Human Activity Recognition Using Temporal Frame Decision Rule Extraction

Amol Patwardhan

Mechanical and Industrial Engineering Department, LSU, apatwa3@lsu.edu

---***---

Abstract – *Activities of humans and their recognition has many practical and real world applications such as safety, security, surveillance, humanoid assistive robotics and intelligent simulation systems. Numerous human action and emotion recognition systems included analysis of position and geometric features and gesture based co-ordinates to detect actions. There exits additional data and information in the movement and motion based features and temporal and time-sequential series of image and video frames which can be leveraged to detect and extract a certain actions, postures, gestures and expressions. This paper uses dynamic, temporal, time-scale dependent data to compare with decision rules and templates for activity recognition. The human shape boundaries and silhouette is extracted using geometric co-ordinate and centroid model across multiple frames. The extracted shape boundary is transformed to binary state using eigen space mapping and parameter dependent canonical transformation in 3D space dimension. The image blob data frames are down sampled using activity templates to a single candidate reference frame. This candidate frame was compared with the decision rule driven model to associate with an activity class label. The decision rule driven and activity templates method produced 64% recognition accuracy indicating that the method was feasible for recognizing human activities.*

Key Words: Emotion Recognition, Human Activity, Crowd Activity, Group Activity, Edge detection, Audi-Video data, 3D sensor, Rule, Kinect, Affect Recognition, Shape Extraction, Decision Model, Template, Temporal.

1. INTRODUCTION

Automatic recognition of human activity has gained attention of researchers in the past decade. The study by Researchers [1] examined the relation of motion and certain human activities. The study analyzed the movement of extracted features and shapes. Some studies [2] also used dynamic features and compared them with temporal templates developed using a key reference frame. Researchers [3] used graphical models to represent activity and used this graphs to recognize human actions. A study [4] used shape extracted from images and matched them with a reference shape. Human behavior analysis was done using probabilities in a study [5]. Researchers [6] used real time evaluation of human actions to check the recognition accuracy of the automated systems. Human gait and facial expression shape analysis was done by Huang and Nixon [7]. The study used canonical space representation for dimensionality reduction. Study by Yamato et. al [8] used HMM in sequential images to detect human actions. Techniques such as fuzzy rules and view-invariant analysis have been done in studies [9], [10] for recognition of human actions. In addition to human activity recognition several studies have focused on recognition of emotions from body display [11], [14], [17], [18], [19], [22], [23]. Additional studies in the last decade have also focused on multimodal emotion recognition instead of unimodal or bimodal data [25], [26], [27],[28] and [35]. Many research studies [36], [37], [38], [39] have also focused on 3D, sensor based human activity recognition, and effective software implementations to support real time processing. This paper focused on using decision rule based evaluation of extracted features by comparing against a candidate human activity shape.

2. METHOD

For the purpose of this study, 7 participants enacted the action of walking in a room and picking up objects. The actions were recorded using 1 camera for the frontal view and another camera for the side view. The human shape was then extracted using geometric model across multiple frames. The extracted shape was transformed to binary state using eigen space mapping and parametric transformation in the canonical space. The image data frames were then down sampled using activity templates to a single candidate frame. This candidate frame was finally compared with the decision rule driven model to associate with an activity class label. For the decision rule driven comparison the window of 16 x 16 was used. For each superimposed frame on the extracted shape image, the pixel intensity and the color was compared with the reference frame. The window was then moved to the next 16 x 16 block on the image. For each comparison the rule was evaluated to see if the difference in the intensities was above or below threshold and the Euclidean distance between colors was above or below the threshold. The results of the 2 rule evaluation was coded into a true = 1 and false = 0 value. Thus a binary matrix consisting of 1s and 0s was obtained. The total of 1s was counted to calculate a similarity score for the image mesh. If the score was higher than the threshold then the activity was recognized as walking in the room and picking up objects else it was classified as the activity of neutral standing pose.

Table -1: List of Actions

List of Actions			
Walk around in room	2 actors	Walk and sit down	2 actors
Pick up book	2 actors	Get up from chair	2 actors
Lift a box	4 actors	Get up from couch	2 actors
Lift a table	5 actors	Sit on couch	2 actors
Move a chair	1 actors	Lift a flower pot	2 actors

The above table shows the list of actions performed by the actors. The actors used this as reference to enact or spontaneous move around the room, sit, pick objects, lift objects, read and get up. The actions were performed individually or in groups. The same actions were repeated in different lighting conditions. The actors were allowed to improvise and deviate from script. This allowed testing the classification on real life like scenario.

3. RESULTS

The recognition accuracy for neutral was better (67% and 75%) as compared to the candidate activity of picking up objects (64% and 71%) for both lighting conditions (dim and controlled).

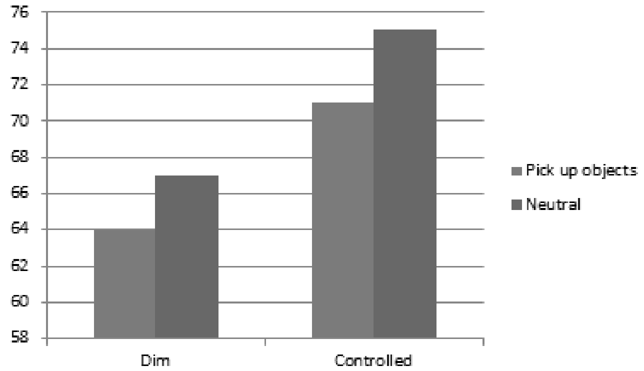

Chart -1: Comparison of Recognition accuracy in dim and controlled lighting for individual person

The accuracy was lower for dim lighting conditions for both the neutral standing pose and the walking and picking up object action.

