Cricket Shot Classification Using Motion Vector

D Karmaker[†], AZM E Chowdhury[†], M S U Miah[*], M A Imran[*] and M H Rahman[*]
Computer Vision Lab
Department of Computer Science
American International University-Bangladesh
{d.karmaker, ehtesham, saef, alimran, hafizur}@aiub.edu

Abstract— **Cricket shots cannot be detected yet from single video sample without multiple view camera and other tools like sonar, speedometer. Extracting salient feature and optical flow from videos of cricket shots is still a challenge. In cricket, body parts movement created several different directional optical flows. So we propose motion estimation approach related to classifying the shots using 3D MACH for action recognition. Our methodology defines 8 classes of angle ranges to detect cricket shots. Our method is grounded on Motion vectors that help to measure the angle of any precise cricket shot. An adequate accuracy level for the shots is established for this particular approach.**

***Keywords-** Cricket; Optical flow; SIFT; Kalman-Filter; Salient; Human pose; Hawk-Eye; Lucas-Kanade; MotionVector; spatiotemporal volume; KLT; LoG; 3D MACH.*

I. INTRODUCTION

Cricket, the game of bat and ball. Different technologies are used in cricket for visualization and coaching [1] [2]. Still satisfactory results for detecting shots are not achieved from recent researches. The problem lies between extractions of features from different video footage from different cameras. Some approaches use field position, camera position, boundary limit, commentary and texts [3]. A machine cannot visualize smoothly like human. Several issues involved such as detection of articulation of human body parts, cricket bat, ball swing and flow. A machine can detect the human body parts from the image [4]. Then, articulating body parts will be recognized [5] using different approaches of action model representation. Moving or articulated body parts tracking and estimation of human pose are much difficult [6] [7] to detect from a single video. Semantic analysis for cricket broadcasting video needs large data set and accuracy level is not satisfactory [8]. Different template based action classification [9] is used to distinguish between different sport activities. It is hard to classify cricket shot as occultation of bat as well as the different parts of body. We propose a distinct approach to classify cricket shots using the motion vector detection [10] for each of the frame using optical flow and transformation of vectors to angle. Here we have implemented for only 4 shots in 4 range of angles of 360 degree. But using this approach we can classify 8 distinct cricket shots. We have established the accuracy level for this method.

II. RELATED WORKS

First, Many researchers have researched on cricket. The Hawk-Eye [2] is an advanced coaching system for cricket. Rahish Tandon and Dr. Amitabha have proposed semantic analysis of broadcasting video [8] involve the use of auxiliary cues to detect events. Another shot boundary detection and shot classification based on multi-scale spatio temporal analysis of color and optical flow features [3].

David Lowe's SIFT [11] and Timor Kadir's research for feature descriptor [12] [13] helps to identify objects direction and with integration with optical flow that can be helpful to detect the bat, ball and body parts movement direction. Bangpeng Yao and Li Fei-Fei have researched about modeling mutual context of object and human pose in human-object interaction activities [6]. Machine also have to recognize the human pose by effective learning approach [14]. Ashwani Aggarwal and Susmit Biswas have proposed a technique for object detection and motion estimation from a MPEG video using background subtraction [10]. Different filter based tracking like Kalman [15], KLT [16] [17] etc approaches are proposed with several modification. Those have played an important role on field of object detection, tracking as well as in action recognition. Mubarak shah, Javed Ahmed and Mikel d. Rodriguiz have proposed an action MACH filter [9] for action recognition, which can distinguish between different sports activities. But these approach take a large learning set to make template class of the action corresponds to the specific sports activity.

Analyzing optical flow, various vision based object tracking [19] and action recognition approach using 3D MACH we proposed our technique to classify cricket shots for detecting several range of angle of optical flow vectors of each pixel of a video [18].

