Statistical Pitch Type Recognition in Broadcast Baseball Videos

Chia-Chang Li[1,*], Chia-Wen Lin[2], and Jen-Yu Yu[3]

[1] Industrial Technology Research Institute
Tainan, Taiwan, ROC

[2] Department of Electrical Engineering
National Tsing Hua University
Hsinchu 30013, Taiwan, ROC
cwlin@ee.nthu.edu.tw

[3] Information and Communications Research Labs
Industrial Technology Research Institute
Hsinchu 310, Taiwan, ROC
KavinYu@eitri.org.tw

Received 28 April 2010; Revised 28 May 2010; Accepted 30 June 2010

Abstract. This paper aims to recognize five baseball pitch types from the broadcast baseball game video. Seeing that each pitch type has its own characteristics in speed, trajectory, acceleration and shape, we propose a baseball pitch type recognition scheme that analyzes the spatio-temporal characteristics of ball trajectories. We model the temporal behavior of pitch type using Hidden Markov Models. Our method further extracts the acceleration features of trajectory and then adopts Bayesian decision to classify the features to further enhance the recognition accuracy.

Keywords: Pattern recognition, sports video analysis, video retrieval.

1 Introduction

Fast growing of multimedia technologies has led to many interesting applications such as event and highlight detection in sports programs. Event detection examines the sports video for events like goal-shootings, scoring a point, foul, etc. Highlights in sports videos are the special events that the audience is especially interested in. The analysis of events or highlights can be carried out by combining various attributes of the video, such as color, motion, texture, object, audio, caption etc. The content analysis problem is to identify relations between the features and the semantic concepts that we want to detect, and then find the procedure that captures the relation.

Features used for the existing events detection approaches can be categorized into three general classes: low-level features [1-6], game-specific features [7-10] and object-related features [11-14]. For example, The method proposed in [2] combines several low-level features and characterizes each highlight event by a Hidden Markov Model (HMM), assuming that most highlights in baseball games are composed of certain scene shot types and these shots exhibit transition in time.. Zhang and Chang [10] proposed an approach that makes use of superimposed caption information for the understanding of a baseball game. In [12], a K-Zone system that tracks the trajectory of a pitch and determines whether each pitch is qualified as a strike or a ball.

Besides qualifying a pitch using a K-zone, viewers are also interested in recognizing the pitch type. Such pitch type information is also kind of useful high-level semantics for indexing and retrieving sports videos. For example, a user can search for video shots containing a specific pitcher's famous pitch. Besides, recently the coaches of teams in Major League Baseball (MLB) have been focusing on data sciences that can assistant them in executing tactics. The pitch statistics are useful for baseball teams, players, sports trainers/analysts, and TV reporters.

Pitchers throw a variety of pitches, each of which possessing a slightly different velocity, trajectory, movement, and arm angle. Although there are more pitch types in real situations, we only consider five most common pitch types: fastball, sinker, slider, changeup, and curveball. In our analysis, we focus on the pitches thrown by a right handed pitcher to a right handed batter. The result, however, may be further generalized to other situations.

- Fastball: the pitch thrown with the maximum speed. The type of fastball is intended to have minimum

* Correspondence author

lateral movement.

- Sinker: a pitch in which the ball is spun so that it drops suddenly and move toward the batter as it reaches the plate.
- Slider: a pitch halfway between a curveball and a fastball, with less break but higher speed than the curve. It tends to drop less and away from the batter more than a curve.
- Changeup: a staple off-speed pitch, usually thrown to look like a fastball but arriving much slower to the plate.
- Curveball: A curveball drops sharply as it reaches the plate and is considered an off-speed pitch.

In this paper, we propose a baseball pitch type recognition scheme that analyzes the spatio-temporal characteristics of ball trajectories. At first, all pitches are then coarsely divided into three groups according to their speed. After normalizing the trajectories by using affine transform, our method represents each trajectory as a set of velocity transition vectors in the temporal domain and models each pitch by using HMMs. In order to improve precision, the acceleration features of trajectory are extracted and then classified using Bayesian decision.

