

SPORTS VIDEO CATEGORIZATION BY MULTICLASS SVM USING HIGHER ORDER SPECTRA FEATURES

Dr. Saju Mohanan,
Department of Information Technology,
Higher College of Technology,
Sultanate of Oman
saju.mohanan@hct.edu.om

Abstract: Video classification is one of the rising fields in the video analysis. A large number of videos are accessed by people's daily from television and internet. It is easy for humans to index the video from the collection of videos which contains news, cartoon, sports, comedy and drama. Among the categories, sports video plays a vital role due to their commercial demand. There is a similarity between the different sports video which makes the classification task difficult. In this study, the sports video categorization for five categories of sports like football, cricket, volleyball, tennis and basketball is presented. The sports video categorization system uses Higher Order Spectra Features (HOSF) for the feature extraction from video frames and multiclass Support Vector Machine (SVM) classifier for the classification of videos. The system gives average classification accuracy of 93.44% using HOSF and multiclass SVM classifier.

Keywords: Sports video categorization, Video frames, higher order spectra features, Multiclass SVM.

I. INTRODUCTION

Sports video categorization by measuring video without thresholding is discussed in [1]. The input video frames are represented with a color histogram and are enhanced to reduce the features. The similarity of the videos is compared by nearest neighbor classifier. Amateur sports video genre categorization is described in [2]. The dense trajectories are calculated by k -means clustering algorithm. The bins and Bag Of Words (BOW) values are calculated by the mean and standard deviation to prevent the trajectories of the players from the audience. Motion boundary histogram and histogram oriented gradients are used as default parameters. The radial basis function based SVM is used for classification.

Sport event videos categorization by trajectory and speed is described in [3]. The speed of motion in the video which represents the trajectory is measured by randomness in position vector magnitude and angular displacements in time series. The classification is made by Hidden Markov Model (HMM). Sports video classification using color feature and HMM is discussed in [4]. The input video frames are extracted by color features due to speed change in video frame. Then the classification is made by HMM.

Automatic sports video genre categorization is discussed in [5]. At first, the videos are significantly chosen by the key-frame clustering. The scale invariant feature transform is used to search the visual words in the Geometry Visual Phrase (GVP). Then GVP and visual words are connected to form the histogram enhancement. Then the SVM classifier is used for classification. Categorization of sports video using camera motion parameters is described in [6]. Initially the video sequence is extracted from the camera motion extraction ratio and camera motion

transition. The nine sports video sequences are categorized by these two features which are analyzed separately.

Categorizing sports video by statistical analyzing and camera motion parameters is discussed in [7]. The video frames are extracted by camera motion parameters like camera motion transition and camera motion extraction ratio. The four types of camera motion parameters like shake, fix, zoom and pan are extracted. They are used to classify the videos. Sports videos for fast unsupervised ego-action learning are described in [8]. The motion feature is used for feature extraction after removing noises like motion blur, splashing water, glare and projectiles. The stacked Dirichlet process mixture model is automatically used for the histogram enhancement of the video frames.

Automatic sports genre classification using edge feature application is discussed in [9]. The input video frames are extracted by edge features. The k -nearest neighbor classifier is used for the classification of videos. Sports genre classification framework based on close up view and pre detection is discussed in [10]. The video key frames are identified by close up views in first level. BOW model is used in second level for feature extraction using scale invariant feature transform. SVM classifier is used for the sports genre classification.

An efficient technique for sports video categorization using HOSF and multiclass SVM classifier is presented. The organization of paper is as follows: In Section 2 materials and methods for sports video categorization are discussed. Section 3 the experimental results and discussion are presented. The last section concludes the sports video categorization.

II. METHODS AND MATERIALS

The sports video categorization system is achieved by two main stages; feature extraction and classification. In this study, HOSF is used for feature extraction and multiclass SVM is used for classification. In feature extraction stage, the HOSF is applied for each video sequence frame by frame to extract the features and saved in database. Then the multiclass SVM classifier is used for the categorization of sports videos. The block diagram of sports video categorization is shown in figure 1.