III. PROPOSED METHODOLOGY

A. Detection of Cricket Shot:

A spatio-temporal maximum average correlation height filter (MACH) [9] is used to recognize action. The filter can capture intra-class variability synthesizing a single action MACH for a class. Cricket shots also related to human pose and action. Here, traditional MACH filter is generalized to video (3D spatiotemporal volume), and vector valued data. We have made the MACH of video of specific shots. This can be considered as training set. We have made the training set from 6 videos for each shots. MACH actually make the cluster of the videos. Now a tested video of shot is correlated with the MACH, it will get the specific frames that matched with shot. In use 6 videos to learn every shots and we check it with another video of that correspondent shot. If the match is occurred it will make another video of the specific frames that

ISBN: 978-1-4799-6211-2 ©2015 IEEE

get match with the test video. When the match of the test action is matched the action is a cricket shot that is detected. The MACH 3D is actually developed by Rodriguez [9] for detect specific action. Using this approach we detect the action as cricket shot.

In our system have to recognize a cricket shot. For that we have made the volume. Figure-1 shows datasets for cover drive. Each action or cricket shot has been performed by 6 different actors which is used to make videos for 3D MACH volume.

Fig 1: Dataset for Cover Drive

The input video is correlated with the volume. In final result of detected video we can see that only 7 frames that are matched for the actual cricket action. But testing video has 31 frames. We only select those frames that have matched that particular action and make another video that depict the matched frames only. Figure-2 demonstrates the output event.

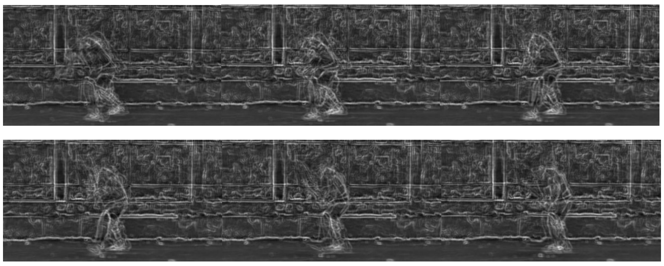

Fig 2: 3D MACH volume for Cover Drive

After recognition of an event i.e. any particular cricket shot, our approach is to detect shot using several range of angle of optical flow vectors of each pixel.

B. Estimation Motion Vector using Optical Flow

We have evaluated the optical flow for each of the frame. There are several approach for evaluating optical flow using a modified approach of Lucas Kanade [16]. Then we got the motion vector for each of the pixel of each of the frame.

C. Calculation of the Angle of the Motion Vector

We have calculated the angle of the each motion vector. We have used here the basic tan inverse method for complex values.

$$\text{vector} = a + b_i$$
$$\text{angle} = \tan^{-1}(b/a)$$

where a = real part of complex value and b = imaginary part of the complex value.

D. Classes of Angle Ranges

All the angles for each vector of each pixel of each frame are clustered at 8 angle range class. The classes are [-90 to -134], [-135 to -179], [-44 to 0], [-45 to -89], [1 to 45], [46 to 90], [91 to 135], [136 to 180]. We have put all the angles of the motion vectors on those correspondent angle class.

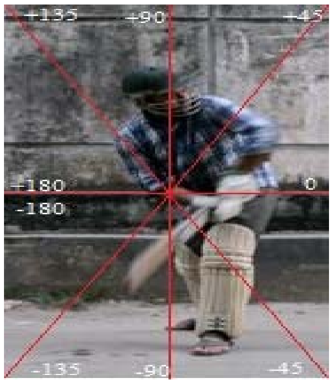

Fig 3: Eight classes of angle ranges

E. Classifying the Shots

1st step is to computer the Laplacian of Gaussian (LoG) of each frame.by using the formula

$$\Delta^2 = \Delta^2 (g * I) = (\Delta^2 g) * I$$

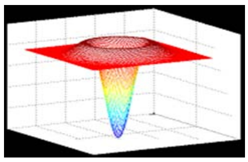

Fig 4: Laplacian of Gaussian (LoG)

Laplacian of Gaussian (LoG) is preferred in this scenario as the numbers of motion vectors are huge in quantity. Based on brightness constancy equation displacement of each pixel in terms of x, y and time is calculated which is represented by u, v and t. where u represents the displacement of X-axis, v represents displacement in Y-axis. Horn & Schunck used smoothness constraint equation in order to find out u,v . As $f_x u + f_y v + f_t = 0$ is under constraint. Along with finding u and v by using Horn & Schunck, we also computed weighted normal flow d by the following equation. $d = \sum W_t f_t / (\text{sqrt}(f_x^2 + f_y^2))$ where W is a weighted value found from parallel from. After

ISBN: 978-1-4799-6211-2 ©2015 IEEE

finding the normal flow for each pixel in the entire frame, all of them are clustered into one group.

