Mosaic Based Characterization of Video Sequences using Fuzzy Inferencing

Pranav Bhushan R.S. Jadon Santanu Chaudhury
Department of Electrical engg., IIT Delhi
New Delhi-110016 INDIA
E-mail: santanuc@cse.iitd.ernet.in

Abstract— **The current paper deals with an approach to characterize and categorize video sequences. Here, image mosaicing is used as a tool to represent the redundant frame based video information into concise and compact scene based information called *mosaics*. Thereafter, fuzzy inferencing is used to classify the mosaics into one of the relevant categories of sports like cricket, golf and football. For categorization mosaic and frame based information has been used in an opportunistic fashion.**

Keywords— **Video-mosaicing, video-classification, fuzzy-inferencing, video-indexing.**

I. Introduction

Characterization and categorization of the video data is itself a complex problem. Several work has already been done in the field of content based retrieval of video. Frame based approach like [3] deals with the representation of video in terms of r-frames, which are used as still images for further analysis. These r-frames are generated for each shot and used for further analysis. The Netra-V system [2] uses a region based content description scheme using low level visual descriptors. It uses an object based video representation for functionalities such as search and retrieval of video objects. The strategy adopted by Netra-V is to build a low-level video content description that can be completely automated. Specific techniques for compact video representations like [4] are used to exploit the contents of video shots and view them as a synoptic frame composed of a mosaic image of the background. Approach like [1] uses mosaicing technique to represent and index video sequences. Its advantage is that it deals with the methods that complement the traditional content based indexing methods. However, in [1] mosaics have not been used for categorization of video sequences.

In the given paper we present an approach to first represent video data in terms of mosaics and there after use these mosaics to classify the video sequences into one of the categories like cricket, golf or football. For categorization we have used fuzzy logic based reasoning scheme. Use of fuzzy techniques enabled us to use approximate attributes and relations for tackling inherent variabilities of the video data. We have proposed a new technique to represent fast spatial configuration change among objects in the video sequence. We have used frame based properties along with mosaic based information in an opportunistic fashion. Experimental results have validated our approach.

II. Mosaic Generation

Collection of overlapping images put together with coordinate transforms relating the different image coordinate systems describe Mosaics. The video information distributed over many frames at the cost of very high temporal redundancy is converted into compact representation in terms of Mosaics. For mosaic generation we have considered the approach as described in [6]. The whole implementation was done on the SGI (IRIX 6.3) Workstation employing Image Vision (il), Digital Media (dmedia), X11 and Image etc libraries. The result for a cricket video sequence (15..50 frames) is as shown in Figure 2. Here the camera is moving straight or rather panning for the straight drive of the cricket shot.

Mosaics [1] can represent video information given the constraints of camera model, motion and lightness changes are accounted for. Otherwise we may also get strange results and are beyond the reach of the convergence of the algorithm.

Fig. 1. Individual Frames Of a video sequence(frames 15,30,50)

Fig. 2. Result for cricket sequence shown in 1

III. Content Extraction from Mosaics

For the process of identification of *Mosaics* we statistically characterize different regions with different key domain-independent parameters covering shape, color and location. This task is accomplished by calculating these features so that a given image or a mosaic can be effectively classified or categorized. We have developed a segmentation algorithm based on the hue, saturation and intensity (HSI) color space for segmenting the mosaics. This is because of the advantage of the HSI color space in characterizing colors; for example the distinction in shades of green like light green, dark green, pale green etc. can be done by considering only hue values. Segmentation is done by transformation from RGB to HSI space through a quantization and thereby grouping the regions falling in the similar hue, saturation and intensity values. Our segmentation approach is somewhat similar to the technique as described in [5]. The color match similarity metric is also taken as described in [5]. Some segmentation results are as shown in Figure 3 for cricket mosaic and in Figure 4 for golf mosaic. The advantage of our scheme is that here we are getting rid of multiple smaller regions which are automatically grouped into bigger regions, having nearly same hue values, for example in the case of pitch as shown in Figure 3. Thus the pitch itself is clear and distinct and shown as an isolated region.

(a)

(b)

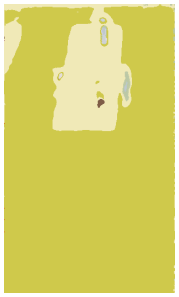

Fig. 3. Segmentation results: (a)Original cricket mosaic, (b) Segmentation using using HSI model

Then several region based features like centroid, area, perimeter, moment invariants, etc. are extracted for further analysis.

