

Group Emotion Detection using Edge Detection Mesh Analysis

Anne Veenendaal, Elliot Daly, Eddie Jones, Zhao Gang, Sumalini Vartak, Rahul S Patwardhan

Abstract— Several emotion recognition system studies have focused mostly on videos and images containing individual actors. This study examines the detection of emotions in a group and crowd settings. Edge detection was used with a Mesh superimposition to extract the features. The feature movement in terms of movement from the reference point was used to track across sequences of images from the color channel. Additionally the video data capturing was done on spontaneous emotions invoked by watching sports events from group of participants. The overall accuracy of the group emotion classifier was 70.9%.

Keywords— Group emotion, crowd emotion, SVM, edge detection, spontaneous emotion recognition.

I. INTRODUCTION

Emotion recognition by computers has several applications in marketing and user profile specific advertisements. Many people watch live sports telecasts in groups. The football games in United States and World cup soccer games or basketball games are viewed by groups of people assembling in a public place, sport bars and a friend's house. The various events during the games invoke several emotions among the supporters of each team. The emotions range from disgust, disappointment or joy depending on whether the supporter's team won or lost. There are incidents of controversial calls, surprise decisions and anguish when a referee makes a bad call or a favorite player is injured and makes an excellent shot, which is revoked. Multimodal emotion recognition has been studied using RGB-D data and audio-visual data. These studies have examined the videos mostly in controlled conditions using individual enactments or spontaneous emotional episodes. In this paper, we focus on spontaneous group emotion detection under indoor lighting. Researchers [1] have examined the side effects of emotional thinking on memory and judgement. Emotion representation has been discussed in the field of psychology [2] using basic emotions such as happy, sad, angry, fear, surprise and disgust. In addition to video-based emotion recognition, researchers [3] have also used vocal features to examine expression of emotions. A survey [4] on emotion detection and modelling using speech, audio and vocal input data has evaluated significance of such modalities in recognition accuracy. Application of emotion estimation in educational settings has been studied by researchers [5]. Studies [6] have used bio-sensors to assist in emotion detection and analyze the psychological and physiological effects of emotions in humans. Researchers [7] have used face, voice and body data to evaluate emotion modelling. A study [8] investigated the co-relation between various input channels in estimation accuracy using neural networks. Researchers [9] used dynamic Bayesian networks for monitoring crowd activity. Researchers [10, [11], [12] have studied intelligent surveillance systems using bio-sensors and bio-inspired devices. Studies [13] have evaluated the connection between body expressions and emotional states. Some studies [14], [15] have focussed on emotion maps and cognitive states and their relation with emotion modelling and expression. Studies [16], [17], [18], [19] have investigated dimensionality issues in emotion modelling, effect of cross-cultural influence of emotion, pedestrian behaviour and crowd simulation for emotion estimation. Research [20], [21] has been done on human gait, pedestrian dynamics especially under influence of alcohol. Studies [22], [23] have focussed on specific emotion such as detection of fear, implementation strategies for automatic systems. Researchers have studied [24], [25] the view-invariant emotion detection for specific human behaviour such as crowd analysis, fatigue and tiredness and sleepiness prediction. In addition to audio-visual modalities studies [26], [27], [28] have focussed on text based sentiment analysis. Several supervised learning techniques (neural network, DBN, HMM) [29], [30], [31], [32], [33] for emotion recognition in various settings (closed spaces, indoor and outdoor) have been used to determine the accuracy of each method. Studies [34], [35], [36], [37], [38] on software implementation of automated continuous computer vision based multimodal emotion recognition techniques have been studied in detail. These studies have implemented novel methods to solve problem of modality fusion using hybrid methods and also provided architectural methods for real time detection. Researchers [39], [40], [41], [42], [43], [44], [45], [46], [47] have developed novel algorithms and behavioural rule based features for multimodal emotion recognition using continuous data and supervised learning.

II. METHOD

4 groups of 5 people for a total of 20 individuals participated in the study. The individuals were all dressed in casual attire and the were aged in the range of 22 to 45. 18 participants were male and 2 were females. The groups gathered for different games of basketball finals. Three different cameras were used to record their reactions for 5 min at the beginning of each half and 5 min before the end of each half. The games usually get interesting at the beginning and towards the end. Tight games invoke more emotions compared to one-sided matches. The non-intrusive way of data capture allowed the subjects to watch the event and express their reactions without consciousness about being recorded on camera.

