Sports classification using cross-ratio histograms

Balamanohar Paluri, Nalin Pradeep S, Hitesh Shah, and Prakash C

Sarnoff Innovative Technologies Pvt Ltd

Abstract. The paper proposes a novel approach for classification of sports images based on the geometric information encoded in the image of a sport's field. The proposed approach uses invariant nature of a cross-ratio under projective transformation to develop a robust classifier. For a given image, cross-ratios are computed for the points obtained from the intersection of lines detected using Hough transform. These cross-ratios are represented by a histogram which forms a feature vector for the image. An SVM classifier trained on aprior model histograms of cross-ratios for sports fields is used to decide the most likely sport's field in the image. Experimental validation shows robust classification using the proposed approach for images of Tennis, Football, Badminton, Basketball taken from dissimilar view points.

1 Introduction

The exponential growth during the last decade of photographic content has fueled the requirement for intelligent content management systems. One of the essential part of such systems is an automated classification of image content. In this paper, we address identification of sports based on the sport field in image using a robust classification mechanism.

Conventional approaches discussed in the literature for sport identification are primarily related to video, for example Wang et. al [1] distinguish sports videos shot using color and motion features. Takagi et. al [2] proposed a HMM based video classification system using camera motion parameters, Kobla et. al [3] applied replay detection,text and motion features with Bayesian classifiers to identify sports video. Messer et. al [4] employed neural networks and the texture codes on semantic cues for the same purpose. More recently, Wang et. al [5] classified sports videos with pseudo-2D-HMM using visual and audio features. In these approaches, sports classification is dependent on cues like color, texture and spatial arrangement, which are not preferable to use due to the variability of these features between images of the same sport for two different fields and views. Also, none of the currently existing approaches applied to classify sports videos can be directly used for classifying images of sports field as they employ temporal features also.

As opposed to existing approaches, we have opted to use the geometric information of a field to identify the sport. The motivation for this approach is attributed to the fact that sports fields have dominant geometric structures on

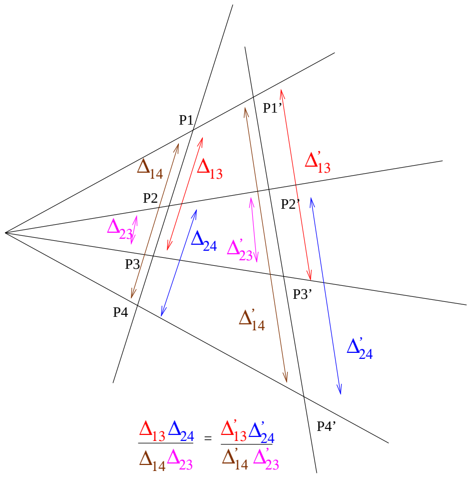


Fig. 1. Invariant nature of cross ratio for four collinear points subjected to a projective transformation

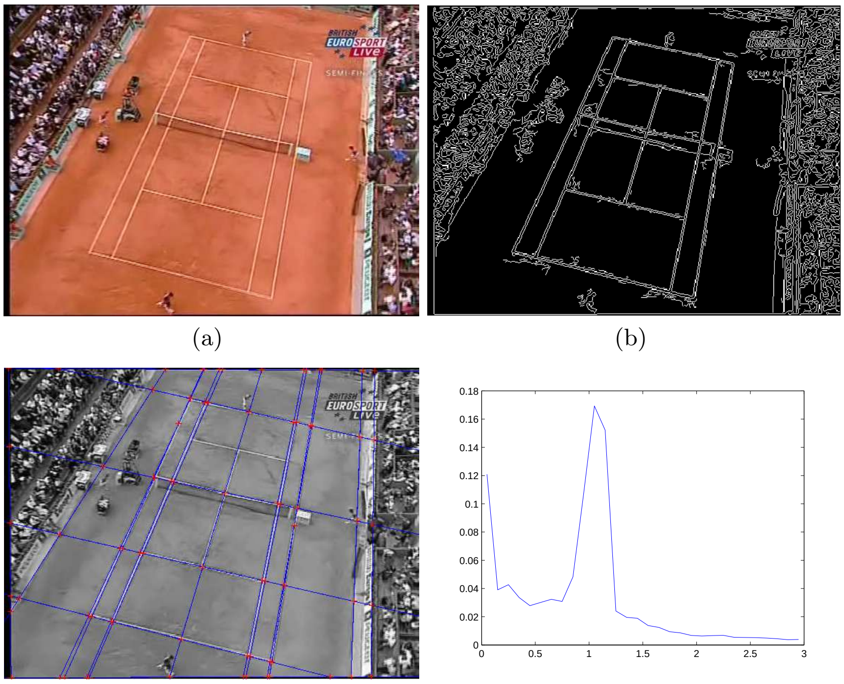
a planar surface. These structures are well defined and uniform across different fields for the same sport. Moreover, these structures consist of lines that are in stark contrast, like white over green ground, to the sports fields making it easy to identify such lines using conventional image processing techniques. These observations necessitated the use of geometric information to develop a robust classifier.

