Development of Behavioral Based System from Sports Video

S.Muthu lakshmi

Research Scholar

N. Akila Banu *Student*

Dr G.V.Uma

Professor, Research Supervisor & Head

email:muthusubramanian2185@gmail.com

Dept. of Information Science & Technology

Anna University, Chennai

Abstract- *A system for detecting and analyzing behavior of a sports person from their facial expression extracted from a sports video from the basis of this project.* ***Shot Segmentation, Object Frame Selection, Image Segmentation, Facial Feature Extraction and Facial Expression Recognition (FER)*** *are the major steps included in developing Cognitive Analysis of Facial Expression Recognition system from Sports Video. The overall outcome involves in making the system recognize the basic expressions such as happy, sad, anger, disgust, surprise and fear along with neutral expression of a sports person.*

Key terms- *Shot Segmentation, Object Frame Selection, Image Segmentation, Facial Feature Extraction and Facial Expression Recognition(FER), Clustering, Thresholding.*

1. INTRODUCTION

With the massive explosion in the area of digital videos, entertainment and communication in our day-to-day life sports video analysis has attracted great attention due to its mass attraction and greater commercial value. For cognitive learning of a sports person behavior which forms as our area of interest all that we first have to do is to analyze the sports video.

Facial Expression Recognition (FER) from videos has been growing interest amongst researchers from various fields, i.e., computer science, engineering, psychology, and neuroscience. This field is getting considerable attention recently because of its wide number of applications but here it’s a way of cognitive learning of a sports person behavior which is our center of focus.

The major objective of developing such system is to understand the objects behavior and store it so that sports professionals can easily retrieve to compare the performed behavior with one's performed by elite persons in order to improve their performance in further training, coaching and competition.

To fulfill such a challenging task the major focus lies on Shot Segmentation and Image Segmentation.

A video is a collection of a huge number of still frames and as the continuous change of frames exhibits a motion like feature, it is known as video. However, the video has different shots[1], which can be defined as a sequence of frames taken by a single camera without any major variation in the color content of consecutive videos. To process any video or to extract object from video it is necessary to segment the video by means of shots.

Another major focus is to develop image segmentation algorithms based on a clustering technique for facial image interpretation. The main goal is to develop new techniques which will result in improved facial expression recognition system for behavioral analysis.

2. RELATED WORKS

Some of the related works from other research is briefly reviewed in three aspects: Segmentation of Video Object, feature extraction and Facial Expression Recognition.

2.1 Segmentation of Video Object (VO)

A VO extraction scheme proposed in [2] had developed an object segmentation method [3], [4] which exploits simple frame difference with edge detection to effectively extract shape information of moving objects

2.2 Features Extraction

Feature Extraction proposed by Yi Ding and Guoliang Fan et al. [5] has adopted the concept of dominant color [6] to estimate the spatial color distribution, and Canny edge detection followed with the Hough transform to detect the yard lines and to compute their angles.

2.3 Facial Expression Recognition

Various methods which have been used in facial expression recognition include Fisher discrimination analysis [12], Locally Linear Embedding [13], Higher Order Singular Value Decomposition [16] and so on [7] [8] [15]. Among them the most successful ones are Neural Network [8] [9] [10] and Support Vector Machine [10] [11] [14].

3. ARCHITECTURE OF COGNITIVE ANALYSIS OF HUMAN BEHAVIOR

Architecture for Cognitive Analysis of Human Behavior in Sports Video is explained in Figure 1 after explanation of each phase given in detail.

Frame Selection:

First the raw sports video is segmented in to number of shots and the object frame where the sports person predominantly appears without any occlusion is selected.

Feature Extraction:

Then the extracted frame i.e., the image of sports person is segmented using Hard C Mean Clustering. Then the region properties of the segmented image are used to extract the facial features.

Facial Expression Recognition:

Finally, classifying expression by thresholding to develop an effective Facial Expression Recognition System for cognitive learning of a sports person Behavior.