After the data was captured, the canny edge detection was applied on every third frame from the video sequence. This caused down sampling of a 24 frame per second to an 8 frames per second sequence. For each image frame the edge detection filter was applied. Then the frame was divided into 20 x 20 meshes and the intersection of the grid lines with the edges was considered as feature co-ordinates. For a consistent feature vector length, each line was further divided into 5 divisions and the features were counted as 1 for an intersection and 0 for no intersection. Thus 20 x 5 features for vertical lines and 20 x 5 features for horizontal lines on the mesh, for a total of 200 feature co-ordinates were obtained to form the feature vector. Additionally the temporal features were also tracked.

The movement of each co-ordinate was tracked across the eight frames. Thus the final feature vector consisted of 200 static and 200 velocity values across 8 frames for a total of 400 features. The best first search technique was used for feature selection.

This resulted in reduced dimensionality of feature vector size and 23 discriminating features were chosen. The sequence of all actions was annotated using three observers to avoid inter-annotator disagreement. The classes used were happy, angry, surprised, sad, disgust, fear and neutral. After the annotation was done, the feature vector was used to train the classifiers using support vector machine and radial basis function as the kernel function with 0.4 as the slack variable. The optimized slack variable was calculated using grid forward search method. The data was split into 70% training and 30% test data. The training was performed using 10-fold cross validation. The optimal threshold for edge detection was obtained using classification tests on sample data set with grid size set to 10 through 50. Similarly the optimal grid size was calculated using classification test and using six basic emotions and the neutral state as the candidate test emotions with the threshold set to 0.2 through 0.8.

III. RESULTS

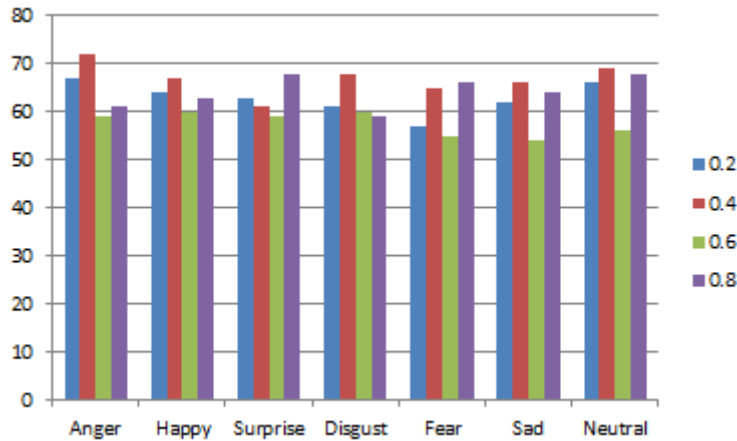


Fig. 4 Thresholds for edge detection

The threshold of 0.2 resulted in better accuracy for anger and neutral emotion. The threshold of 0.4 resulted in best accuracy for most of the emotions except fear and surprise.

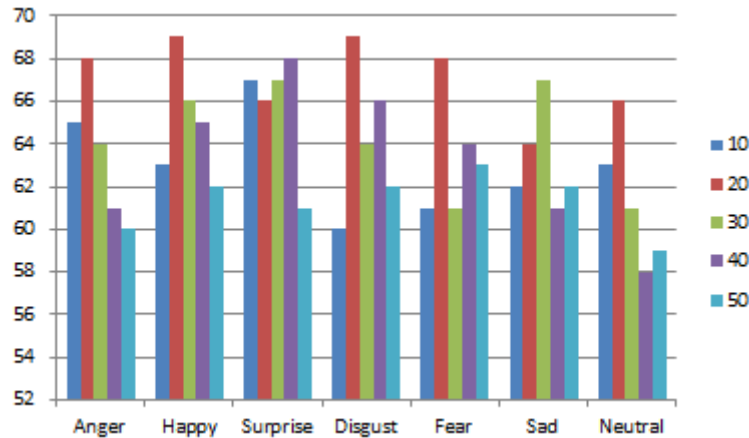


Fig. 5 Thresholds for grid size

0	1	2	3	4	5	6		
0.678	0.036	0.046	0.039	0.081	0.072	0.048	0	Anger
0.099	0.766	0.025	0.05	0.022	0.005	0.032	1	Happy
0.109	0.023	0.72	0.037	0.062	0.031	0.018	2	Surprise
0.024	0.03	0.113	0.704	0.065	0.025	0.039	3	Sad
0.034	0.039	0.119	0.074	0.672	0.022	0.04	4	Fear
0.058	0.019	0.058	0.022	0.041	0.729	0.075	5	Disgust
0.085	0.04	0.068	0.027	0.065	0.025	0.691	6	Neutral