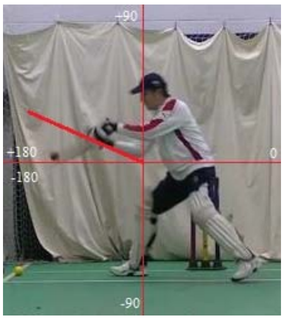

Fig 5: Square-cut shot angle

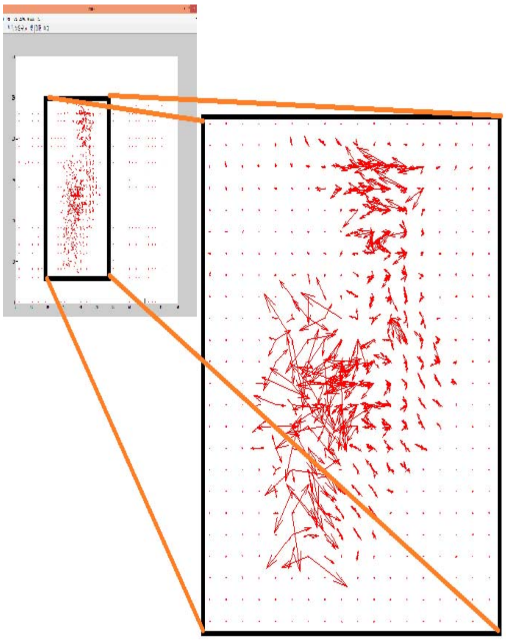

Fig 6: Weighted optical flow

After retrieving all the motion vectors from each frame, we did split in into 11 fragments. Each is represented as a histogram that contains only those motion vectors that fall under that category. Each histogram is computed and the average of all the motion vectors are calculated

Fig 7: Histogram Fragmentation

All For each of the angle class we have sum all the vectors of correspondent angle class. We have taken the class that have maximum (positive or negative) total vector. And this indexed angle class defines the specific shot.

S = {s1, s2, s3, s4, s5, s6, s7, s8}; each element represent range of angles single shot.

V = {v1, v2, v3, v4, v5, v6, v7, v8}; each element represent summation of the motion vectors related to the angle of ranges.

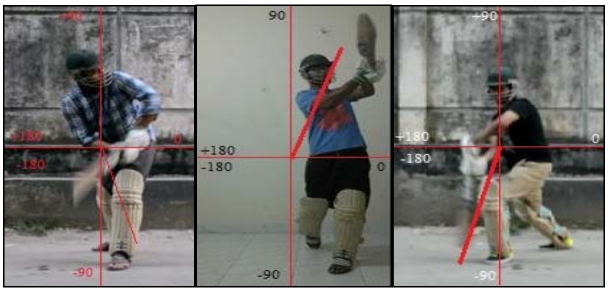

Fig 8: (a) Flick shot angle. (b) Hook shot angle. (c) Off-Drive shot angle

IV. ALGORITHM OF THE FIRST APPROACH

The pseudo-code of the first approach: Classes of angle range

```
vecset = 1 * 8 complexarray
WHILE(NumberOfFrames)
    frameno = no. of frames
    lap = laplacian(frame)
    FOR-EACH of the pixel
          vector = motionvector()
          angle = tan^-1(imag(vector)/real(vector))
          IF angle: 0 to 45
              Vecset[1] = vecset[1] + vector
          END iF
          IF angle: 46 to 90
              Vecset[2] = vecset[2] + vector
          END iF
          IF angle: 91 to 135
              Vecset[3] = vecset[3] + vector
          END iF
          IF angle: 136 to 180
              Vecset[4] = vecset[4] + vector
          END iF
          IF angle: -179 to -135
              Vecset[5] = vecset[5] + vector
          END iF
          IF angle: -134 to -90
              Vecset[6] = vecset[6] + vector
          END iF
          IF angle: -89 to -45
              Vecset[7] = vecset[7] + vector
          END iF
          IF angle: -44 to 0
              Vecset[8] = vecset[8] + vector
        END iF
        END FOR-EACH
END WHILE
Y: Max (positive or negative) vector from vecset array
Compute weighted normal flow
Correspondent angle range of Y defines the shot.
```