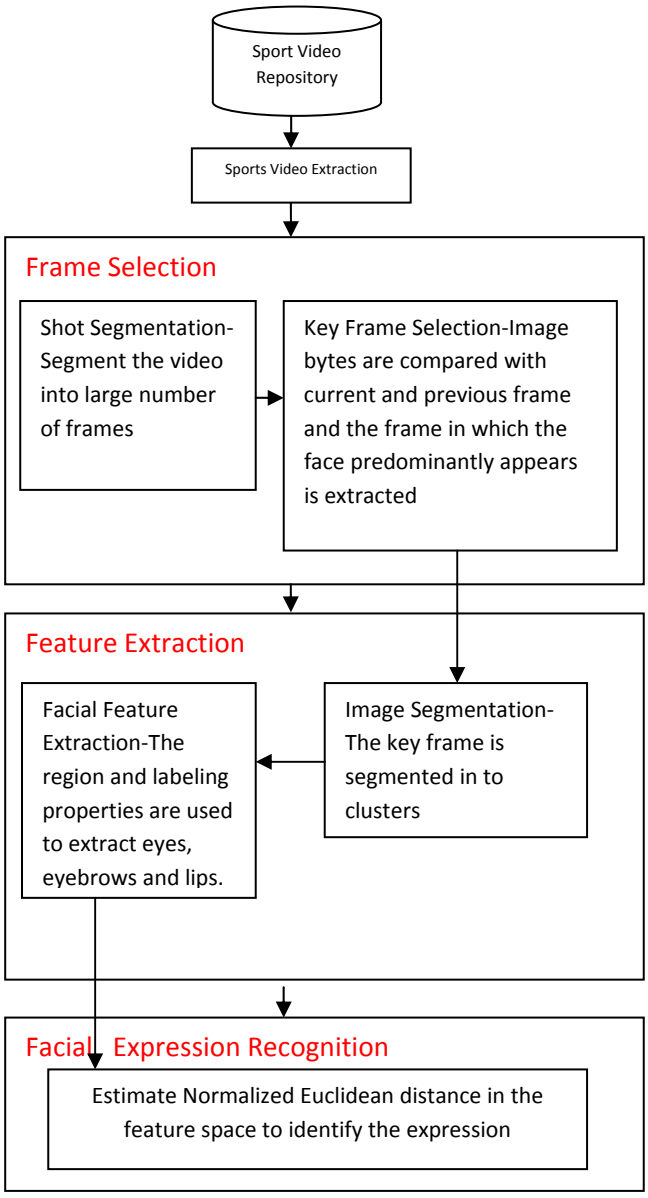


Figure 1: Architecture for Cognitive Analysis of Human Behavior in Sports Video

4. PROPESED MODEL

In order to create Cognitive Analysis of Facial Expression Recognition system from Sports Video to analyze a sports person behavior, *Shot Segmentation, Object Frame Selection, Image Segmentation, Facial Feature Extraction and Facial Expression Recognition (FER)* are the major steps involved. Fig. 2 explains the flow diagram of the proposed system.

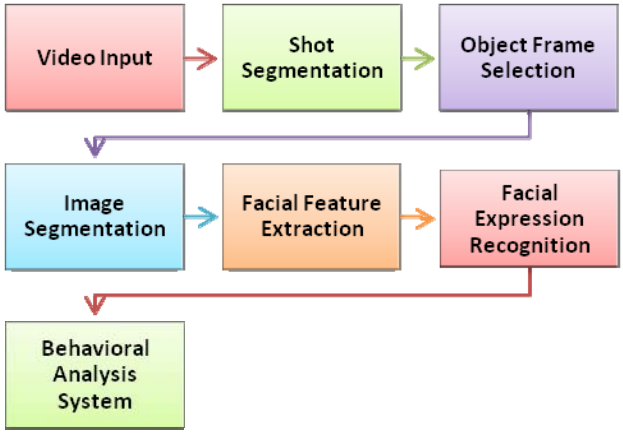


Figure 2: Flow Diagram for Cognitive Analysis of Human Behavior in Sports Video

3.1 Shot Segmentation

A frequent first step for the majority of video content analysis techniques is to segment video into elementary shots, each continuous in time and space[17]. Given below is the algorithm for implementing Shot Segmentation.