2 Propose Statistical Pitch Type Recognition

2.1 System Overview

In order to perform pitch type recognition, ball trajectories need to be extracted. As shown in Fig. 1, we implement a trajectory tracking module by modifying the approaches proposed in [13][15]. In a pitch shot, the ball candidates in each frame are first detected using several filters including color, position, size, and shape filters. Initially, there may exist several ball candidates detected in one frame since the actual ball object may be misdetected because the ball may merge into white regions or be blurred with background. Our method uses a Kalman filter to track the ball positions so as to eliminate unreasonable ball candidates. After the trajectory tracking, the ball trajectories of the candidate set may be discontinuous. We then apply the trajectory interpolation and extension according to the physics-based method of ball trajectory [15] that utilizes the physical characteristics that a ball's motion trajectory is parabolic curve at vertical and is a straight line at horizontal in the time axis. Finally, for a pitch shot, only one trajectory is valid after checking the physical characteristics of detected trajectory candidates.

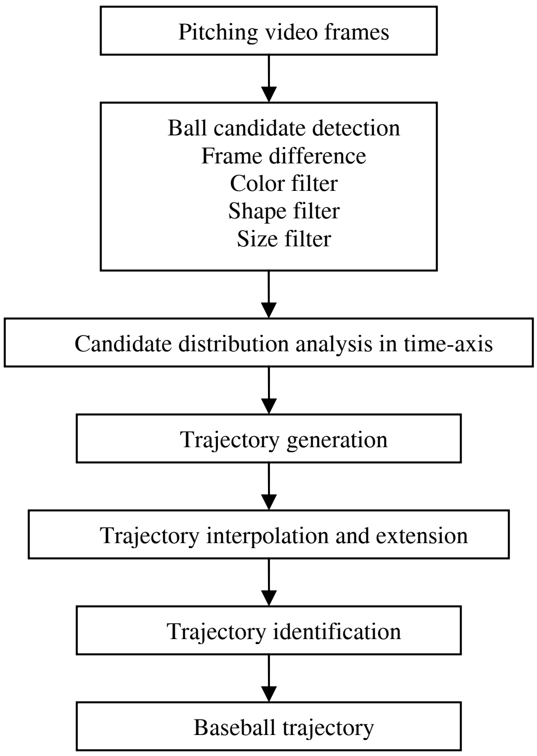


Fig. 1. The flowchart of baseball trajectory detection

Fig. 2 illustrates the processing flow of the proposed algorithm. After extracting the ball trajectory, we analyze the speed, trajectory, and acceleration features of pitches to identify the pitch types. These pitches are first classified into three groups according to the pitch speed, each of the groups including two pitch types. The pitch speed thresholds are empirically set according to speed statistics collected from MLB baseball videos. Subsequently, the trajectory of each pitch is extracted, normalized, and then represented by a set of velocity vectors. These sets of velocity vectors are grouped by *K*-means clustering and the results are then analyzed using HMMs to classify two pitch types in each group.

