilised to avoid over-fitting issue in classification process. Lastly, SVM classifier is used to distinguish the aggressive behaviour from the normal behaviour. The experimental results proved that the effectiveness of the method with 96% of accuracy rate.

In brief, they implemented bag-of-feature approach since this method can represents video sequence as unordered sets of local space-time characteristics which are quantized into codebooks or discrete vocabularies. Furthermore, this approach is followed by classification method that has been used to conduct the classification. Besides, bag-of-feature has been frequently used for human action recognition and categorization due to its good performance and simplicity. While, skeleton graph is applied because of its capability to allow localization of the coordinates of each human body parts in the form of vectors. Consequently, it provides an efficient shape descriptor for image recognition in this system.

Principal component analysis: Wu *et al*. (2005) have tackled human abnormal behaviour detection by using Principal Component Analysis (PCA) and Support Vector Machine (SVM). Their system applied foreground detection, segmentation, tracking, body part labelling and optical flow analysis in order to detect running people, person carrying a long bar, bending down movement while most are standing or walking and hand waving in crowded environment.

A statistical background model was utilised for the foreground detection technique in order to find the foreground pixels. The optical flow method was then used to detect the foreground blobs by finding big blobs. Next, the blobs that carry multiple peoples were separated into individual component when trapping occurs during segmentation process. For the tracking stage, distance and colour information were combined to track multiple peoples.

They have used PCA and optical flow approach in different cases for abnormal behaviour detection. For the first method, they have implemented PCA for selecting the ideal features since the information size is too big. The range from the centroid to each region are computed and changed into coefficients by means of discrete fourier transformation. Then, normal and abnormal behaviours are distinguished by using SVM classifier.

Next, an optical flow approach is applied as the second method to compute the rate of each pixel. The rate of each pixel is refined by using 3×3 pattern and a predetermined threshold is used to obtain the binary velocity image. The authors have set the grey value to be 0 if the value of filtered pixel is smaller than the threshold and otherwise, it is set to be 255. The noise is removed by grouping fragmented components into blobs. The experimental results showed that the algorithm works well and produce robust abnormal human behaviour detection in crowded environment and successfully differentiate running, carrying a long bar, bending down or waving hand activities for both proposed methods.

Researchers implemented PCA technique to find good characteristics by reassembling information and make the characteristics to be highly noticeable. They have also used SVM classifier because of its

effectiveness and has a better principle compared to other neural networks such as fisher linear discriminant and nearest neighbour algorithm.

CONCLUSION

In conclusion, vision-based aggressive human behaviour detection system has been widely studied. The overview of related works for human behaviour detection system have been explored to enhance the knowledge in this topic, especially in aiding algorithm selection to find the aggressive behaviour for a particular environment. This kind of technology is expected to assist safety personnel to detect anomalous activities that may cause violent actions. Hence, this system is capable to improve situational awareness among the security personnel, especially when they are in charge of monitoring people in prison or detention rooms. An enclosed environment to detect kicking, pushing and boxing behaviour in real-time applications is being used as a benchmark for comparisons in this technological innovation. For future works, more sample of aggressive behaviours can be analysed like bittting, headbutting or hair pulling to make this system more robust and effective either in enclosed or unenclosed room applications.

ACKNOWLEDGEMENT

This research is supported by Ministry of Higher Education Malaysia (FRGS/2/2014/TK03/UKM/02/05) and University Kebangsaan Malaysia (IP-2014-031).

REFERENCES

Appiah, K., A. Hunter, J. Owens, P. Aiken and K. Lewis, 2009. Autonomous real-time surveillance system with distributed IP cameras. Proceedings of the 3rd ACM/IEEE International Conference on Distributed Smart Cameras, August 30-September 2, 2009, Como, Italy, pp: 1-8.

Blasch, E., 2009. NAECON08 grand challenge entry using the belief filter in audio-video track and ID fusion. Proceedings of the IEEE National Aerospace and Electronics Conference, July 21-23, 2009, Dayton, OH., USA., pp: 296-303.

Chen, D., H. Wactlar, M.Y. Chen, C. Gao, A. Bharucha and A. Hauptmann, 2008. Recognition of aggressive human behavior using binary local motion descriptors. Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, August 20-25, 2008, Vancouver, BC., Canada, pp: 5238-5241.

Ezzahout, A. and R.O.H. Thami, 2013. Conception and development of a video surveillance system for detecting, tracking and profile analysis of a person. Proceedings of the 3rd International Symposium on ISKO-Maghreb, November 8-9, 2013, Marrakech, Morocco, pp: 1-5.