Algorithm for Shot Segmentation of a Sports Video:-

Input: Sports Video

Initialization: Initialize Processor, source handler, start frame and end frame.

Steps:

The steps involved in shot segmentation algorithm are as follows:

1. Open the video source
2. Create a processor for the video source and start
3. Configure the processor thus obtaining track controls
4. Get the output source from the processor and set it to source handler. Source handler is used to display information about each frame of data received.
5. For each frame buffer
    a. Convert buffer data to pixel data
    b. Save the jpeg to file
6. Until end frame is reached

Output: Sports Video segmented in to frames stored as jpg images.

A series of frames with uninterrupted camera motion is known as shot, while a series of shots that are consistent from the narrative as well as the users point of view is known as a clip.

Video Segmentation is a major maneuver in image sequence analysis. Video segmentation splits a video file into shots which is illustrated as a contiguous sequence of video frame that are recorded from a single camera operation.

3.2 Object Frame Selection

To select the frame in which the object predominantly appears, comparing each frame with the previous frame to extract the number of bytes occupied by each frame. A frame with higher number of byte size has more resolution clarity of the sports person captured distinctly. Graph Cut Algorithm[29] is used to extract the foreground object.

3.3 Image Segmentation

Segmentation is an important part of any automated image recognition system because at this moment only the objects of interest are extracted for further processing, such as, description or recognition. Image segmentation is done by supervised and unsupervised clustering techniques. The unsupervised clustering techniques used here is Hard C Means. Given below is the algorithm for implementing Image Segmentation.

Algorithm for Image Segmentation:-

Input: Sports Person Image with object alone i.e., segmented foreground.

Initialization: Initialize Partition Matrix, U(r) ; cluster to 2 ; convergence to 0.01 and receive the data matrix.

Steps:

The steps involved in the Hard C means algorithm are as follows:

1. Receive the data matrix x
2. Fix the number of clusters, c ($2 \le c \le n$), where n is the length of the image data
3. Initialize the partition matrix, U(r)
4. Calculate the cluster centers, Y using equation
5. $$y_j = \frac{\sum_{k=1}^{n} (u_{jk})\ x_k}{\sum_{k=1}^{n} (u_{jk})}$$
6. Calculate the Euclidean distance matrix, D using equation

   $$E_j(x_k) = \| x_k - y_j \|^2$$
7. Update the partition matrix, U(r) using equation

   $$u_{jk} = \begin{cases} 1 & if\ E_j(x_k) \le E_l(x_k) \quad \forall j,k,l \quad and \quad l \ne j \\ 0 & Otherwise \end{cases}$$
8. Check for convergence

   If max || U(r) - U(r+1)||< ε stop else repeat steps 4–6

Output: Set of segments that collectively cover the entire image, or a set of contours.

3.4 Facial Feature Extraction

When receiving a clustered image and extracting the object features, not only the facial features are extracted but also objects body features. Then using various region properties on the segmented image the system retrieves only the objects facial features. Given below is the algorithm for implementing facial feature extraction.

Algorithm for Facial Feature Extraction:-

Input: Segmented image obtained from Hard C Mean Algorithm

Initialization : Convert the gray scale cluster to black and black cluster to white, imadjust()

Steps :

- Label all the regions, bwlabel()
- Get the region properties of the image, this will segment the portions of hair, brows, eyes nose and lips, and all the body feature are extracted, regionprops()
- Using area , orientation the needed features alone can be extracted.

Output: Extracted features for FER analysis, such as eyes brow, eyes and lips.

3.5 Facial Expression Recognition(FER)

Finally various dimensionalities and thresholding of the extracted features are used to classify the expression and develop an effective Facial Expression Recognition System for understanding the behavior.