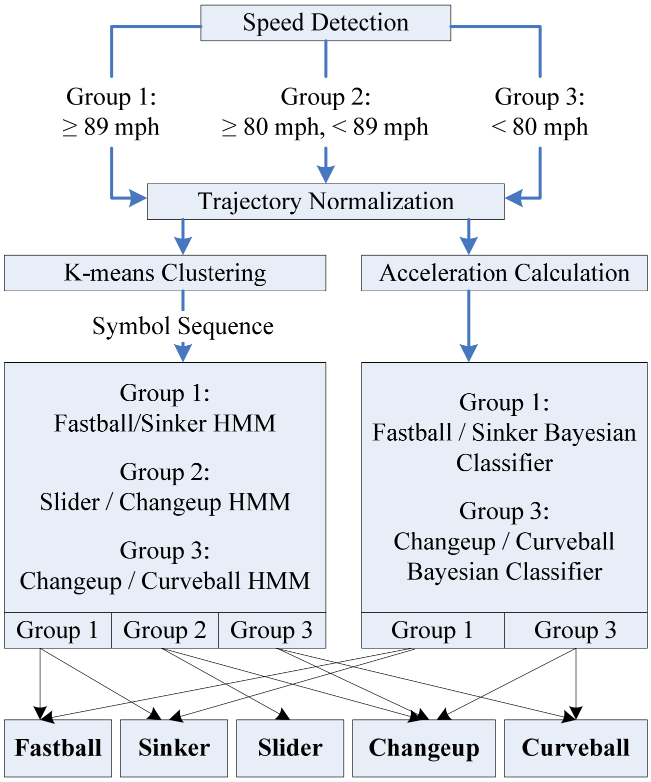


Fig. 2. Processing flow of the proposed algorithm

The acceleration features are then extracted from the velocity vectors and fed into Bayesian classifiers to perform another classification for group 1 and group 3. These results of different classifiers are fused to obtain final classification for all groups.

2.2 Pitch Speed Classifier

Pitch speed may vary from pitch to pitch. Generally, one pitch with a higher speed has less movement (such as fastball), whereas one with off-speed has more movement (such as curve). We collected 10~15 clips for each pitch type and compute the speed distribution for different pitchers. Although the speed of the same type of pitches thrown by different pitchers varies, its range is usually rather stable. The five pitch types are classified into three groups according to the speed distribution as shown in Table 1. These thresholds can be adapted to the pitchers' actual statistics.

Table 1. Pitch speed threshold of each group

Group	Speed Range	Pitch
Group 1	more than 89 mph	Fastball Sinker
Group 2	between 80 and 89 mph	Slider Changeup
Group 3	less than 89 mph	Changeup Curveball

2.3 Trajectory Classifier

After roughly grouping the five types using the speed classifier, we further analyze the pitch types according to ball trajectory along time. The trajectory classifier divides the trajectory into several directions in the time domain and characterizes the temporal behaviors of different pitch types by HMMs.

The position of the *k*-th trajectory in frame *i* can be written as

$$\pi_i^k = \left[x_i^k \; y_i^k \right]^T \tag{1}$$

where x_i^k and y_i^k are the spatial coordinate in the *i*-th frame. Each trajectory can thus be represented as

$$T_k = \begin{bmatrix} x_1^k & x_2^k & \dots & x_N^k \\ y_1^k & y_2^k & \dots & y_N^k \end{bmatrix}^T \tag{2}$$

where *N* indicates the frame index of the *k*-th trajectory.

Because the start position and the end potion may not be aligned from pitch to pitch, the trajectory has to be aligned and normalized. We apply the following affine transformation to normalize each trajectory:

$$\begin{bmatrix} x_i^{k\prime} \\ y_i^{k\prime} \end{bmatrix} = s \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_i^k - x_0^k \\ y_i^k - y_0^k \end{bmatrix} + \begin{bmatrix} -x_0^k \\ -y_0^k \end{bmatrix} \tag{3}$$

where *s* is the scaling factor to normalize the length from the start point to the end point for each trajectory. Fig. 3 shows the original trajectories and normalized trajectories, respectively, for two types.

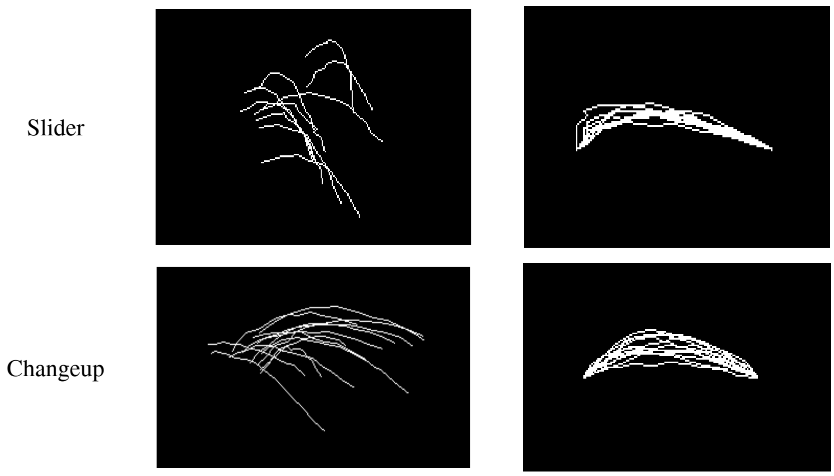


Fig. 3. *K*-means clustering result of trajectory features (*K* = 6)