Fig. 6 Classification accuracy for six emotions

IV. CONCLUSIONS

The threshold for edge detection that yielded the best recognition results was 0.4. The classification results for happiness was the highest at 76.6% followed by disgust 72.9% and surprise 72%. The recognition rate for fear was the lowest with 67.2%. The overall accuracy of the group emotion recognition process was 70.9%. The grid size of 20 resulted in the best accuracy for 5 out of 7 emotion classes (including neutral class). This study evaluated a novel technique that implemented image processing steps to extract the edges and then extract the features for chaotic scenes resulting from expression of emotions in group spectator settings. The group of people in the scene resulted in occluded view where many other techniques are not accurate because of the lack of patterns. As a result the action based recognition techniques cannot be applied in such scenarios. The techniques mentioned in the paper showed promising results to overcome this limitation of view-dependence and lack of sufficient training data.

REFERENCES

- [1] Bower, G.H., Cohen, P.R.: Emotional influences in memory and thinking: Data and theory. In: *Affect and Cognition*, vol. 1 (1982)
- [2] Ekman, P.: Expression and the nature of emotion. *Approaches to Emotion* 3, 19–344 (1984)
- [3] Juslin, P.N., Scherer, K.R.: Vocal expression of affect. *The new handbook of Methods in Nonverbal Behavior Research*, 65–135 (2005)
- [4] Scherer, K.R.: Vocal communication of emotion: A review of research paradigms. *Speech Communication* 40, 227–256 (2003)
- [5] Mota, S., Picard, R.W.: Automated posture analysis for detecting learner’s interest level. In: *2003 Conference on Computer Vision and Pattern Recognition Workshop, CVPRW 2003*, vol. 5. IEEE (2003)
- [6] Villon, O., Lisetti, C.: A user-modeling approach to build user’s psycho-physiological maps of emotions using bio-sensors. In: *The 15th IEEE International Symposium on Robot and Human Interactive Communication, ROMAN 2006* (2006)
- [7] Bänziger, T., Grandjean, D., Scherer, K.R.: Emotion recognition from expressions in face, voice, and body: The Multimodal Emotion Recognition Test (MERT). *Emotion* 9(5), 691 (2009)
- [8] Jockusch, J., Ritter, H.: An instantaneous topological mapping model for correlated stimuli. In: *International Joint Conference on Neural Networks, IJCNN 1999*, vol. 1. IEEE (1999)
- [9] Chiappino, S., et al.: Event based switched dynamic bayesian networks for autonomous cognitive crowd monitoring. In: *Wide Area Surveillance*, pp. 93–122. Springer, Heidelberg (2014)
- [10] Dore, A., et al.: A bio-inspired system model for interactive surveillance applications. *Journal of Ambient Intelligence and Smart Environments* 3(2), 147–163 (2011)
- [11] Dore, A., Regazzoni, C.S.: Interaction Analysis with a Bayesian Trajectory Model. *IEEE Intelligent Systems* 25(3), 32–40 (2010)
- [12] Dore, A., Cattoni, A.F., Regazzoni, C.S.: Interaction modeling and prediction in smart spaces: a bio-inspired approach based on autobiographical memory. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans* 40(6), 1191–1205 (2010)
- [13] Damasio, A.R.: *The feeling of what happens: Body and emotion in the making of consciousness*. EBook (1999)
- [14] Kohonen, T.: *Self-organizing maps*, vol. 30. Springer (2001)
- [15] Ortony, A.: *The cognitive structure of emotions*. Cambridge University Press (1990)
- [16] Russell, J.A., Lewicka, M., Niit, T.: A cross-cultural study of a circumplex model of affect. *Journal of personality and social psychology* 57(5), 848 (1989)
- [17] Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal statistical Society* 39(1), 1–38 (1977)
- [18] Schwarz, G.: Estimating the dimension of a model. *The annals of statistics* 6(2), 461–464 (1978)
- [19] Baig, M.W., Barakova, E., Regazzoni, C.S., Rauterberg, M.: Realistic Modeling of Agents in Crowd Simulations. In: *2014 5th International Conference on Intelligent Systems Modelling & Simulation (ISMS)*, pp. 507–512. IEEE (2014)
- [20] Helbing, D., Molnar, P.: Social force model for pedestrian dynamics. *Physical review E* 51(5), 4282 (1995)
- [21] 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.
- [22] 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.
- [23] 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.
- [24] 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.
- [25] Anne Veenendaal, Elliot Daly, Eddie Jones, Zhao Gang, Sumalini Vartak, Rahul S Patwardhan, *Group Emotion Detection using Edge Detection Mesh Analysis*, *Computer Science and Emerging Research Journal*, vol 2, 2014.
- [26] 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.
- [27] 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.
- [28] 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.
- [29] 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.
- [30] 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.
- [31] 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.
- [32] 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.