A. HOSF Feature extraction

HOSF extracts both phase information and amplitude of the given input. The cumulant spectra and moment spectra which are present in the HOSF can be used for deterministic and random process. The bi-spectrum for one-dimensional function is a two-dimensional function. So, the bi-spectrum of two-dimensional images is a four-dimensional function. The third order statistics namely, the bi-spectrum is used. The HOSF is also used in analysis of electro-encephalogram signals [11] and de-noising method for noise estimation [12]. The bi-spectrum is defined by,

$$H(h_1, h_2) = T[Z(h_1) Z(h_2) Z * (h_1 + h_2)] \quad (1)$$

where $Z(h)$ is the Fourier transform and $T[\cdot]$ denotes the expectation operation. The one-dimensional projection and two-dimensional bi-spectrum is employed by the average bi-spectrum such as it is defined by,

$$k_\theta(l) = \iint_s m(u, v) \cdot \gamma(u \sin \theta - v \cos \theta - l) dudv \quad (2)$$

where $k_\theta(l)$ is the projection of the image $m(u,v)$ at angle θ . Then the each one-dimensional projection is computed by the average bispectrum is defined by,

$$Bi(x_1, x_2) = \langle F(x_1).F(x_2).F(x_1 + x_2) \rangle \quad (3)$$

where $F(x_1)$ is the Fourier transform of $k_\theta(l)$. In this study HOSF is used to extract the input video frame by frame in the video sequences.