ISBN: 978-1-4799-6211-2 ©2015 IEEE

V. Experiment Results

We have experimented on this approach. For the approach we have figured out summation of vectors corresponds to 4 shots. We can detect 8 shots right now with the approach. We have tested on 4 shots and the results of 3 shots are visualized on figure 8. Through the approach we have shown the accuracy level of our tested 4 shots on Table III. We have also shown the angle ranges and the vector summation for these shots on Table I. The threshold levels are shown on Table II.

We have made the accuracy level of our classifier based on the shots of specialized cricket coaches and player and set of shots are used for testing.

TABLE I. Vector Sum and Angle Range of Four Shot Class

Shot	Square Cut	Hook	Flick	Off Drive
Angle Range (in degree)	+136 to +180	+46 to +90	-45 to -89	-90 to -134
Vector Summation	-4.96e+03 +1.73e+03i	+8.96e + 02 +1.9e + 02i	+6.81e+02 -2.2e+03i	-6.19e+02 +2.37e+03i

TABLE II. Threshold Level for the 4 Shots.

Shot	Square Cut	Hook	Flick	Off Drive
Threshold (Max. Correlation)	0.73239	0.65141	0.52361	0.78307

TABLE III. Accuracy Level of Our Approach

Shot	Square Cut	Hook	Flick	Off Drive
Accuracy	61.22%	53.32%	62.74%	63.57%

VI. Future Works

In our approach, we can modify the process by establishing an effective bat tracker. By doing this we can eliminate the motion vector of other than cricket bat. So from this we will increase the angle accuracy as well as the detection accuracy. In the second approach, we can lessen the processing time as well as effective learning like SVM (support vector machine) and Bag of Words. We will in merge with action recognition, we can do skeleton based action detection [20] as well as with the bat angle. So this approach will more be convenient.

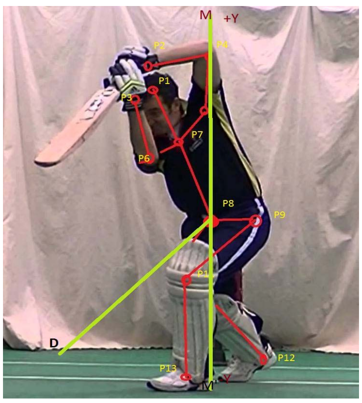

Fig 9: 3D model of skeleton

Our ultimate goal is to develop a vision based automated cricket coaching system. As Computer vision is fully depends on the color of pixels. Using single camera it is hard to make best precision of flow of bat and actual 3D model of angle. In future we will make a 3D model of skeleton [20] (see figure-9) occultation and with the movement of different body parts we will define the action along with angle of flow of bat. Our future approach will expanded to bowler action as well as bowling classification using the same method applied here. Making an intelligent cricket coaching system for all aspects of people as well as for cricket boards and academies is big challenge. We are trying to make as much efficient and accurate as a human eye can do.

VII. Conclusions

In this paper, proposed approach for detection of the specific cricket shots included several vision based tracking approaches and their extensive implementation in a consistent manner.

In our approach we have utilized the MACH filter [9] and correlated the test video with MACH train set. After getting the maximum correlation threshold we define different shots. After analysis of optical flow we have got the summation of motion vector results corresponds to the angle ranges for each of the angle classes. After getting the maximum correlation threshold we define different shots. We have stated an accuracy level of our proposed method according to real life cricket shots. To the best of our knowledge, these approaches of detection and classifying cricket shots have not been considered for studies earlier. According to the several vision based approaches and their application, our approach can be effective and efficient.

References

[1] H. Collins and R. Evans, “Hawkeye second edition public understanding revised 07 - clean,” 2008.

[2] H. Collins and R. Evans, “hawkeye-final-submission,” 2012.

[3] D. S. Rughwani, “Shot Classification and Semantic Query Processing on Broadcast Cricket Videos Shot Classification and Semantic Query Processing on Broadcast Cricket Videos,” no. September, 2008.

