Fear Detection with Background Subtraction from RGB-D data.

Anne Veenendaal, Elliot Daly, Eddie Jones, Zhao Gang

***Abstract*— This paper examines the automated detection of human activity inside a hallway using depth data for background subtraction and hall way walls as reference frame. The study focusses on panic and fear as the main activity for recognition. Many times, humans try to evade, run to a safe place because of fear caused by violence, presence of active shooter or because of natural calamity such as earthquake and tornado. The research first uses a series of images to extract the human blob using a background to foreground subtraction approach. The blob is then normalized using a reference frame of hallway, followed by edge detection and then features are extracted by continuous application of a mesh. Further, a support vector machine (SVM) classifier was used to detect activity representing fear. The results indicate recognition accuracy of 74.8% for real time unconstrained video image frames.**

***Keywords*— Fear, Evade, Human Activity, Threat, Surveillance, Sensor, SVM, 3D tracking, Emotion, Hand, Body, Face, Legs.**

I. Introduction

More and more incidents of violence in public are being reported. Sometimes there are natural calamities such as tornado, earthquake, flash floods, tsunamis. In such life threatening scenarios humans take evasive actions. They panic and run to take shelter and express fear. Surveillance systems equipped with detection of such behaviour can provide alerts and notify the authorities to take immediate action to resolve the issue. This paper examines detection of human activities in the hallway specifically during the event of some threat. Haritaoglu et. al [1] have implemented a surveillance system to detect human activity in outdoor environments. The study used infrared sensing and grey scale image processing for object detection. In a research by Stauffer [2], similarities in activities was examined to find patterns in real time. Elgammal et. al [3] used kernel density estimation to separate background from foreground. Researchers [4], [5], [6], [7], [8] have analyzed view independent human gait detection, using view calibration, multi-view recording, shape and motion features. The studies [9], [10], [11], [12], [13], [14], [15] on human activity recognition have used various techniques such as probabilistic classification, HMM and model based approaches. Readers are directed towards studies [16] through [36] for surveys, unimodal, bimodal and multimodal emotion recognition techniques as well as software strategies for real time implementation of such systems. Researchers [37] through [53] have focused on capturing data using sensors, motion based human activity recognition, 3D feature creation and using color and depth data (RGB-D) for segmentation and action recognition.

II. Method

In this research 28 participants were asked to enact actions (7 action categories in total) representing fear and normal walk through the hallway. The video footage was captured using an infrared sensor and video camera to store the color and depth data. The sessions were repeated under different lighting conditions, clothes color, wall color. The data was then fed to the automatic background subtraction system (AUBSS). The background subtraction used two channels for the data. The first channel was the depth frame and the second channel was the color frame. In the first step the total available depth information was calculated. Then the depth was split into 1000 sub divisions. Then the frame which was farthest was processed first.

Each pixel in the frame was analysed for similarity in color in the HSV space with the threshold of 20. This resulted in blobs of pixel with similar colors. This was repeated for each sub division of the depth frame. Once a collection of blobs was obtained the images were superimposed with a reference frame. The reference frame was constructed by first performing edge detection to detect the walls in the hallway. Then the reference frame was used to eliminate the blobs which were outside the frame boundaries. A grid image was applied on each blob and the 3-D co-ordinates and color of the intersecting point were taken as the feature for the particular frame. A feature vector was constructed from all the depth sub divisions. This was used for training the SVM based classifier for each class of action. The slack variable was set to 0.3 and optimized using grid forward search. The data was split into 80% training data and 20% test data. The training was done using 10-fold cross validation. The radial basis function was used as the kernel for the training.