The normalized trajectory is then represented by sequential velocity vectors in the time domain, in which each velocity $(v_i^k,\ v_i^k)^T$ of the *k*-th trajectory is written as

$$\begin{bmatrix} v_{x,i}^k \\ v_{y,i}^k \end{bmatrix} = \begin{bmatrix} x_{i+1}^{k}{}' - x_i^{k\prime} \\ y_{i+1}^{k}{}' - y_i^{k\prime} \end{bmatrix} \quad \text{for } i = 1, 2, \cdots, N-1 \tag{4}$$

The temporal statistical behavior of each ball trajectory can be characterized by modeling the sequential velocity vectors using HMMs. According to our observations, the directions of the velocity vectors provide useful information in pitch type classification. In our method, the pitch trajectories are first classified using *K*-means clustering. The classification information is subsequently used as the observation symbols in an HMM classifier for pitch type recognition.

A. K-Means Clustering

We first classify the velocity vectors into several classes using *K*-means clustering [16]. In our method, the *K*-means clustering was designed such that the directions of the data in the same class are as close as possible after classification. Besides, the velocity vectors are classified according to its maximum projection to a unit mean vector instead of a nearest mean vector. Fig. 4 illustrates the classification result using *K*-means clustering for 856 velocity vectors when $K = 6$. The classes are subsequently used as the observation symbols of HMM.

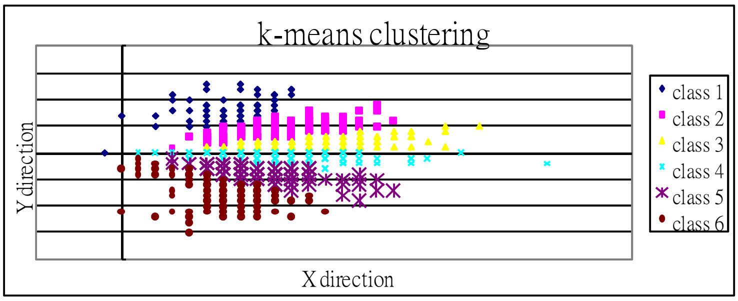


Fig. 4. *K*-means clustering result of trajectory features ($K = 6$)

B. Hidden Markov Models (HMM)

In our method, after applying *K*-means clustering, the temporal statistical behavior of each ball trajectory is modeled using HMMs [16]. An HMM can be described by a transition probability matrix **A**, an initial state probability distribution **π**, and a set of probability density functions for observations **B**. The implementation of HMM includes the following five elements:

- The number of states: This number is usually unknown prior to obtaining HMMs. We select 4 to 12 state numbers in our experiments to find the optimal one.
- Observation symbols: The observation symbols are obtained from the *K*-means clustering result. Because a baseball pitch trajectory consists of about 14 velocity vectors, we use 6 to 8 classes in our experiment based on the assumption that the adjacent velocity vectors in time domain are similar.
- The state transition probability matrix: We use a typical left-right HMM structure. The state transition probability can be learnt from the training data.
- The observation probability distribution in each state: It also can be learnt from the training data.
- The initial state distribution: In the left-to-right HMM we adopted, we set the initial state distribution $\pi = [1, 0, 0, \ldots, 0]$.

2.4 Acceleration Classifier

A breaking ball, such as a slider or a sinker, always changes the direction of movement as it reaches the plate. Each pitch type has an individual way of movement. For example, a slider moves left and down, and a sinker moves right and down in the video frame. There are different characteristics in terms of acceleration from pitch to pitch. Although a pitch changes the direction of movement obviously as it reaches the plate, the acceleration of the rear trajectory is not significantly distinct from others.