[4] J. Shotton, T. Sharp, and A. Kipman, “Real-time human pose recognition in parts from single depth images,” Communications of the . ,2013. [Online]. Available: http://dl.acm.org/citation.cfm?id=2398381.

[5] M. Andriluka, S. Roth, and B. Schiele, “Pictorial structures revisited: People detection and articulated pose estimation,” Computer Vision and Pattern, pp. 1014–1021, 2009. [Online]. Available http://ieeexplore.ieee.org/xpls/abs all.jsp?arnumber=5206754

[6] B. Yao and L. Fei-Fei, “Modeling mutual context ofobject and human pose in human-object interaction activities,” . . . and Pattern Recognition (CVPR) 2010 IEEE, vol.1,b 2010. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs all.jsp?arnumber=5540235.

[7] S. H. Zhang Z, “Skeleton body pose tracking from efficient three-dimensional motion estimation and volumetric reconstruction,” International Journal of Computer Science and . . . , 2013. [Online]. Available:http://www.opticsinfobase.org/abstract.cfm?uri=ao-51-23-5686 http://ijcsmc.com/docs/papers/April2013/V2I4201376.pdf

[8] R. Tandon, “Semantic Analysis of a Cricket Broadcast Video,” pp. 1–9, 2009.[Online].Available:http://www.cse.iitk.ac.in/users/rashish/FinalReport.pdf

[9] M. D. Rodriguez, J. Ahmed, and M. Shah, “Action MACH A Spatio-temporal Maximum Average Correlation Height Filter for Action Recognition.”

[10] A. Aggarwal, S. Biswas, and S. Singh, "Motion Vector Based Moving Object Detection and Tracking in the MPEG Compressed Domain," pp. 121–130, 2006.

[11] D. G. Lowe, "Distinctive image features from scaleinvariant keypoints," International Journal of Computer Vision, vol. 60, no. 2, pp. 91–110, 2004.

[12] T. Kadir, P. Hobson, M. Brady, and J. Close, "FROM SALIENT FEATURE TO SCENE DESCRIPTION,"pp. 2–5.

[13] J. Shi, "Good Features to Track 2 Two Models of Image Motion 3 Computing Image Motion," no. June,1994.

[14] C. M. Bishop, Pattern Recognition and Machine Learning, 2006, vol. 4. [Online].Available:http://www.library.wisc.edu/selectedtocs/bg0137.pdf

H. Patel and D. Thakore, "Moving Object Tracking Using Kalman Filter," International Journal of Computer Science, 2013. [Online]. Available:http://www.opticsinfobase.org/abstract.cfm?uri=ao51235686http://ijcsmc.com/docs/papers/April2013/V2I4201376.pdf.

[15] S. Baker and I. Matthews, "Lucas-kanade 20years on: A unifying framework," International journal of computer vision, vol. 56,no. 3, pp. 221–255,2004.[Online]. Available:http://link.springer.com/article/10.1023/B:VISI.0000011205.11775.fd

[16] B. Duvenhage, J. Delport, and J. de Villiers, "Implementation of the Lucas-Kanade image registration algorithm on a GPU for 3Dcomputational platform stabilisation," Proceedings of the 7th, 2010. [Online]. Available:http://dl.acm.org/citation.cfm?id=1811172

[17] B. Horn and B. Schunck, "Determining optical flow," 1981 Technical, vol. 2, no.April,pp. 326–332, 1981. [Online]. Available: http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=1231

[18] M A U Rahman, M S U Miah, M A Fahad and D Karmaker, " SHIMPG: Simple human interaction with machine using Physical Gesture" Control Automation Robotics & Vision (ICARCV), 2014 IEEE. [Online]. Available: http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=7064322.

[19] D karmaker, H Rahman, M S Rahaman, M S Rahman, "Global Motion tracking with six parameter model", ARPN Journal of Systems and Software,Volume 1 No. 5, 2011. [Online]. Available: http://scientific-journals.org/journalofsystemsandsoftware/archive/vol1no5/vol1no5_3.pdf.

[20] Gall, J., Stoll, C., De Aguiar, E., Theobalt, C., Rosenhahn, B., Seidel, H.P.: Motion capture using joint skeleton tracking and surface estimation. 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops,CVPR Workshops 2009 (2009)1746–1753