We also evaluated the recognition accuracy for a group a people performing the neutral standing pose and pickup up objects actions under dim and controlled lighting.

In this scenario the recognition accuracy for neutral was better (63% and 71%) as compared to the candidate activity of picking up objects (59% and 65%) for both lighting conditions (dim and controlled). The accuracy was lower for dim lighting conditions for both the neutral standing pose and the walking and picking up object action.

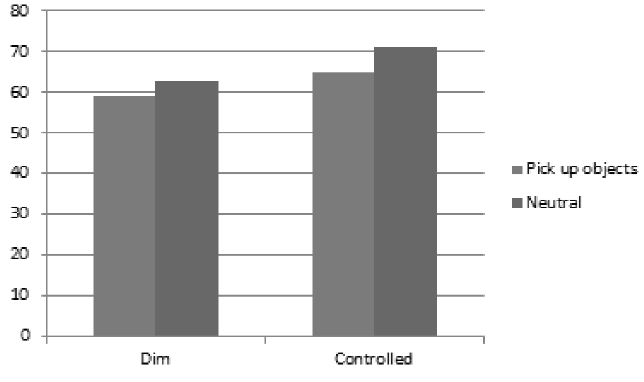

Chart -2: Comparison of Recognition accuracy in dim and controlled lighting for group of people.

But overall there was a drop in the recognition accuracy when compared with individual's performing the same actions instead of a group of people.

4. CONCLUSIONS

We did not find conclusive evidence that the decision rule based pixel analysis inside a 16 x 16 window was better for human activity recognition. The neutral pose recognition accuracy was better than picking up object accuracy. This was expected because the neutral position does not contain noise whereas the picking up object action involves a lot of movement which if not captured in the rule based evaluation may result in lower recognition accuracy. As a future scope other techniques such as HMM, 3D RGB-D sensor based tracking and image processing will be explored and compared with the method in this study to improve recognition accuracy.

REFERENCES

[1] A.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.

[2] A.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.

[3] A.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.

[4] A.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.

[5] A.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.

[6] A.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.

[7] A.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.

[8] A.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.

[9] A.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.

[10] A.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.

[11] A.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.

[12] A.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.

[13] A.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.

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

[15] 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.

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

[17] 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.

[18] A. S. Patwardhan, "An Architecture for Adaptive Real Time Communication with Embedded Devices," LSU, 2006.

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

[20] 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).

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

[22] A. S. Patwardhan, "Embracing Agile methodology during DevOps Developer Internship Program", IEEE Software (in review), 2016.

[23] A. S. Patwardhan, "Edge Based Grid Super-Imposition for Crowd Emotion Recognition", International Research Journal of Engineering and Technology (IRJET), May. 2010.

[24] A. S. Patwardhan, "Human Activity Recognition Using Temporal Frame Decision Rule Extraction", International Research Journal of Engineering and Technology (IRJET), May. 2010.

[25] A. S. Patwardhan, "Low Morale, Depressed and Sad State Recognition in Confined Spaces", International Research Journal of Engineering and Technology (IRJET), May. 2011.

[26] A. S. Patwardhan, “View Independent Drowsy Behavior and Tiredness Detection”, International Research Journal of Engineering and Technology (IRJET), May. 2011.

[27] A. S. Patwardhan, “Sensor Based Human Gait Recognition for Drunk State”, International Research Journal of Engineering and Technology (IRJET), May. 2012.

[28] A. S. Patwardhan, “Background Removal Using RGB-D data for Fright Recognition”, International Research Journal of Engineering and Technology (IRJET), May. 2012.

[29] A. S. Patwardhan, “Depth and Movement Data Analysis for Fight Detection”, International Research Journal of Engineering and Technology (IRJET), May. 2013.

[30] A. S. Patwardhan, “Human Action Recognition Classification using HMM and Movement Tracking”, International Research Journal of Engineering and Technology (IRJET), May. 2013.

[31] A. S. Patwardhan, “Feedback and Emotion Polarity Extraction from Online Reviewer sites”, International Research Journal of Engineering and Technology (IRJET), May. 2014.

[32] A. S. Patwardhan, “Call Tree Detection Using Source Code Syntax Analysis”, International Research Journal of Engineering and Technology (IRJET), May. 2014.

[33] A. S. Patwardhan, “Walking, Lifting, Standing Activity Recognition using Probabilistic Networks”, International Research Journal of Engineering and Technology (IRJET), May. 2015.

[34] A. S. Patwardhan, “Online News Article Temporal Phrase Extraction for Causal Linking”, International Research Journal of Engineering and Technology (IRJET), May. 2015.

[35] A. S. Patwardhan, “Online Comment Processing for Sentiment Extraction”, International Research Journal of Engineering and Technology (IRJET), May. 2016.

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

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

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

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