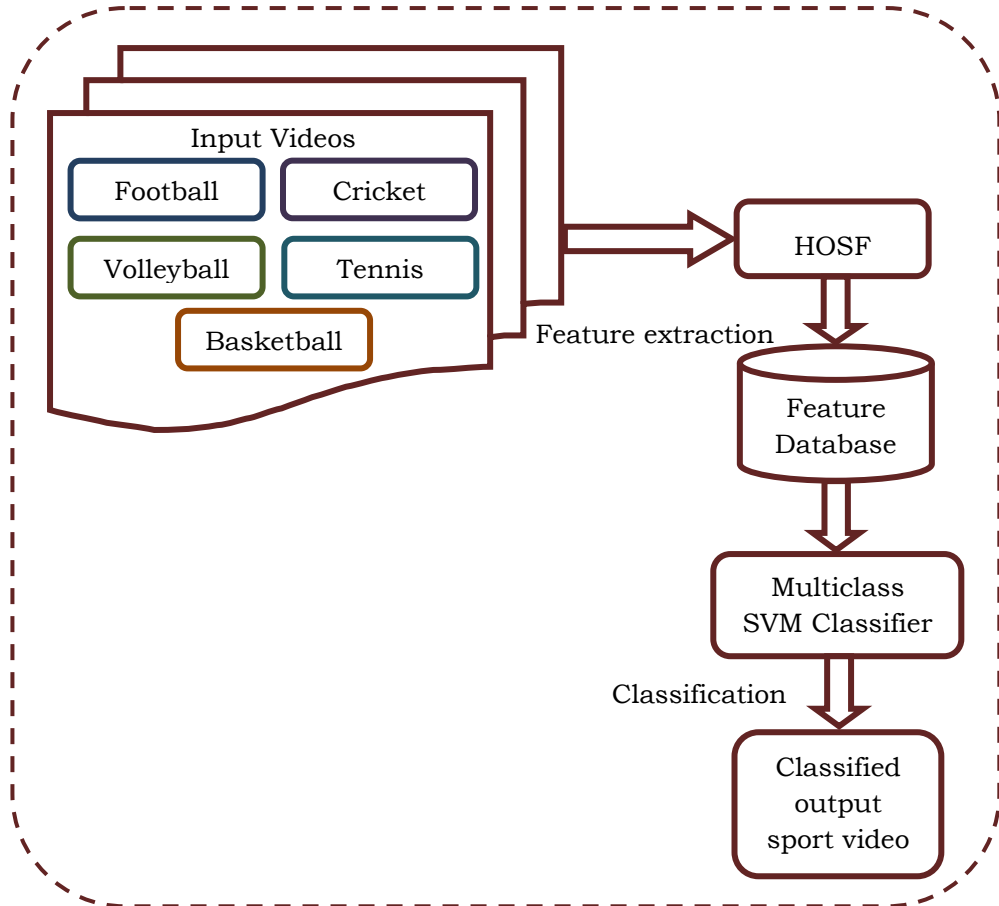


Fig. 1 Sports video classification system-workflow

B. Multiclass SVM classifier

SVM is a machine learning algorithm with superior properties. SVM classifier is easy to understand, robust and performs well in real time situations. Comparing to others SVM has global solutions. The SVM classifier automatically finds the support vectors for the better classification. Hyperplane which is constructed by the SVM classifier separates into two categories. SVM classifier has a high dimensional data and superior generality. Consider the input samples y_i and their corresponding target z_i . Then the hyperplane that separates the training sample set has the form,

$$(\delta \cdot z) + c = 0 \quad (4)$$

The SVM classifier constructs the hyperplane with maximum distance between the different classes. The set of vectors is separated by the hyperplane and the distance between the closest vectors in the hyperplane. SVM classifier is also used in glaucoma diagnosis [13], brain tumor detection [14] and license plate localization [15]. The multiclass SVM performs the classification of unknown sports video into the predefined class of video. The extracted feature from the input video sequence which is stored in feature database is given to multiclass SVM classifier for the categorization purpose among the five classes. The system uses k -fold cross validation. The k -fold cross validation technique is also used in supervised adaption technique [16]. The SVM classifier uses 10-fold cross validation for the classification. The SVM construct the hyperplane to separate the given number of classes.

III. RESULTS AND DISCUSSION

The performance of the sports video categorization system is evaluated by datasets which are obtained by television broadcast from different sessions. Figure 2 shows the sample video frames from each category in the database.



Fig. 2 Sample video frames in the database for five classes

Five hundred (500) video clips are collected from the television for performance evaluation. Among these 500 video clips, one hundred (100) video clips are from five sports category like cricket, football, volleyball, tennis and basketball. Each sports video clip is composed of 500 frames with 128x128 resolutions with 20 second duration. From the each category video the 50 clips for testing and 50 clips for training are randomly selected.

At first, the input video sequences are separated into frame by frame for performance evaluation. Then the HOSF technique is used to extract the features and stored in the database. Then, the multiclass SVM classifier with 10-fold indices is used to perform the sports video categorization. The performance of the sports video categorization is measured in terms of accuracy. Table 1 shows the classification accuracy of sports videos in each category.

TABLE 1 Classification accuracies of sports videos in each category

10-Fold cross validation	Classification Accuracy (%)				
	Cricket	Football	Volleyball	Tennis	Basketball
1	95	90	88	91	96
2	90	93	91	90	95
3	95	98	90	93	93
4	92	95	93	94	91
5	97	100	95	95	90
6	99	91	97	90	91
7	100	96	96	88	93
8	98	95	92	85	97
9	90	93	90	93	96
10	96	92	91	98	95
Average	95.2	94.3	92.3	91.7	93.7

From Table 1, it is observed that the maximum classification accuracy of 95.2% is obtained by classifying the cricket videos. The minimum classification accuracy of 91.7% is obtained by the tennis class using multiclass SVM. The other sports video classes like football, volleyball and basketball produces 94.3%, 92.3% and 93.7%. Only cricket and football classes produce 100% accuracy using multiclass SVM classifier in a single fold at least. The lowest classification accuracy of 85% is produced by tennis class using HOSF and multiclass SVM classifier.

IV. CONCLUSION

An efficient method for sports video categorization among different classes using HOSF and multiclass SVM is presented. The different sports video classes like, cricket, football, volleyball, tennis and basketball are used in this study. The input video clips are converted into frame by frame. Each video clips consist of 500 frames. These video frames are given to the HOSF module for feature extraction and stored in the database. Then the multiclass SVM classifier with 10-fold cross validation is used for the classification of sports videos into their respective category. Results show that the system provides average classification accuracy of 93.44% using HOSF and multiclass SVM classifier. Among the five different sports video category, the classification of cricket videos gives the better classification accuracy.