5. RESULT & OUTPUT

The outcome of each phase of the system is given below. Figure 3 shows some the basic expressions such as *happy, sad, anger, disgust, surprise and fear along with neutral expression* of a sports person. Even though there are lots of facial expressions only seven are identified as universal expressions. When it comes to a sports person he expresses his behavior more predominantly through his facial expressions. A sports person is happy and surprised when one wins or achieves their goal or target. Similarly a sports person is sad or disgust when one loses or fails to achieve their goal or target. Anger and fear are the expressions commonly seen when a sports person is started towards or about to achieve their goal or target.

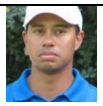
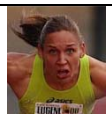
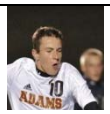
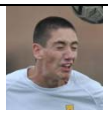
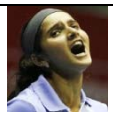
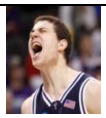
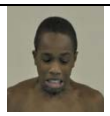
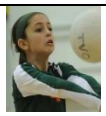
Expressions	*Some of the example datasets*		
happy			
sad			
anger			
disgust			
surprise			
fear			
neutral			

Figure 3: Some examples of basic expressions of sports persons.

The input and output of the FER system is shown in Figure 4. Figure 4: (a) shows human facial features being extracted. Figure 4: (b) shows a sports video. Figure 4: (c) shows image of the sports person being segmented and feature being extracted.

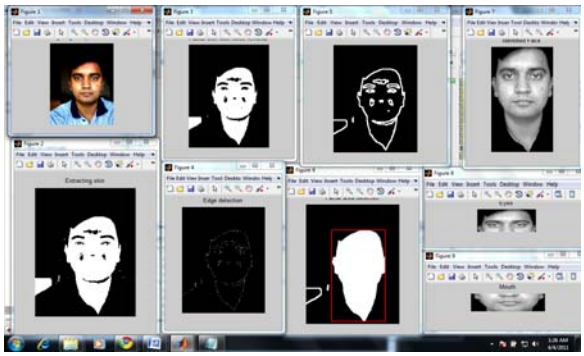


Figure 4: (a) Human face detection and features extraction

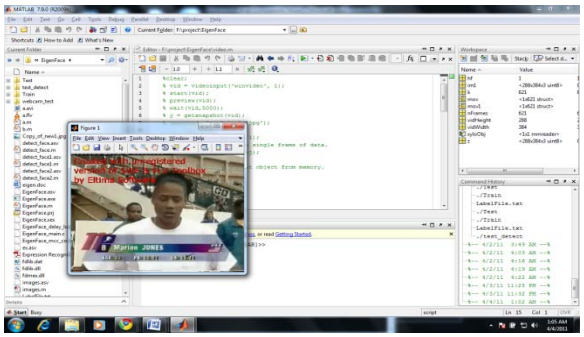


Figure 4: (b) A sports video

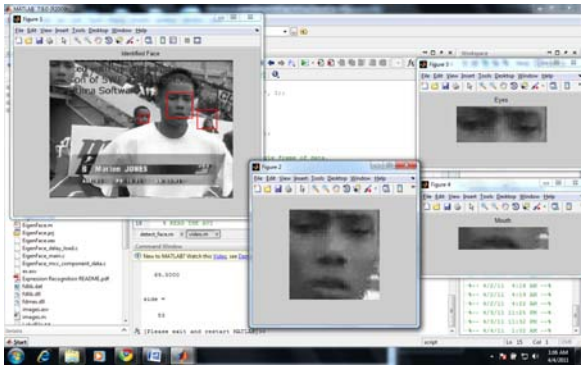


Figure 4: (c) Sports person face detection and features extraction

6. CONCLUSION

A more rising research area which has been in the limelight recently amongst researchers and experimenters is the developing facial expression recognition system. However proposed work is with respect to sports domain. The first and foremost process of the proposed system is it fragmented the long video sequence into shots. Subsequently, the object i.e., the sports man is extracted. With the extracted object the system segment the image in to clusters. Then the regions properties are applied segmented image are used to extract the facial features. Finally, expressions are classified by thresholding thus developing an effective Facial Expression Recognition System for analysis of the sports person Behavior. Thus a cognitive human behavioral system for sports domain have been developed by analyzing facial expression such as happy, sad, anger, disgust, surprise and fear along with neutral expression from shots extracted from sports video.