The idea of using geometric information like projective invariance has been used for object recognition [6–8]. But to the best of our knowledge, there has been no prior work done related to the classification of sport's field images using projective invariance. In the proposed approach, we exploit the idea of invariance of cross ratio under projective transformation. Thus in any view of a sport's field four corresponding co-linear points have the same cross ratio. Since this is true only for the four corresponding points, we use a histogram of cross-ratio for a given sport's field image to describe a model for that sport. To ensure selection of the same points in each frame, we limit our consideration to points of intersection of dominant lines only.

The paper is organized as follows. Section 2 explains the proposed approach. The results are propounded in section 3. Finally, the section 4 concludes the paper by suggesting the future work.

2 Proposed approach

The planar geometric structures in images of a sport's field, i.e. lines to demarcate a sport's field, taken from different view points, are related with each other under a projective transformation. Such a transformation does not preserve distances or angles, however incidence and cross-ratios are invariant under it, see



(a) (b) (c) (d)

Fig. 2. Interim steps for calculation of the feature vector: (a) Input image (b) Canny edge detections (c) Hough transform based line detections and the points of interesection of the lines (d) Histogram of cross ratios calculated using the intersection points.

[9]. In the following paragraphs, we build upon the invariant property of both these attributes to develop a robust classifier for sports images.

2.1 Cross ratio histogram

For a line, incidence relations are : 'lies on' between points and lines (as in 'point P lies on line L'), and 'intersects' (as in 'line L1 intersects line L2'). The cross ratio is a ratio of ratios of distances. Given four collinear points $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{p}_3$, and $\mathbf{p}_4$ in $\mathcal{P}^2$ and denoting the Euclidean distance between two points $\mathbf{p}_i$ and $\mathbf{p}_j$ as Δ_{ij}, cross ratio can be defined as shown in figure 1,

$$\tau_{\mathbf{p_1p_2p_3p_4}} = \frac{\Delta_{13}\Delta_{24}}{\Delta_{14}\Delta_{23}}. \qquad (1)$$

Although cross ratio is invariant once the order of the points has been chosen, its value is different depending on this order. Four points can be chosen in $4! = 24$ different orders, but in fact only six distinct values are produced, which are related by the set

$$\{\tau, \frac{1}{\tau}, 1-\tau, \frac{1}{1-\tau}, \frac{\tau-1}{\tau}, \frac{\tau}{\tau-1}\}. \tag{2}$$

Thus, for given four points, we use minimum of the above cross ratio value as a representative for them.

To classify a given unknown sport's field image, a canny edge detection [10] is done on the image. The dominant lines in the image are then identified using a Hough Transform [11]. Since in a sport's field lines drawn on the field are dominant as compared to other edges produced due to noise or player, these dominant lines are are easily identified using the above approach. On each line obtained, we find intersecting points with the rest of the lines. These intersection points are then used for a cross ratio calculation as they will be consistently detected in any view of a sport's field. Also for all the coplanar lines, such points of intersection will correspond to the same physical point on the field irrespective of the viewing angle of the camera due to the invariance of incidence relations. From the entire set of points obtained in a line, we form subsets of four points each. For each subset we calculate the representative cross-ratio as explained using the set in equation (2).

Thus for a given image, we obtain a large number of these cross ratio values which encode the geometric structures present in the field. Ideally, if we are able to consistently reproduce the corresponding points in each input image for a given sport, we can match the cross ratio values individually and obtain a good classification. But this is not the case normally, there are two issues which make the direct comparison of cross ratio values infeasible. Firstly, it is challenging to consistently establish correspondence in points across images of a sport's field from different view points because some points might not be detected due to noise in the image and as well due to the orientation of the sports fields being significantly different. Secondly, there might be small variation in the cross ratio values due to noise in measurement of lines. To overcome these issues, a histogram based representation of cross ratios is done.

For each new image to be classified, the above steps are done to obtain a histogram of cross ratios. This histogram is then used as a feature vector describing the dominant geometric structures in the image. A trained SVM classifier is used to decide the class to which each feature vector and its corresponding image belong to. The entire process with intermediate results is pictorially depicted in figure 2.