Hou, A.L., J.L. Guo, C.J. Wang, L. Wu and F. Li, 2013. Abnormal behavior recognition based on trajectory feature and regional optical flow. Proceedings of the 7th International Conference on Image and Graphics, July 26-28, 2013, Qingdao, China, pp: 643-649.

Ju, J., B. Ku, D. Kim, T. Song, D.K. Han and H. Ko, 2015. Online multi-person tracking for intelligent video surveillance systems. Proceedings of the IEEE International Conference on Consumer Electronics, January 9-12, 2015, Las Vegas, NV., USA., pp: 345-346.

Kim, J.S., D.H. Yeom and Y.H. Joo, 2011. Fast and robust algorithm of tracking multiple moving objects for intelligent video surveillance systems. IEEE Trans. Consum. Electron., 57: 1165-1170.

Lim, H.S., M.Z. MatJafri, K. Abdullah, C.J. Wong and N.M. Saleh, 2006. PM10 retrieval technique from internet protocol camera imageries. Proceedings of the International Conference on Computing and Informatics, June 6-8, 2006, Kuala Lumpur, Malaysia, pp: 1-6.

Moon, H.M., C.H. Seo, Y. Chung and S.B. Pan, 2009. Privacy protection technology in video surveillance system. Proceedings of the 4th International Conference on Embedded and Multimedia Computing, December 10-12, 2009, Jeju, South Korea, pp: 1-6.

Murata, A., T. Kanagawa, N. Hata and T. Hayami, 2012. Effects of penalty and probability of punishment on cooperative behavior in 2-person Prisoner's dilemma situation. Proceedings of the SICE Annual Conference, August 20-23, 2012, Akita, Japan, pp: 2144-2150.

Murthy, O.V.R., I. Radwan, A. Dhall and R. Goecke, 2013. On the effect of human body parts in large scale human behaviour recognition. Proceedings of the International Conference on Digital Image Computing: Techniques and Applications, November 26-28, 2013, Hobart, Australia, pp: 1-8.

Nirjon, S., C. Greenwood, C. Torres, S. Zhou and J.A. Stankovic *et al*., 2014. Kintense: A robust, accurate, real-time and evolving system for detecting aggressive actions from streaming 3D skeleton data. Proceedings of the IEEE International Conference on Pervasive Computing and Communications, March 24-28, 2014, Budapest, Hungary, pp: 2-10.

Ouanane, A. and A. Serir, 2013. New paradigm for recognition of aggressive human behavior based on bag-of-features and skeleton graph. Proceedings of the 8th International Workshop on Systems, Signal Processing and their Applications, May 12-15, 2013, Algiers, Algeria, pp: 133-138.

Pang, J.M., V.V. Yap and C.S. Soh, 2014. Human behavioral analytics system for video surveillance. Proceedings of the IEEE International Conference on Control System, Computing and Engineering, November 28-30, 2014, Batu Ferringhi, Malaysia, pp: 23-28.

Stratton, G., R. Murphy, M. Rosenberg, P. Fergus and A. Attwood, 2012. Creating intelligent environments to monitor and manipulate physical activity and sedentary behavior in public health and clinical settings. Proceedings of the IEEE International Conference on Communications, June 10-15, 2012, Ottawa, ON., Canada, pp: 6111-6115.

Theodoridis, T. and H. Hu, 2013. Modeling aggressive behaviors with evolutionary taxonomers. IEEE Trans. Hum.-Mach. Syst., 43: 302-313.

Wilson, B.D., 2008. Development in video technology for coaching. Sports Technol., 1: 34-40.

Wu, X., Y. Ou, H. Qian and Y. Xu, 2005. Detection system for human abnormal behavior. Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, August 2-6, 2005, Edmonton,Canada, pp: 1204-1208.

Zhen, C. and Y. Su, 2010. Research on the image capture technology in the video screen. Proceedings of the 3rd International Conference on Advanced Computer Theory and Engineering, Volume 5, August 20-22, 2010, Chengdu, China, pp: 581-584.

Zulkifley, M.A., N.S. Samanu, N.A.A.N. Zulkepeli, Z. Kadim and H.H. Woon, 2016. Kalman filter-based aggressive behaviour detection for indoor environment. Proceedings of the 7th International Conference on Information Science and Applications, February 15-18, 2016, Ho Chi Minh City, Vietnam, pp: 829-837.