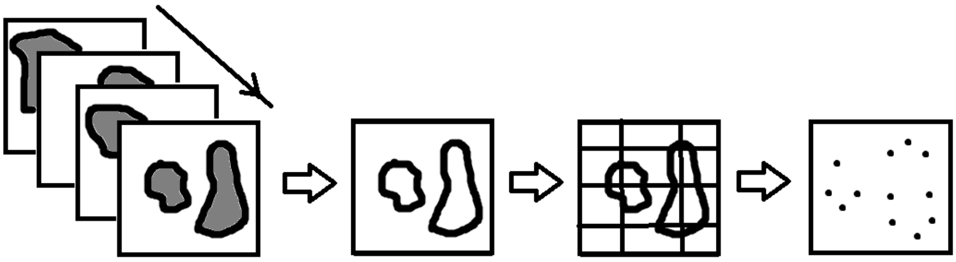

Fig. 1. Blob extraction and grid based feature co-ordinate extraction.

Fig. 2. Edge detection on the background subtracted human shapes.

III. Results

Results for classification of actions in controlled lighting.

0.789	0.025	0	0.063	0	0.088	0.034	Run Forward
0.128	**0.79**	0.004	0.034	0.033	0	0.011	Duck
0.15	0.028	**0.771**	0	0.052	0	0	Crawl
0.016	0.05	0.086	**0.651**	0.054	0.054	0.09	Run Away
0.027	0.038	0.03	0.021	**0.879**	0	0.005	Push against doors
0.068	0.019	0	0.025	0	**0.866**	0.023	Push against wall
0.022	0.051	0.009	0.013	0.022	0.029	**0.853**	Normal walk

The overall accuracy of the activity recognition system under controlled lighting was 80%.

Results for classification of actions in dim lighting.

0.698	0.022	0.02	0.066	0.068	0.078	0.047	Run Forward
0.126	**0.775**	0.004	0.034	0.032	0	0.03	Duck
0.139	0.025	**0.713**	0.032	0.048	0.029	0.014	Crawl
0.031	0.052	0.09	**0.679**	0.056	0.056	0.035	Run Away
0.028	0.039	0.032	0.086	**0.74**	0.032	0.043	Push against doors
0.066	0.018	0.057	0.024	0	**0.795**	0.04	Push against wall
0.022	0.05	0.009	0.031	0.022	0.029	**0.838**	Normal walk

The overall accuracy of the activity recognition system was 74.8% which was lower than the accuracy in controlled lighting.

IV. Conclusions

The drop in the accuracy in activity recognition under dim lighting showed a limitation in the system and more work needs to be done to make the system more generalizable and robust. The research provided benchmark fear detection results data for further analysis. The processing performance of the implementation was real time which indicated that the method proposed was not computationally intensive as initial thought. As future scope some improvements in the mesh based feature extraction techniques need to be explored to improve the classification accuracy and discriminatory power of the feature vectors.

References

[1] I. Haritaoglu, D. Harwood and L. S. Davis, "W4: Real-time surveillance of people and their activities", IEEE Trans. Pattern Anal. Machine Intell, vol. 22, pp. 809-830, 2000.

[2] 2. C. Stauffer and W. E. L. Grimson, "Learning patterns of activity using real-time tracking", IEEE Trans. Pattern Anal. Machine Intell, vol. 22, pp. 747-757, 2000

[3] A. Elgammal, R. Duraiswami, D. Harwood and L. S. Davis, "Background and foreground modelling using nonparametric kernel density estimation for visual surveillance", Proc. IEEE, vol. 90, no. 7, pp. 1151-1163Huang, P., Harris, C., Nixon, M., 1999. Human gait recognition in canonical space using spatio-temporal templates. IEE Proc. Vision, Image Signal Process., 93–100.

[4] Liang Wang, Tieniu Tan, Weiming Hu, Huazhong Ning, "Automatic Gait Recognition Based on Statistical Shape Analysis", IEEE Transactions on Image Processing, Vol 12, No 9, September, 2003.

[5] W. Zeng, C. Wang "Human gait recognition via deterministic learning.", Neural Network, vol 35, pp 92-102, November, 2012.