2.2 Support Vector Machine classifier

Since multiple sports are being considered, there is a need for a multi-class SVM classifier. SVM is originally designed for binary classifications only. However, multiple strategies of extending the binary classifier to a multi-class classifier have been discussed in literature such as One-against-all [12], one-against-one [13], half-against-half [14] etc. In case of one-against-all(OVA), the N class

	Horizontal line	Vertical line
Tennis	1.0208	1.0989
Badminton	1.08	1.412
Football	1.333	1.05
Basketball	0.91	0.96

Table 1. Cross ratio values for horizontal and a vertical line, on the outer boundary of the sport field, in each sports. These values are calculated based on the standard field measurements directly.

problem is decomposed into series of N binary class problems. In one-aginast-one(OVO), for a N class problem $\frac{N(N-1)}{2}$ binary classifiers are trained and the appropriate class is found by a voting scheme. For a Half-against-Half(HAH) extension, a classifier is built by recursively dividing the training dataset of N classes into two subsets.

We have used a modified OVA approach. OVA is one of the earliest approaches for multi-class SVM and gives good results for most of the problems. For an N-class problem, it constructs N binary SVMs, each binary SVM is trained with all the samples belonging to one class as positive samples and all the samples from the rest of the classes as negative. Given a sample to classify, all the N SVMs are evaluated and the label of the class that has the largest value of the decision function is chosen.

One drawback of this method is that when the results from multiple classifiers are combined, each classifier gets the same importance irrespective of its competence. Hence, instead of directly comparing the decisions, we use reliability measures to make the final decision which makes the multi-class framework more robust. The static reliability measure proposed in [15] has been used for our multi-class framework.

3 Experimental Results

Images of Tennis, Basketball, Badminton, and Football fields have been used for experimental evaluation since they possess a good geometric structure in their respective fields. First column of Table 1 shows the cross ratio values considering points of intersection of the vertical lines with the outer boundary horizontal line of the field. Similarly, second column shows the cross ratio value for points of intersection of horizontal lines with the outer vertical line of the field. It can be observed that the difference of geometric structure in a sport's field is reflected well in the cross ratio values and hence these can be used as a basis for classification of the sport's field. Cross ratio histogram for images of two different sports fields for the all the four sports considered is shown in figure 3. It can be noted that for the same sport images taken on different fields from different view points generate similar cross ratio histograms.

We have collected 200 images of each sport's field from videos, Internet; and by taking photos of the various sport's field we ensure that the images for

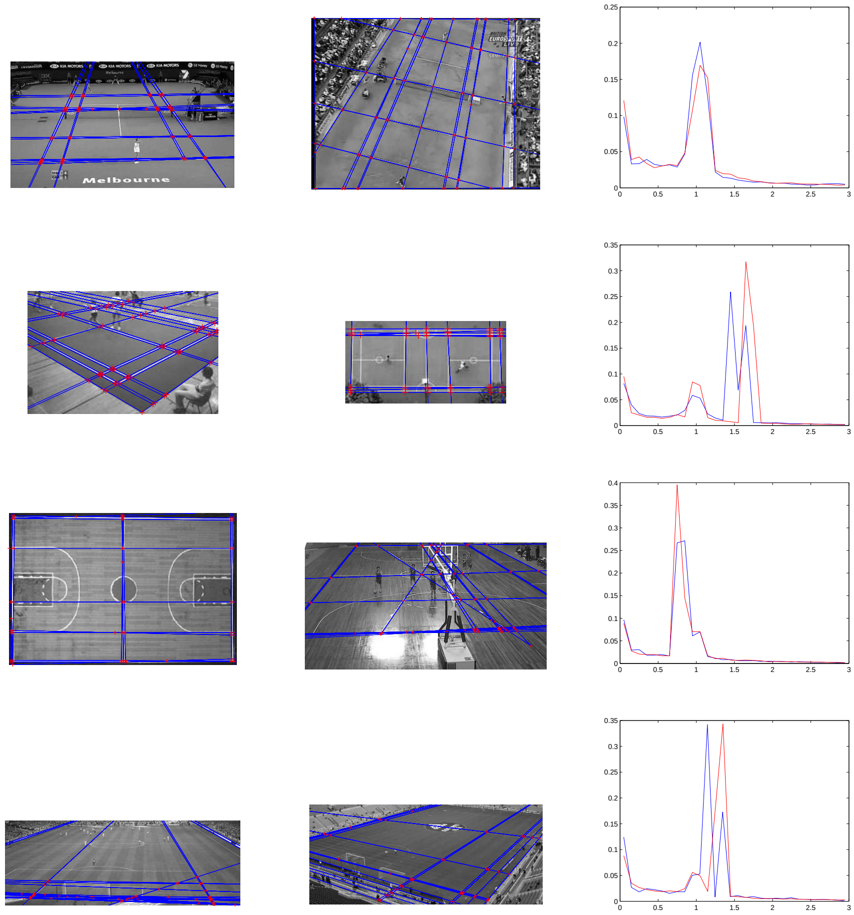


Fig. 3. First and second column show two seperate views of the sport field with detected line and points of intersections overlaid on the gray image. The third column shows the cross ratio histogram for both the views (blue represent the histogram for first column image and red for the second column image in each row).

