A ROBUST HOUGH-BASED ALGORITHM FOR PARTIAL ELLIPSE DETECTION IN BROADCAST SOCCER VIDEO

Xinguo Yu a, b, Hon Wai Leong b, Changsheng Xu a, and Qi Tian a

a Institute for Infocomm Research, 21 Heng Mui Keng Terrace, Singapore 119613, {xinguo, xucs, tian}@i2r.a-star.edu.sg
b School of Computing, National University of Singapore, Singapore 117543, leonhw@comp.nus.edu.sg

ABSTRACT
This paper presents a robust Hough-based algorithm for partial slightly-oblique ellipse detection in broadcast soccer video. The successful identification of the ellipses will significantly facilitate soccer video analysis. The existing standard and various modified ellipse Hough transforms measure a cell in the Hough space as though the ellipse defined by the cell were a complete ellipse. Hence, they are not robust when they are applied to detect the partial ellipses appearing in broadcast soccer video. This paper proposes a new measure function that is able to fairly measure whole and partial ellipses. With the improved measure function, we propose an algorithm to detect the partial ellipses in broadcast soccer video. The proposed algorithm first estimates the target ellipse by using the symmetry of the ellipse and the domain knowledge of soccer video. Then for each estimated ellipse, the algorithm searches around the estimated ellipses to find the ellipse with the highest measured value. Our algorithm is efficient and memory-little, i.e. it overcomes two main problems of the standard ellipse Hough transform. More importantly, the proposed algorithm is much more robust than the existing ellipse Hough transforms. Experimental results show that the proposed algorithm achieves above 96% recall and 100% precision. Our algorithm may be the first one that is able to detect ellipses from commercial video in pseudo real-time.

Key Words: Soccer Video, Ellipse Detection, Hough Transform, Measure Function.

1. INTRODUCTION

Ellipse detection in broadcast soccer video has attracted a growing interest as its results can significantly facilitate soccer video analysis [9-11]. With the results of ellipse detection, we can estimate the ball size [10-11], facilitate the video indexing and summarization, compute the view position [9], and proceed to do the camera calibration.

In soccer video analysis, the work in [9] detected the existence of the ellipse to compute which portion of the soccer field is in the frame. However, the work can only detect the ellipse that is almost complete as it uses the least squared fitting on the edge points. The work in [10-11] detected the existence of ellipse and computed the size of the ellipse to estimate the ball size. However, these methods did not address how to robustly detect the partial ellipse in broadcast soccer video. The detection performances have a lot of room for improvement. In ellipse detection, many algorithms have been proposed in last four decades. Most of them belong to three categories: Ellipse Hough Transform (EHT) [2-3, 5-8, 12], Least Squared Fitting (LSF) [4], and Invariant Pattern Filter (IPF) [1]. Unfortunately, all work in these three categories did not solve the problem of how to fairly measure partial ellipses. The EHT is the most common method for ellipse detection. The idea of the EHT is to gather the evidence of the ellipse occurrence in the Hough space. There are two ways to collect the evidence. One way (used by the Standard EHT (SEHT) as well as its variants) is that each sample point from the image votes for all the cells that might have produced the sample point [5]. The other way (used by the combinatorial ellipse Hough transform (CEHT) as well as its variants) is that each 5-point combination votes for one cell determined by this combination [3]. The votes are counted in an accumulator array, called the absolute measure function (AMF). For a given cell, the value of the function is the total times that sample points or the combinations of sample points have voted for it. There are two problems when the EHTs including the SEHT and CEHT are applied to detect the almost-whole ellipses. One is that in its simplest implementation it requires a lot of computation. The other is that it requires a large memory space to maintain its accumulator array. Various modifications to them have been proposed to overcome or alleviate the problems [2, 5-8, 12]. Yuen et al. [12] proposed a two-stage EHT to detect the partially-occluded ellipses. However, their method requires that the arcs of targets are smooth as the method uses the tangents of the arcs to compute the center of target ellipses. To our best knowledge, all modified EHTs have not changed the AMF. However, the AMF has a bias against the partial ellipses as the partial ellipse definitely has fewer evidences than the complete ellipse with the same size. As a result, the existing EHTs are not robust when they are applied to detect the half ellipse appearing in the noisy image due to the use of the AMF. Hence, the existing EHTs cannot robustly detect the partial ellipses in broadcast soccer video.

Filters in the IPF [1] also measure each hypothesized ellipse though it were a complete ellipse being bias against partial ellipses. As for the LSF, it cannot be applied to fit partial ellipses at all [4] as it optimizes the ellipse parameters only close to the obtained sample points.