REFERENCES

- [1]. P. Mutchima, and P. Sanguansat, "A novel approach for measuring video similarity without threshold and its application in sports video categorization", IEEE International Conference on Pervasive Computing, Signal Processing and Applications, 2010, pp. 868-872.
- [2]. S.M. Safdarnejad, X. Liu, and L. Udpa, "Genre categorization of amateur sports videos in the wild", IEEE International Conference on Image Processing, 2014, pp. 1001-1005.
- [3]. S. Susan, S. Chaurawat, V. Nishad, M. Sharma, and S. Sahay, "Speed and trajectory based sports event categorization from videos", IEEE International Conference on Signal Processing and Communication, 2016, pp. 496-501.
- [4]. J. Hanna, F. Patlar, A. Akbulut, E. Mendi, and C. Bayrak, "HMM based classification of sports videos using color feature", IEEE International Conference Intelligent Systems, 2012, pp. 388-390.
- [5]. Y. Dong, J. Zhang, X. Chang, and J. Zhao, "Automatic sports video genre categorization for broadcast videos", IEEE Conference on Visual Communications and Image Processing, 2012, pp. 1-5.
- [6]. S. Takagi, S. Hattori, K. Yokoyama, A. Kodate and H. Tominaga, "Sports video categorizing method using camera motion parameters", International Conference on Multimedia and Expo, Vol. 2, 2003, pp. II-461.
- [7]. S. Takagi, S. Hattori, K. Yokoyama, A. Kodate, and H. Tominaga, "Statistical analyzing method of camera motion parameters for categorizing sports video", 2003, pp. 222-225.
- [8]. K.M. Kitani, T. Okabe, Y. Sato, and A. Sugimoto, "Fast unsupervised ego-action learning for first-person sports videos", In CVPR IEEE, 2011, pp. 3241-3248.
- [9]. Y. Yuan, and C. Wan, "The application of edge feature in automatic sports genre classification", IEEE Conference on Cybernetics and Intelligent Systems, Vol. 2, 2004, pp. 1133-1136.
- [10]. J. Zhang, Y. Dong, K. Tao, and X. Chang, "A modified sports genre categorization framework based on close-up view pre-detection", IEEE International Conference on Broadband Network and Multimedia Technology, 2011, pp. 301-305.
- [11]. S.A. Hosseini, M.A. Khalilzadeh, M.B. Naghibi-Sistani, and V. Niazmand, "Higher order spectra analysis of EEG signals in emotional stress states", IEEE International Conference on Information Technology and Computer Science, 2010, pp. 60-63.
- [12]. J. Liu, and P. Wu, "Higher order spectra denoising method of phase matching noise estimation", International Conference on Pervasive Computing, Signal Processing and Applications, 2010, pp. 811-814.
- [13]. C. Srinivasan, S. Dubey, and T.R. Ganeshbabu, "Complex Texture Features For Glaucoma Diagnosis Using Support Vector Machine", International Journal of MC Square Scientific Research, Vol. 7, No. 1, 2015, pp. 84-90.
- [14]. T.S. Kumar, K. Rashmi, S. Ramadoss, L.K. Sandhya, and T.J. Sangeetha, "Brain tumor detection using SVM classifier", IEEE International Conference on Sensing, Signal Processing and Security, 2017, pp. 318-323.
- [15]. D. Yang, J. Kong, N. Du, X. Li, and X. Che, "A novel approach for license plate localization based on SVM classifier", IEEE International Conference on Information Management and Engineering, 2010, pp. 655-660.

- [16]. D.R. Caon, A. Amehraye, J. Razik, G. Chollet, R.V. Andreão, and C. Mokbel, "Experiments on acoustic model supervised adaptation and evaluation by k-fold cross validation technique", International Symposium On I/V Communications and Mobile Network, 2010, pp. 1-4.