7. REFERENCES

[1] Geeta P, Vasumathi Narayanan(2008) A Survey of Content-Based Video Retrieval Journal of Computer Science 4(6):474-486.

[2] Huang-Chia Shih, Jenq-Neng Hwang and Chung-Lin Huang, "Content-Based Attention Ranking Using Visual and Contextual Attention Model for Baseball Video", IEEE Transaction on multimedia, Vol.11, NO.2, Feb 2009.

[3] W. Hua, M. Han, and Y. Gong, "Baseball scene classification using multimedia features," in Proc. IEEE Int. Conf. Multimedia and Expo, 2002.

[4] L. Xie, S.-F. Chang, A. Divakaran, and H. Sun, "Structure analysis of soccer video with hidden Markov models," in Proc. Int. Conf. Acoustic, Speech and Signal Processing, Orlando, FL, 2002.

[5] Yi Ding and Guoliang Fan, "Sports Video Mining via Mulitichannel Segmental Hidden Markov Model", IEEE Transaction on multimedia, Vol.11, NO.7, Nov 2009.

[6] Ekin, A. Tekalp, and R. Mehrotra, "Automatic soccer video analysis and summarization," IEEE Trans. Image Process., vol. 12, no. 7, pp. 796-807, Jul. 2003.

[7] P. Aleksic and A. Katsaggelos. Automatic facial expression recognition using facial animation parameters and multistream hmms. *Information Forensics and Security, IEEE Transactions on*, 1:3–11, March 2006.

[8] Kotsia and I. Pitas. Facial expression recognition in image sequences using geometric deformation features and support vector machines. *IEEE Transactions on Image Processing*, 16(1):172–187, 2007.

[9] L. Ma and K. Khorasani. Facial expression recognition using constructive feedforward neural networks. *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, 34(3):1588–1595, 2004.
[10] H. Kobayashi, A. Tange, and F. Hara. Real time recognition of six basic facial expressions. *Proceedings of the IEICE General Conference*, 1995:256, 19950327.
[11] P. Michel and R. E. Kaliouby. Real time facial expression recognition in video using support vector machines. In *ICMI '03: Proceedings of the 5th international conference on Multimodal interfaces*, pages 258–264, New York, NY, USA, 2003. ACM.
[12] Y. Shinohara and N. Otsu. Facial expression recognition using fisher weight maps. In *FGR*, pages 499–504. IEEE Computer Society, 2004.
[13] Y.-K. Wu and S.-H. Lai. Facial expression recognition based on supervised lle analysis of optical flow and ratio image. *International Computer Symposium*, Nov 2006.
[14] Q. Xu, P. Zhang, W. Pei, L. Yang, and Z. He. A facial expression recognition approach based on confusion-crossed support vector machine tree. In *IIH-MSP*, pages 309–312. IEEE Computer Society, 2006.
[15] L. Yin and X. Wei. Multi-scale primal feature based facial expression modeling and identification. In *FGR '06: Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition*, pages 603–608, Washington, DC, USA, 2006. IEEE Computer Society.
[16] H. Wang and N. Ahuja. Facial expression decomposition, 2003.
[17] Hoi-Kok Cheung, Wan-Chi Sui, Feng(2005) New Block-Based Motion Estimation For Sequence with Illumination Variation And Its Application To Video Mosaicking. In: proceedings of IEEE International Symposium on Circuits and Systems 3:2903-2906.