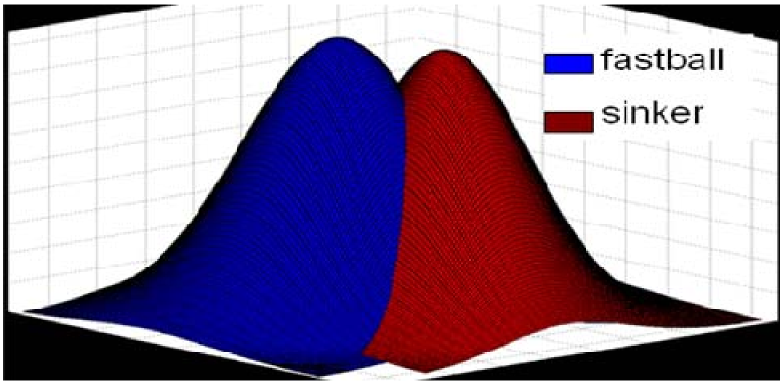
Because the instantaneous acceleration between frames is usually very small (about 1~2 pixels/s^2), the acceleration feature would be sensitive to the noise in ball position detection. We thus propose to use the sum of the whole trajectory to obtain a more reliable feature. Fig. 5 illustrates the acceleration distribution for the three groups listed in Table 1. We can observe that the acceleration distributions of group1 and group 3 can be easily separated, but this is not suitable for group 2. We therefore use the acceleration feature to classify pitch types only for group 1 and group 3.

Bayesian classifier [8] is a statistical method used to solve pattern recognition problems, when those problems are posed in a particular way. To apply the method, given the acceleration measurements, the probability distribution of a pitch needs to be estimated. Assuming that the acceleration distributions of pitches are general multivariate normal, in which the covariance matrices are different for each category, the following discriminant function is adopted.

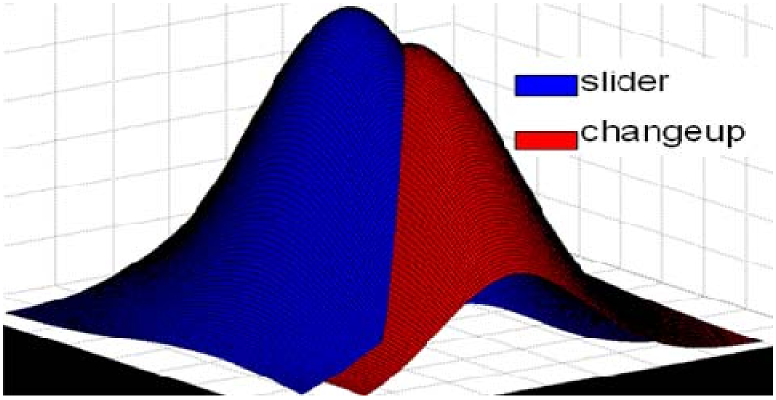
$$g_i(x) = \mathbf{x}^t\mathbf{W}_i\mathbf{x} + \mathbf{w}_i^t\mathbf{x} + w_{i0} \text{ for } i = 1, 2 \tag{5}$$

where $\mathbf{x} = \begin{bmatrix} a_x^k & a_y^k \end{bmatrix}$, $\mathbf{W}_i = -\frac{1}{2}\mathbf{\Sigma}_i^{-1}$, and $\mathbf{w}_i = \mathbf{\Sigma}_i^{-1}\boldsymbol{\mu}_i$ and $w_{i0} = -\frac{1}{2}\boldsymbol{\mu}_i^t\mathbf{\Sigma}_i^{-1}\boldsymbol{\mu}_i - \frac{1}{2}\ln|\mathbf{\Sigma}_i| + \ln P(\omega_i)$

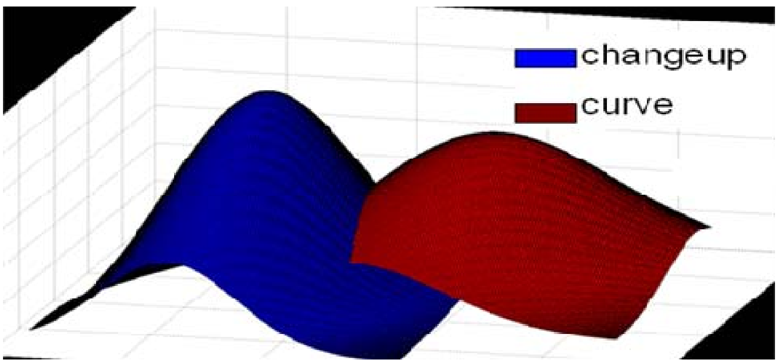
where **x** denotes the measurement of the pitch acceleration, $g_i(\mathbf{x})$ represents the probability of a pitch given the acceleration measurement **x**. The parameters of this discriminant function (μ_i, Σ_i) can be estimated from the training data. Since we initially divide five pitch types into three groups according to the speed measurements and there are two kinds of pitches in each group, the discriminant function is a two-category case. Consequently, we compare $g_1(\mathbf{x})$ with $g_2(\mathbf{x})$ and choose a higher value as the classification result.