IV. Video Sequence Classification

For video sequence classification we have computed fuzzy attributes of the segmented regions and spatial relations between these regions. Fuzzy rules based on these features have been used for categorization of the video sequences.

A. Fuzzy Attributes

Different fuzzy variable like green score, pitch score, yellow color match, rectangular shape match, man score, elliptical score, surroundedness in grass ,etc. are considered. Here, we describe the scheme adopted for some of the fuzzy attribute

Fig. 4. Segmentation results: (a) Original golf mosaic, (b) Segmentation using HSI Model

computation.

A.1 Green Score

The variable green score is determined by taking into account the total green score of the image. This is done by counting all the green color pixels having hue value near to 120 degrees. We normalize the green counted pixels by the area of image so as to get the over all green amount in the image. This calculated value is applied to the fuzzy classifier to fuzzy value for the green score.

A.2 Pitch Score

The pitch score is calculated by the color and the shape match of the image regions. For representing the somewhat rectangular shape of the region in the image we have calculated a mean feature vector from an ensemble of images. Since pitch is somewhat rectangular in shape so the match score with respect to the moments of a rectangular region is fuzzified.

A.3 Player Score

The values of the player score are also calculated wrt. to moment value match of a somewhat elliptical blob surrounded by green grass. The score is also calculated in a manner similar to that used for the pitch score case.

A.4 Pitch color score

Here we also describe another variable called pitch color score. The value we have chosen is found to be nearly same match for pale-dull-yellow colored regions. Since these regions can only represent the pitch. Its value is taken as the mean of hue,saturation and intensities values for a set of images. Match score with respect to this cluster is fuzzified.

B. Fuzzy Rules

The motivation for fuzzy rules can be in the terms of human perception of knowledge. There is a need to represent this knowledge based on out past experiences and skill. For example a cricket scene may consist of a field with certain degree of green color, a pitch somewhat rectangular in shape (may be of slightly yellowish-brown in color surrounded by grass), and some players in the ground. Such knowledge can be used for further rule evaluation and classification of mosaics or video. For the current approach we have employed several rules for our fuzzy inferencing engine. Some of them can be elaborated here as:

- Rule1: If (Green-score is high) AND (Pitch-score is negligible AND (Man-score is single) THEN the match may be a *GOLF* sequence.
- Rule2: If (Green-score is medium) AND (Pitch-score is negligible) AND (Man-score is single) THEN the match may be a *GOLF* sequence.
- Rule3: If (Green-score is high) AND (Pitch-score is negligible) AND (Man-score is many) THEN the match may be a *FOOTBALL* sequence.
- Rule4: If (Green-score is medium) AND (Pitch-score is negligible) AND (Man-score is many) THEN the match may be a *FOOTBALL* sequence.
- Rule5: If (Green-score is small) AND (Pitch-score is negligible) AND (Man-score is many) THEN the match may be a *FOOTBALL* sequence.
- Rule6: If (Green-score is high) AND (Pitch-score is large) AND (Man-score is single) THEN the match may be a *CRICKET* sequence.
- Rule6: If (Green-score is high) AND (Pitch-score is large) AND (Man-score is single) THEN the match may be a *CRICKET* sequence.

- Rule7: If (Green-score is high) AND (Pitch-score is large) AND (Man-score is many) THEN the match may be a *CRICKET* sequence.
- Rule8: If (Green-score is high) AND (Pitch-score is medium) AND (Man-score is single) THEN the match may be a *CRICKET* sequence.
- Rule9: If (Green-score is high) AND (Pitch-score is medium) AND (Man-score is many) THEN the match may be a *CRICKET* sequence.
- Rule10: If (Green-score is medium) AND (Pitch-score is large) AND (Man-score is many) THEN the match may be a *CRICKET sequence.*