- [33] 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.
- [34] A. S. Patwardhan, 2016. "Structured Unit Testable Templated Code for Efficient Code Review Process", PeerJ Computer Science (in review), 2016.
- [35] 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.
- [36] A. S. Patwardhan and G. M. Knapp, "Affect Intensity Estimation Using Multiple Modalities," Florida Artificial Intelligence Research Society Conference, May. 2014.
- [37] 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.
- [38] A. S. Patwardhan, "An Architecture for Adaptive Real Time Communication with Embedded Devices," LSU, 2006.
- [39] A. S. Patwardhan and G. M. Knapp, "Multimodal Affect Analysis for Product Feedback Assessment," IIE Annual Conference. Proceedings. Institute of Industrial Engineers-Publisher, 2013.
- [40] 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).
- [41] A. S. Patwardhan and G. M. Knapp, "EmoFit: Affect Monitoring System for Sedentary Jobs," preprint, arXiv.org, 2016.
- [42] A. S. Patwardhan, J. Kidd, T. Urena and A. Rajagopalan, "Embracing Agile methodology during DevOps Developer Internship Program", IEEE Software (in review), 2016.
- [43] A. S. Patwardhan, "Edge Based Grid Super-Imposition for Crowd Emotion Recognition", International Research Journal of Engineering and Technology (IRJET), May. 2010.
- [44] A. S. Patwardhan, "Analysis of Software Delivery Process Shortcomings and Architectural Pitfalls", PeerJ Computer Science (in review), 2016.
- [45] A. S. Patwardhan, "Multimodal Affect Recognition using Kinect", ACM TIST (in review), 2016.
- [46] A. S. Patwardhan, "Augmenting Supervised Emotion Recognition with Rule-Based Decision Model", IEEE TAC (in review), 2016.
- [47] A. S. Patwardhan, Jacob Badeaux, Siavash, G. M. Knapp, "Automated Prediction of Temporal Relations", Technical Report. 2014.
- [48] A. S. Patwardhan, "Edge Based Grid Super-Imposition for Crowd Emotion Recognition", International Research Journal of Engineering and Technology (IRJET), May. 2010.
- [49] A. S. Patwardhan, "Human Activity Recognition Using Temporal Frame Decision Rule Extraction", International Research Journal of Engineering and Technology (IRJET), May. 2010.
- [50] A. S. Patwardhan, "Low Morale, Depressed and Sad State Recognition in Confined Spaces", International Research Journal of Engineering and Technology (IRJET), May. 2011.
- [51] A. S. Patwardhan, "View Independent Drowsy Behavior and Tiredness Detection", International Research Journal of Engineering and Technology (IRJET), May. 2011.
- [52] A. S. Patwardhan, "Sensor Based Human Gait Recognition for Drunk State", International Research Journal of Engineering and Technology (IRJET), May. 2012.
- [53] A. S. Patwardhan, "Background Removal Using RGB-D data for Fright Recognition", International Research Journal of Engineering and Technology (IRJET), May. 2012.
- [54] A. S. Patwardhan, "Depth and Movement Data Analysis for Fight Detection", International Research Journal of Engineering and Technology (IRJET), May. 2013.
- [55] A. S. Patwardhan, "Human Action Recognition Classification using HMM and Movement Tracking", International Research Journal of Engineering and Technology (IRJET), May. 2013.
- [56] A. S. Patwardhan, "Feedback and Emotion Polarity Extraction from Online Reviewer sites", International Research Journal of Engineering and Technology (IRJET), May. 2014.
- [57] A. S. Patwardhan, "Call Tree Detection Using Source Code Syntax Analysis", International Research Journal of Engineering and Technology (IRJET), May. 2014.
- [58] A. S. Patwardhan, "Walking, Lifting, Standing Activity Recognition using Probabilistic Networks", International Research Journal of Engineering and Technology (IRJET), May. 2015.
- [59] A. S. Patwardhan, "Online News Article Temporal Phrase Extraction for Causal Linking", International Research Journal of Engineering and Technology (IRJET), May. 2015.
- [60] A. S. Patwardhan, "Online Comment Processing for Sentiment Extraction", International Research Journal of Engineering and Technology (IRJET), May. 2016.