[6] Lin, Kuo-Wei and Wang, Shu-Ting and Chung, Pau-Choo and Yang, Ching-Fang, "A New View-Calibrated Approach for Abnormal Gait Detection", Advances in Intelligent Systems and Applications - Volume 2, Proceedings of the International Computer Symposium ICS 2012 Held at Hualien, Taiwan, December 12--14, 2012.

[7] A Puad Ismail, "Abnormal gait detection using Hexagonal method on Model based Front view model.", Journal of Electrical Systems, 2015. S. M. Metev and V. P. Veiko, Laser Assisted Microtechnology, 2nd ed., R. M. Osgood, Jr., Ed. Berlin, Germany: Springer-Verlag, 1998.

[8] F. Niu and M. Abdel-Mottaleb, “View-Invariant Human Activity Recognition Based on Shape and Motion Features,” in Proceedings of the IEEE Sixth International Symposium on Multimedia Software Engineering, pp. 546-556, 2004.

[9] M. Z. Uddin, J. J. Lee, and T.-S. Kim, “Independent Shape Component-Based Human Activity Recognition via Hidden Markov Model,” Applied Intelligence, DOI:10.1007/s10489008-0159-2, 2009.

[10] M. Z. Uddin, J. J. Lee and T.-S. Kim, “Human Activity Recognition Using Independent Component Features from Depth Images,” in Proceedings of the 5th International Conference on Ubiquitous Healthcare, pp. 181-183, 2008.

[11] N. D. Thang, Y.-K Lee, S.-Y. Lee, and T.-S. Kim, “Estimation of 3-D Human Body Posture via Co-Registration of 3-D Human Model and Sequential Stereo Information,” Applied Intelligence, DOI: 10.1007/s10489-009-0209-4, 2010.

[12] R. Cucchiara, C. Grana, A. Prati, and R. Vezzani, “Probabilistic posture classification for human-behavior analysis,” IEEE Trans. Syst. Man, and Cybern. A, vol. 35, no. 1, pp. 42–54, 2005.

[13] R. Hamid, Y. Huang, and I. Essa, “ARGMode–Activity recognition using graphical models”, in Proc. Conf. Comput. Vision Pattern Recog., vol. 4, pp. 38–45, Madison, Wisconsin, 2003.

[14] P. S. Huang, C. J. Harris, and M. S. Nixon, “Canonical space representation for recognizing humans by gait or face,” in Proc. IEEE Southwest Symp. Image Anal. Interpretation, pp. 180–185, 1998.

[15] J. Yamato, J. Ohya, and K. Ishii, “Recognizing human action in time-sequential images using hidden Markov model,” in Proc. IEEE CVPR, pp. 379–385, 1992.

[16] L. X. Wang and J. M. Mendel, “Generating fuzzy rules by learning from examples,” IEEE Trans. Syst., Man Cybern., vol. 22, no. 6, pp. 1414–1427, 1992.

[17] K. Etemad and R. Chellappa, “Discriminant analysis for recognition of human face images,” J. Opt. Soc. Am. A, Vol. 14, pp. 1724–1733, 1997.

[18] A. S. Patwardhan, 2016. “Structured Unit Testable Templated Code for Efficient Code Review Process”, PeerJ Computer Science (in review), 2016.

[19] A. S. Patwardhan, and R. S. Patwardhan, “XML Entity Architecture for Efficient Software Integration”, International Journal for Research in Applied Science and Engineering Technology (IJRASET), vol. 4, no. 6, June 2016.

[20] A. S. Patwardhan and G. M. Knapp, “Affect Intensity Estimation Using Multiple Modalities,” Florida Artificial Intelligence Research Society Conference, May. 2014.

[21] A. S. Patwardhan, R. S. Patwardhan, and S. S. Vartak, “Self-Contained Cross-Cutting Pipeline Software Architecture,” International Research Journal of Engineering and Technology (IRJET), vol. 3, no. 5, May. 2016.

[22] A. S. Patwardhan, “An Architecture for Adaptive Real Time Communication with Embedded Devices,” LSU, 2006.