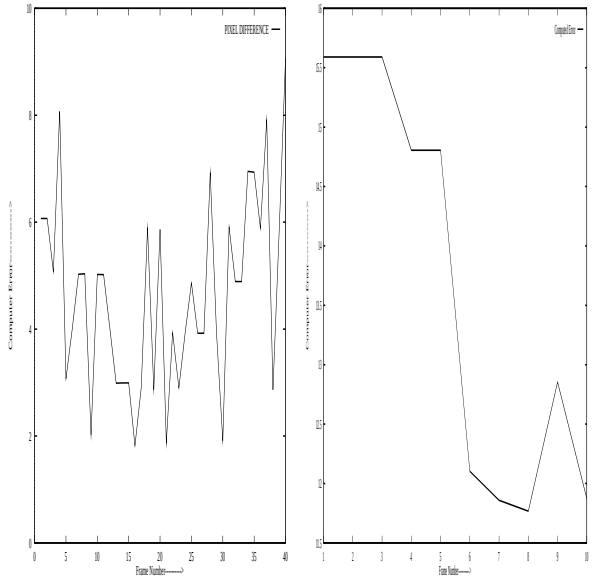

Fig. 5. Graph showing error for (a) Football sequence (b) cricket sequence

C. Exploiting Frame Based Information

Since panoramic mosaics can not capture dynamic information content of the video we propose to combine frame based attributes with mosaic based information for fuzzy categorization. In the context of current problem domain we have used frame based scheme for computation of the change in spatial configuration among the players. The change can be calculated in terms of a changing polygon shape made by the initial points at the corners of the image and the players present in the image.So, for the cases of sequences having fast relative motion between the objects here, *players* the shape of the polygon changes and the amount of error is significant. As contrast to rather static objects or *players* the subsequent error is less. Some of the results of the football and cricket sequences are as shown in Figure 5.

From the graphs it is clear that there is a large variation in the error for the football sequence (frames) as compared to that of cricket. In other words we can say that,in general, there is a large variation associated with the football sequence because of the fast motion of the players. This attribute is captured through a fuzzy variable and encoded in fuzzy rules. These rules are fired when the maximum membership value obtained for the categories after mosaic based processing does not exceed 0.6.

File Name	Cricket	Golf	Football
(c1.ppm)	0.608050	0.466407	0.280638
(c2.ppm)	0.949325	0.230734	0.438150
(c3.ppm)	0.590072	0.283898	0.272341
(c4.ppm)	0.678893	0.329069	0.313321
(c5.ppm)	0.907895	0.397286	0.483226
(c6.ppm)	0.604081	0.452852	0.278806
(c7.ppm)	0.419162	0.452209	0.193460
(g1.ppm)	0.354282	0.783548	0.163515
(g2.ppm)	0.510706	0.783637	0.441208
(fb3.ppm)	0.399765	0.556414	0.184507

TABLE I
RESULTS SHOWING FUZZY SCORES OF MOSAICS FOR CRICKET,GOLF AND FOOTBALL SEQUENCES

V. CLASSIFICATION RESULTS

The fuzzy classification results of mosaics of some sequences based on the fuzzy analysis are as shown in table I. Entries (1-7) are cricket mosaics, (8,9) are golf mosaics and (10) is a football sequence. Results corresponding to 8 and 10 indicate the requirement for frame based processing. After frame based processing membership score for football in case of 10 increases to 0.7012. The result for 8 does not show clear disambiguation. In terms of overall statistics, using a large set of example sequences we have achieved an overall accuracy of around 89

VI. Conclusion

From the results of the fuzzy classification it is clear that the categorization and classification of the sequences is done effectively. The approach for the spatial configuration change can be further employed as a frame based approach along with mosaics as an aid to represent games like football etc.

References

[1] P. Anandan and M. Irani. Video Indexing Based on Mosaic Representations. In *Proceedings of IEEE*, volume 86, pages 905–921, May 1998.

[2] Y. Deng and B.S. Manjunath. Netra-v towards an Object based Video Representation. *IEEE Tran. on Cirts and Systems for Video Technology*, 8(5), Sept. 1998.

[3] Myron Flickner et al. Query by image and video content : the Qbic system. *IEEE Computer*, pages 23–32, September 1995.

[4] Mark Gelgon and Patrick Bouthemy. Determining a structured spatio-temporal representation of video content for efficient visualisation and indexing. In *IRISA/INRIA, Campus Universitarie de Beaulieu, France.*

[5] J.R. Smith and Shih-Fu Chang. VisualSEEK: a fully automated content-based image query system. *ACM Multimedia*, November 1996.

[6] Richard Szeliski. Video Mosaics for Virtual Environments. *IEEE Computer Graphics and Applications*, 16(2):22–30, March 1996.