Sport	Tennis	Badminton	Basketball	Football	Non-sport
Tennis	96.66	2	0.66	0	0.66
Badminton	2.66	94	0.66	1.33	1.33
Basketball	0.66	2	91.33	3.33	2.66
Football	0.66	0.66	3.33	90.66	4.66
Non-sport	2	1	5	8	84

Table 2. Classification accuracies in percentage for each class. The overall classification accuracy is 91.857 %

each class were taken with different viewpoints and illumination conditions. Out of these images, 50 images from each class were used for generating the cross ratio histogram feature vector to train the SVM classifier. The remaining 150 images from each class were mixed to form a collective database of 600 images which has been used as test set to evaluate the performance of the trained SVM classifier. To test robustness of the proposed system when presented with non-sport images, we introduced 100 random images that did not belong to any of the four sports field.

For SVM classifier, we experimented with Polynomial kernel function and Radial basis kernel functions with a variety of parameter values. Among the both, we observed that the second order polynomial function performed with higher accuracy.

The performance metric is defined as the percentage of correct classification per sport. Classification accuracies for the 4 types of sports is given in Table 2. Overall classification error on the test dataset was observed to be 8.14 % . The results clearly indicate that for real world datasets under different conditions and viewpoints, the classification accuracies are very good. Also we noted that most of the non-sport images were also segregated well and the few which were not classified as one of the four sports consisted of images of buildings.

4 Conclusion

The paper proposed an approach for sports image classification based on geometric structure of the sports field. Our experimental validation shows encouraging results and high accuracy of classification even for widely separated views of a sport's field. One of the observation from experimental results was that the proposed approach performs very well when used for classifying images of the sports fields it was trained for. But the performance drops, from the classification error of 6.83% to 8.14%, with introduction of random images having prominent geometric structures.

Thus, currently we are working on to improve the system's ability to discern between sports and non-sports images using other invariant cues from geometry of the sports fields to improve robustness and accuracy of the proposed approach. To address this, we are exploring a relative spatial distribution of intersection points and invariant measures for lines under projective transformation to further augment the feature vector for the classifier. Also, we are gathering more dataset for sports like Hockey, Baseball, and Lacrosse to extend the classifier.

References

1. Wang, D.H., Tian, Q., Gao, S., Sung, W.K.: News sports video shot classification with sports play field and motion features. In: ICIP. (2004) 2247–2250
2. Takagi, S., Hattori, S., Yokoyama, K., Kodate, A., Tominaga, H.: Sports video categorizing method using camera motion parameters. In: ICME '03: Proceedings of the 2003 International Conference on Multimedia and Expo, Washington, DC, USA, IEEE Computer Society (2003) 461–464
3. Kobla, V., DeMenthon, D., Doermann, D.: Identifying sports videos using replay, text, and camera motion features (2000)
4. Messer, K., Christmas, W., Kittler, J.: Automatic sports classification. icpr **02** (2002) 21005
5. Wang, J., Xu, C., Chng, E.: Automatic sports video genre classification using pseudo-2d-hmm. In: ICPR '06: Proceedings of the 18th International Conference on Pattern Recognition, Washington, DC, USA, IEEE Computer Society (2006) 778–781
6. Lei, G.: Recognition of planar objects in 3-d space from single perspectiveviews using cross ratio. IEEE Transactions on Robotics and Automation **6**(4) (1990) 432–437
7. Weiss, I., Ray, M.: Model-based recognition of 3d objects from single images. IEEE Transactions on Pattern Analysis and Machine Intelligence **23**(2) (2001) 116–128
8. Rajashekhar, Chaudhuri, S., Namboodiri, V.: Image retrieval based on projective invariance. (2004) I: 405–408
9. Mundy, J.L., Zisserman, A.: Geometric invariance in computer vision. MIT Press, Cambridge, MA, USA (1992)
10. Canny, J.: A computational approach to edge detection. IEEE Trans. Pattern Anal. Mach. Intell. **8**(6) (1986) 679–698
11. Hough, P.: Method and means for recognizing complex patterns. In: US Patent 3,069,654. (1962)
12. Vapnik, V.: Statistical Learning Theory. (Wiley)
13. KreBel, U.H.G.: Pairwise classification and support vector machines. (1999) 255–268
14. Lei, H., Govindaraju, V.: Half-against-half multi-class support vector machines. In: Multiple Classifier Systems. (2005) 156–164
15. Yi Liu, Z.Y.: One-against-all multi-class svm classification using reliability measures. In: International Joint Conference on Neural Networks. Volume 2. (2005) 849–854