Object Tracking and Classification in Sport Videos

Mustafa Teke
Graduate School of Informatics
mustafa.teke@gmail.com

Masoud Sattari
Department of Computer Engineering
masoud.sattari@gmail.com

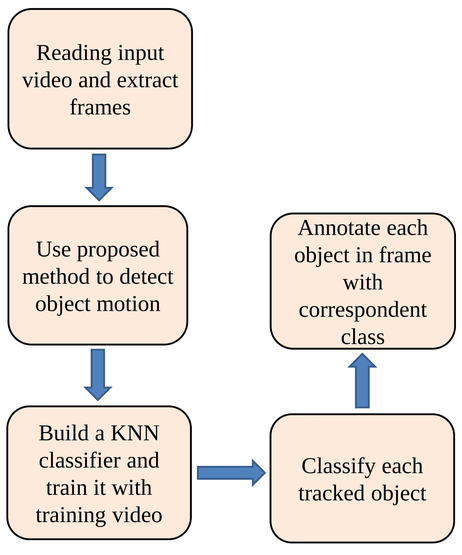
Introduction

- Different methods is used for object tracking:
  - Mean shift
  - Lucas-Kanade Tracker
  - Codebook
  - Motion histogram
  - Kalman filter
- In this project we will track players, referee and ball in a football video. After tracking the moving objects, our task is to classify them to 4 classes:
  - Red players
  - White players
  - Referee
  - Ball
- For tracking Adaptive background mixture models is used.
- KNN classification is applied to classify each object
- In order to evaluate the method we calculate Precision/Recall for both object tracking and classification
- Our results are compared to results of Codebook and Motion histogram

Dataset

- Training and test videos are acquired from VS-PETS 2003
- It contains videos from outdoor football
- There are two fixed camera in the corners of play field
- Videos are 720*576 and 25 frame/second

Work Flow



Algorithm

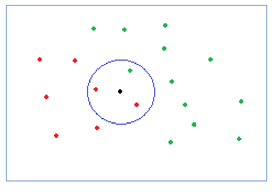
- Images of the scene without the moving objects exhibit some regular behavior that can be well described by a statistical model.
- If we have a statistical model of the scene, an intruding object can be detected by spotting the parts of the image that don't t the model. This process is usually known as background subtraction.
- A common bottom-up approach is applied and the scene model has a probability density function for each pixel separately.
- A pixel from a new image is considered to be a background pixel if its new value is well described by its density function.
- For a static scene the simplest model could be just an image of the scene without the moving objects.
- Gaussian mixture model is used to estimate appropriate values for the variances of the pixel intensity levels from the image.

Classification

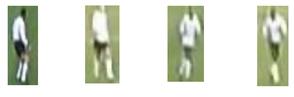
- K-Nearest Neighbors (K-NN) is used to classify each object.
- K-NN aims to collects k (n) nearest neighbor. It starts with window V_n. Window size is increased until it covers k samples.

$$p_n(x) = \frac{k_n/n}{V_n}$$

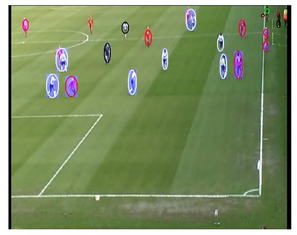
- In this method classification is done based on majority voting.



- K-NN is a very effective against noisy data
- Learning is very simple but classification is costly
- For each region we get RGB color and calculate the Mean value of them also, we calculate Standard Deviation.
- There is a pattern in different regions for these features so; they will be used to classify new regions.
- From a training video we have selected 30 frames and among these frames we have selected 70 regions to train to classifier.



R (mean)	G (mean)	B (mean)	Standard Deviation
154	168	133	17
171	179	144	18
158	171	138	16
173	194	163	15



- Red mask is for Red team
- White mask is for White team
- Black mask is for Referee
- Green is for Ball

Classification Results

TP	FP	FN	Precision	Recall
6	1	1	0.857142857	0.857142857
7	2	1	0.777777778	0.875
8	0	0	1	1
8	1	1	0.888888889	0.888888889
8	3	0	0.727272727	1
7	1	1	0.875	0.875
7	2	1	0.777777778	0.875
6	1	1	0.857142857	0.857142857

Classification Results for Team A (White Team)

TP	FP	FN	Precision	Recall
6	2	3	0.75	0.666667
6	0	1	1	0.857143
6	1	2	0.857143	0.75
6	1	1	0.857143	0.857143
6	3	1	0.666667	0.857143
7	2	0	0.777778	1
7	1	0	0.875	1

Classification Results for Team B (Red Team)

TP	FP	FN	Precision	Recall
1	2	0	0.333333	1
1	1	0	0.5	1
2	2	0	0.5	1
2	1	0	0.666667	1
2	0	0	1	1
1	2	1	0.333333	0.5
2	2	0	0.5	1

Classification Results for Referee

Compare with other methods

TP	FP	FN	precision	recall
16	3	3	0.842105	0.842105
17	4	2	0.809524	0.894737
18	5	3	0.782609	0.857143
18	6	3	0.75	0.857143
19	5	2	0.791667	0.904762
19	3	2	0.863636	0.904762
18	3	3	0.857143	0.857143
18	4	3	0.818182	0.857143
18	4	3	0.818182	0.857143
18	5	3	0.782609	0.857143
17.9	4.2	2.7	0.811566	0.868922

Tracking result for Codebook method

	TP	FP	FN	Precision	Recall
Proposed Method	16.33	2.83	2.72	0.86	0.86
Motion Histogram	14.88	3.71	3.94	0.81	0.79
Codebook	17.90	4.20	2.70	0.81	0.87

Comparing our method with Codebook and Motion histogram

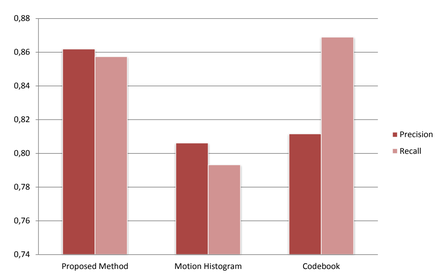


Diagram presentation of table above

Conclusion

Several works has been done in this field and various methods have been proposed. In this work we combine object targeting with classification technique I we hope to get good result. As diagram Shows comparing other two methods our algorithm has a high precision but for recall the proposed method is less than Codebook recall.

References

[1] Yu, X. and Farin, D., “Current and emerging topics in sports video processing,” IEEE International Conference on Multimedia and Expo, 2005. ICME 2005.
[2] Yu, X.; Xu, C.; Tian, Q., “A ball tracking framework for broadcast soccer video”, in Proc. of ICME 2003
[3] Yu, X. and Xu, C. and Leong, H.W. and Tian, Q. and Tang, Q. and Wan, K.W., “Trajectory-based ball detection and tracking with applications to semantic analysis of broadcast soccer video,” in Proceedings of the eleventh ACM international conference on Multimedia, 2003

Source Code is available in http://mustafateke.com/projeler/sports-videos-tracking/ under GPL.