(a)



(b)



(c)

Fig. 5. Acceleration distributions of trajectories in three groups: (a) group 1 (fastball and sinker), (2) group 2 (slider and changeup), and (3) group 3 (changeup and curve)

2.5 Fusion of Classifiers

The final probability for each pitch is computed by a weighted sum of two probabilities of HMM and Bayesian as

$$P_{final} = \alpha \cdot P_{HMM} + (1-\alpha) \cdot P_{Bayesian} \quad (6)$$

where P_{HMM} and $P_{Bayesian}$ are the probabilities computed by HMM and Bayesian decision respectively; α is a weighting factor, which is set to be 0.5 empirically.

3 Experimental Results

The database used in our experiments comprises six baseball games recorded from MLB baseball videos with a frame rate of 30 fps. These games involve different teams, pitchers, stadiums, and TV companies. The database has totally 195 pitch trajectories in which 75 trajectories are used as training data and 120 trajectories as test data. The ground truth was established based on human evaluation of experts. Table 2 lists the training data number and test data number for each pitch.

Table 2. Training data number and test data number

Pitch Type	Total	Training sample	Testing sample
Fastball	60	20	40
Sinker	52	20	32
Slider	20	10	10
Changeup	42	15	27
Curveball	21	10	11

We experimented with several observation symbol numbers and states numbers for HMMs as shown in Table 3, where M is the observation symbol number and N is the states number. The results show that HMMs with M = 6 and N = 10 or M = 6 and N = 12 all achieve very promising accuracy, whereas HMMs with M = 7 and N = 12 is particularly accurate in classifying pitch types for group 1.

Table 3. Classification accuracy with different observation symbol numbers (M) and state numbers (N)

(a)

M = 6 , N = 10				
Group	Pitch Type	HMM	Bayesian	Final
Group 1	Fastball	82.50%	82.50%	87.50%
	Sinker	0.8125	62.50%	81.25%
Group 2	Slider	100.0%	(n/a)	100.0%
	Changeup	85.18%	(n/a)	85.18%
Group 3	Changeup	96.29%	70.37%	96.29%
	Curve	81.81%	81.81%	81.81%

(b)

M = 6 , N = 12				
Group	Pitch Type	HMM	Bayesian	Final
Group 1	Fastball	82.50%	82.50%	90.00%
	Sinker	81.25%	62.50%	78.12%
Group 2	Slider	100.0%	(n/a)	100.0%
	Changeup	88.89%	(n/a)	88.89%
Group 3	Changeup	96.29%	70.37%	96.29%
	Curve	90.91%	81.81%	90.91%

(c)

$M = 7$, $N = 12$				
Group	Pitch Type	HMM	Bayesian	Final
Group 1	Fastball	82.50%	82.50%	90.00%
	Sinker	78.12%	62.50%	81.25%
Group 2	Slider	80.00%	(n/a)	80.00%
	Changeup	85.18%	(n/a)	85.18%
Group 3	Changeup	92.59%	70.37%	92.59%
	Curve	81.81%	81.81%	72.73%

(d)

$M = 7$, $N = 12$				
Group	Pitch Type	HMM	Bayesian	Final
Group 1	Fastball	87.50%	82.50%	90.00%
	Sinker	90.63%	62.50%	93.75%
Group 2	Slider	80.00%	(n/a)	80.00%
	Changeup	81.48%	(n/a)	81.48%
Group 3	Changeup	88.89%	70.37%	85.19%
	Curve	90.91%	81.81%	81.82%

4 Conclusion

In this paper, we proposed a novel baseball pitch recognition method. In our method, the trajectory of each pitch is first extracted automatically. All pitches are then coarsely divided into three groups according to their speed. After normalizing the trajectories by affine transform, we represent each trajectory as a set of velocity transition vectors in the temporal domain and model each pitch by using HMMs. In order to improve precision, we have proposed to extract the acceleration of trajectories and classify them by using Bayesian decision. Our experimental results show that the proposed method can recognize the baseball pitches with high accuracy.