[23] A. S. Patwardhan and G. M. Knapp, “Multimodal Affect Analysis for Product Feedback Assessment,” IIE Annual Conference. Proceedings. Institute of Industrial Engineers-Publisher, 2013.

[24] A. S. Patwardhan and G. M. Knapp, “Aggressive Action and Anger Detection from Multiple Modalities using Kinect”, submitted to ACM Transactions on Intelligent Systems and Technology (ACM TIST) (in review).

[25] A. S. Patwardhan and G. M. Knapp, “EmoFit: Affect Monitoring System for Sedentary Jobs,” preprint, arXiv.org, 2016.

[26] A. S. Patwardhan, J. Kidd, T. Urena and A. Rajagopalan, “Embracing Agile methodology during DevOps Developer Internship Program”, IEEE Software (in review), 2016.

[27] A. S. Patwardhan, “Analysis of Software Delivery Process Shortcomings and Architectural Pitfalls”, PeerJ Computer Science (in review), 2016.

[28] A. S. Patwardhan, “Multimodal Affect Recognition using Kinect”, ACM TIST (in review), 2016.

[29] A. S. Patwardhan, “Augmenting Supervised Emotion Recognition with Rule-Based Decision Model”, IEEE TAC (in review), 2016.

[30] A. S. Patwardhan, Jacob Badeaux, Siavash, G. M. Knapp, “Automated Prediction of Temporal Relations”, Technical Report. 2014.

[31] S. D’Mello and A. Graesser, “Multimodal Semi-Automated Affect Detection from Conversational Cues, Gross Body Language, and Facial Features,” User Modeling and User-Adapted Interaction, vol. 10, pp. 147-187, 2010.

[32] T. Baenziger, D. Grandjean, and K.R. Scherer, “Emotion Recognition from Expressions in Face, Voice, and Body. The Multimodal Emotion Recognition Test (MERT),” Emotion, vol. 9, pp. 691-704, 2009.

[33] C. Busso et al., “Analysis of Emotion Recognition Using Facial Expressions, Speech and Multimodal Information,” Proc. Int’l Conf. Multimodal Interfaces, T.D.R. Sharma, M.P. Harper,G. Lazzari, and M. Turk, eds., pp. 205-211, 2004.

[34] N. Sebe, I. Cohen, and T.S. Huang, “Multimodal Emotion Recognition,” Handbook of Pattern Recognition and Computer Vision, World Scientific, 2005.

[35] R. Cowie, E. Douglas-Cowie, N. Tsapatsoulis, G. Votsis, S. Kollias, W. Fellenz, and J. Taylor, “Emotion Recognition in Human-Computer Interaction,” IEEE Signal Processing Magazine, vol. 18, no. 1, pp. 32-80, 2001.

[36] S. Carlsson and J. Sullivan, “Action recognition by shape matching to key frames,”in Proc. IEEE Comput. Soc. Workshop Models versus Exemplars in Comput. Vision, pp. 263–270, Miami, Florida, 2002.

[37] A. Kapoor and R.W. Picard, “Multimodal Affect Recognition in Learning Environments,” Proc. 13th Ann. ACM Int’l Conf. Multimedia, pp. 677-682, 2005.

[38] O. Oreifej and Z. Liu. HON4D: Histogram of oriented 4D normals for activity recognition from depth sequences. In IEEE Conference on Computer Vision and Pattern Recognition, pp 716–723, 2013

[39] L. Xia, C.-C. Chen, and J. K. Aggarwal. View invariant human action recognition using histograms of 3D joints. 2012 IEEE Computer Vision and Pattern Recognition Workshops, pp 20–27.

[40] J. Wang, Z. Liu, Y. Wu, and J. Yuan. Mining actionlet ensemble for action recognition with depth cameras. In CVPR, pp 1290-97, 2012.

[41] Sijin Li, Weichen Zhang, and Antoni B. Chan. Maximum-Margin Structured Learning with Deep Networks for 3D Human Pose Estimation. Intl. Conf. on Computer Vision (ICCV), Santiago, 2015.