References

[1] Y. Rui, A. Gupta, A. Acero, "Automatically Extracting Highlights for TV Baseball Programs," *Proceedings of ACM Multimedia*, pp. 105-115, 2000.

[2] P. Chang, M. Han, Y. Gong, "Extract Highlights from Baseball Game Video with Hidden Markov Models," *Proceedings of IEEE International conference of Image Processing,* Vol. 1, pp. 609-612, 2002.

[3] M. Han, W. Hua, W. Xu, Y. Gong, "An Integrated Baseball Digest System Using Maximum Entropy Method," *Proceedings of ACM Multimedia*, pp. 347-350, 2002.

[4] H. C. Shih and C. L. Huang, "Detection of the Highlights in Baseball Video Program," *Proceedings of IEEE International conference of Multimedia and Expo*, Vol. 1, pp. 595-598, 2004.

[5] C. C. Cheng and C. T. Hsu, "Fusion of Audio and Motion Information on HMM-Based Highlight Extraction for Baseball Games," *IEEE Transactions on Multimedia*, Vol. 8, No. 3, pp. 585-599, 2006.

[6] W. Zhou, A. Vellaikal, C. C. J. Kuo, "Rule-Based Video Classification System for Basketball Video Indexing," *Proceedings of ACM Multimedia*, pp. 213-216, 2000.

[7] C. H. Liang, W. T. Chu, J. H. Kuo, J. L. Wu, W. H. Cheng, "Baseball Event Detection Using Game-Specific Feature Sets and Rules," *Proceedings of IEEE International Symposium on Circuits Systems*, Vol. 4, pp. 3829-3832, 2005.

[8] W. T. Chu and J. L. Wu, "Integration of Rule-Based and Model-Based Decision Methods for Baseball Event Detection," *Proceedings of IEEE International conference of Multimedia and Expo*, 2005.

[9] W. N. Lie and S. H. Shia, "Combining Caption and Visual Features for Semantic Event Classification of Baseball Video," *Proceedings of IEEE International conference of Multimedia and Expo*, pp. 1254-1257, 2005.

[10] D. Zhang and S. F. Chang, "Event Detection in Baseball Video Using Superimposed Caption Recognition," *Proceedings of ACM Multimedia*, pp. 315-318, 2002.

[11] X. Yu, Q. Tian, K. W. Wan, "A Novel Ball Detection Framework for Real Soccer Video," *Proceedings of IEEE International conference of Multimedia and Expo.*, Vol. 2, pp. 265-268, 2003.

[12] A. Gueziec, "Tracking Pitches for Broadcast Television," *Computer*, Vol. 35, No.3, pp. 38-43, 2002.

[13] H. Shum and T. Komura, "A Spatiotemporal Approach to Extract the 3D Trajectory of the Baseball from a Single View Video Sequence," *Proceedings of IEEE International conference of Multimedia and Expo*, Vol. 3, pp. 1583-1586, 2004.

[14] G. S. Pingali, Y. Jean, I. Carlbom, "Real Time Tracking for Enhanced Tennis Broadcasts," *Proceedings of IEEE Computer Vision and Pattern Recognition,* pp. 260-265**,** 1998.

[15] W. T. Chu, C. W. Wang, J. L. Wu, "Extraction of Baseball Trajectory and Physics-Based Validation for Single-View Baseball Video Sequences," *Proceedings of IEEE International conference of Multimedia and Expo*, pp. 1813-1816, 2006.

[16] R. D. Duda, P. E. Hart, D. G. Stork, "Pattern Classification, 2nd ed.," *John Wiley & Sons Inc*, USA, 2001.