[42] S. Ji, W. Xu, M. Yang, and K. Yu. 3d convolutional neural networks for human action recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(1):221-231, 2013.

[43] P. Luo, X. Wang, and X. Tang. Pedestrian parsing via deep decompositional neural network. In ICCV, pages 2648-55, 2013.

[44] J. K. Aggarwal and M. S. Ryoo. Human activity analysis: A review. ACM Computing Survey, 43(3):16, 2011.

[45] U. Maurer, A. Smailagic, D. P. Siewiorek, and M. Deisher. Activity recognition and monitoring using multiple sensors on different body positions. In Intl Workshop on Wearable and Implantable Body Sensor Networks, pp 113–116, 2006.

[46] L. Chen, H. Wei, J. M. Ferryman. A survey of human motion analysis using depth imagery. Pattern Recognition Letters, 34:1995, 2013.

[47] W. Li, Z. Zhang, and Z. Liu. Action recognition based on a bag of 3D points. IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp 9-14, 2010.

[48] J. Shotton, T. Sharp, A. Kipman, A. W. Fitzgibbon, M. Finocchio, A. Blake, M. Cook, and R. Moore. Real-time human pose recognition in parts from single depth images. Communications of the ACM, 56(1):116–124, 2013.

[49] J. Sung, C. Ponce, B. Selman, and A. Saxena. Unstructured human activity detection from RGBD images. In IEEE Intl Conference on Robotics and Automation, pp 842–849, 2012.

[50] C. Zhang, Y. Tian. RGB-D based daily living activity recognition. Journal of Computer Vision and Image Processing, 2(4), Dec. 2012.

[51] D. Gurkaynak and H. Yalcin, Recognition and Classification of Human Activity from RGB-D Videos. IEEE Conference on Signal Processing and Communication Applications (SIU 2015), pp 1642-1646, 2015.
[52] R. Poppe, "Vision-based human motion analysis: an overview," Comput. Vision and Image Understan., vol. 108, pp. 4–18, 2007.
[53] A. F. Bobick and J. W. Davis, "The recognition of human movement using temporal templates," IEEE Trans. Pattern Anal. Machine Intell., vol. 23, no. 3, 2001.
[54] Anne Veenendaal, Elliot Daly, Eddie Jones, Zhao Gang, Sumalini Vartak, Rahul S Patwardhan, Drunken Abnormal Human Gait Detection using Sensors, Computer Science and Emerging Research Journal, vol 1, 2013.
[55] Anne Veenendaal, Elliot Daly, Eddie Jones, Zhao Gang, Sumalini Vartak, Rahul S Patwardhan, Fear Detection with Background Subtraction from RGB-D data, Computer Science and Emerging Research Journal, vol 1, 2013.
[56] Anne Veenendaal, Elliot Daly, Eddie Jones, Zhao Gang, Sumalini Vartak, Rahul S Patwardhan, Code Definition Analysis for Call Graph Generation, Computer Science and Emerging Research Journal, vol 1, 2013.
[57] Anne Veenendaal, Elliot Daly, Eddie Jones, Zhao Gang, Sumalini Vartak, Rahul S Patwardhan, Multi-View Point Drowsiness and Fatigue Detection, Computer Science and Emerging Research Journal, vol 2, 2014.
[58] Anne Veenendaal, Elliot Daly, Eddie Jones, Zhao Gang, Sumalini Vartak, Rahul S Patwardhan, Group Emotion Detection using Edge Detecttion Mesh Analysis, Computer Science and Emerging Research Journal, vol 2, 2014.
[59] Anne Veenendaal, Elliot Daly, Eddie Jones, Zhao Gang, Sumalini Vartak, Rahul S Patwardhan, Polarity Analysis of Restaurant Review Comment Board, Computer Science and Emerging Research Journal, vol 2, 2014.
[60] Anne Veenendaal, Elliot Daly, Eddie Jones, Zhao Gang, Sumalini Vartak, Rahul S Patwardhan, Sentiment Analysis in Code Review Comments, Computer Science and Emerging Research Journal, vol 3, 2015.
[61] Anne Veenendaal, Elliot Daly, Eddie Jones, Zhao Gang, Sumalini Vartak, Rahul S Patwardhan, Temporal Analysis of News Feed Using Phrase Position, Computer Science and Emerging Research Journal, vol 3, 2015.
[62] Anne Veenendaal, Elliot Daly, Eddie Jones, Zhao Gang, Sumalini Vartak, Rahul S Patwardhan, Decision Rule Driven Human Activity Recognition, Computer Science and Emerging Research Journal, vol 3, 2015.
[63] Anne Veenendaal, Elliot Daly, Eddie Jones, Zhao Gang, Sumalini Vartak, Rahul S Patwardhan, Depression and Sadness Recognition in Closed Spaces, Computer Science and Emerging Research Journal, vol 4, 2016.
[64] Anne Veenendaal, Elliot Daly, Eddie Jones, Zhao Gang, Sumalini Vartak, Rahul S Patwardhan, Dynamic Probabilistic Network Based Human Action Recognition, Computer Science and Emerging Research Journal, vol 4, 2016.
[65] Anne Veenendaal, Elliot Daly, Eddie Jones, Zhao Gang, Sumalini Vartak, Rahul S Patwardhan, Fight and Aggression Recognition using Depth and Motion Data, Computer Science and Emerging Research Journal, vol 4, 2016.
[66] Anne Veenendaal, Elliot Daly, Eddie Jones, Zhao Gang, Sumalini Vartak, Rahul S Patwardhan, Sensor Tracked Points and HMM Based Classifier for Human Action Recognition, Computer Science and Emerging Research Journal, vol 5, 2016.
[67] A. S. Patwardhan, "Edge Based Grid Super-Imposition for Crowd Emotion Recognition", International Research Journal of Engineering and Technology (IRJET), May. 2010.
[68] A. S. Patwardhan, "Human Activity Recognition Using Temporal Frame Decision Rule Extraction", International Research Journal of Engineering and Technology (IRJET), May. 2010.
[69] A. S. Patwardhan, "Low Morale, Depressed and Sad State Recognition in Confined Spaces", International Research Journal of Engineering and Technology (IRJET), May. 2011.
[70] A. S. Patwardhan, "View Independent Drowsy Behavior and Tiredness Detection", International Research Journal of Engineering and Technology (IRJET), May. 2011.
[71] A. S. Patwardhan, "Sensor Based Human Gait Recognition for Drunk State", International Research Journal of Engineering and Technology (IRJET), May. 2012.
[72] A. S. Patwardhan, "Background Removal Using RGB-D data for Fright Recognition", International Research Journal of Engineering and Technology (IRJET), May. 2012.
[73] A. S. Patwardhan, "Depth and Movement Data Analysis for Fight Detection", International Research Journal of Engineering and Technology (IRJET), May. 2013.
[74] A. S. Patwardhan, "Human Action Recognition Classification using HMM and Movement Tracking", International Research Journal of Engineering and Technology (IRJET), May. 2013.
[75] A. S. Patwardhan, "Feedback and Emotion Polarity Extraction from Online Reviewer sites", International Research Journal of Engineering and Technology (IRJET), May. 2014.
[76] A. S. Patwardhan, "Call Tree Detection Using Source Code Syntax Analysis", International Research Journal of Engineering and Technology (IRJET), May. 2014.
[77] A. S. Patwardhan, "Walking, Lifting, Standing Activity Recognition using Probabilistic Networks", International Research Journal of Engineering and Technology (IRJET), May. 2015.
[78] A. S. Patwardhan, "Online News Article Temporal Phrase Extraction for Causal Linking", International Research Journal of Engineering and Technology (IRJET), May. 2015.
[79] A. S. Patwardhan, "Online Comment Processing for Sentiment Extraction", International Research Journal of Engineering and Technology (IRJET), May. 2016.