In this paper, we propose an algorithm to detect the partial ellipses in broadcast soccer video. To build the robust algorithm we propose an ellipse measure function which is able to fairly measure the whole and partial ellipses (or ellipse rings). The proposed algorithm has three main components: ellipse estimation, ellipse search, and ellipse refinement. In the ellipse estimation component, we first detect the central line in soccer field. Once the central line is found, we rotate the frame to make the central line to be vertical. Then we use two methods to estimate parameters of the target ellipse: template matching and point statistics. In the ellipse search component, we search around the estimated ellipses to obtain the ellipse with the highest measure value. Finally, the found ellipse in the transformed frame is converted (mapped back) into the original frame. To get the accurate parameters we refine the ellipse. Our algorithm is very fast as we estimate the possible locations and sizes of the target ellipse efficiently by making use of the symmetry of ellipse and the domain knowledge of soccer video. Unlike the existing EHTs, the proposed algorithm measures cells in the Hough space one by one. To save memory our

0-7803-8603-5/04/$20.00 ©2004 IEEE.
algorithm only memorizes the cell with the highest measure value. Hence, unlike the existing EHTs, our algorithm is memory-little as it does not memorize all values of the measure function. More importantly, the experimental results show that the proposed algorithm is much more robust than the EHTs using the AMF. The proposed algorithm targets to detect the ellipses with more than half of their areas visible in the frame. Detecting the ellipses that has less than the half appearing in the frame is left to the ellipse tracking algorithm but not to be discussed in this paper.

The rest of the paper is structured as follows. Section 2 presents the proposed algorithm. Section 3 shows the experimental results of the ellipse detection algorithm presented in Section 2. We conclude the paper in Section 4.

2. ALGORITHM OF ELLIPSE DETECTION

For each frame of broadcast soccer video, we propose a Hough-based algorithm to detect the ellipse using three components as shown in Figure 1. In the rest of this section, we describe each component in Figure 1 in turn.

2.1. Estimation of the Ellipse

The ellipse estimation component is to estimate the parameters of the target ellipse in the frame. In almost all the frames with more than half ellipses, the central line of the soccer field appears in the frame. Hence we first detect the central line and rotate the frame to make the central line to be vertical. As a result, the ellipse becomes horizontal as it is perpendicular to the central line. Then, we estimate the horizontal ellipse in the rotated frame through sample point statistics and template matching. Notice before the rotation we may shift the frame a little so that the transformed frame can contain most sample points. Hence, generally we do a transform to the frame including shift and rotation.

Detection of the central line: For the given frame, we segment pixels with line color and all other pixels are painted in the field color. For each pixel \((i, j)\), we define \(\Psi(i, j)\) as follows.

\[
\Psi(i, j) = \begin{cases} 1, & \text{pixel } (i, j) \text{ is of the line color,} \\ 0, & \text{otherwise}. \end{cases}
\]  

In the segmented image, many pixels do not belong to the central line. According to statistics, the central line is almost vertical (the angle between the central line and the vertical line is less than \(\pm 5^\circ\)). With this knowledge, we propose a filter \(\mathcal{R}(\bullet)\) to remove the pixels not belonging to long almost vertical line. \(\mathcal{R}(\bullet)\) paints the pixel \((i, j)\) with the field color if \(\Phi(i, j) = 0\).

\[
\Phi(i,j) = \begin{cases} 1, & \text{if } \left| \sum_{k} \Psi(i, j - k) \right| / L > \text{threshold}, \\ 0, & \text{otherwise}. \end{cases}
\]

After we apply \(\mathcal{R}(\bullet)\) to the segmented image, most of pixels that do not belong to the central line are removed, i.e. we visually identify the central line as illustrated in Figure 2. Now we find the parameters of the central line. We compute the histogram that reflects the number of pixels with the line color in each column. From the histogram, we obtain the central line location and its width. This is actually a simplified straight line Hough transform. As a result, our procedure is very fast for finding the central line.

Rotation of the frame: Now we find the slope of the line. We rotate the frame in \(0, \pm 1, \pm 2, \pm 3, \pm 4, \pm 5\) degrees respectively. In each rotated frame, we find the parameters of the central line. We compute the histogram that reflects the number of pixels with the line color in each column. From the histogram, we obtain the central line location and its width. This is actually a simplified straight line Hough transform. As a result, our procedure is very fast for finding the central line.
Estimation of the ellipse in the transformed frame: We have two ways to estimate the parameters of the ellipse. One way is to estimate the parameters by template matching to find the characteristic points on the ellipse as shown in Figure 4(a). In the frame, the central line may cross the ellipse at two points. So we create a template for such cross points and search for such points along the vertical central line. We can compute the length of minor axis, the center, and the major axis if we find such two points. Along the major axis we search for the farthest points of the ellipse by template matching to find the length of the major axis of the ellipse. We may get more than one when we search for the farthest points of the ellipse due to noise. So far we have obtained several estimated cells in the Hough space. The other way is to estimate the parameters of the ellipse by sample point statistics as shown in Figure 4(b). After we obtain the central line, we can roughly segment the ellipse in the transformed frame. In the segmented image, we can compute the location of major axis of the ellipse by finding the average row of qualified sample points. Along the major axis we search for the farthest points of the ellipse by finding all points with line color on the major axis. Through such search we obtain the length of the major axis. Project the sample points horizontally to compute the length of minor axis.

Figure 4. The estimated ellipses. Each pair of black ellipses indicates an estimated ellipse. In (a), two ellipses are estimated by template matching; in (b), two ellipses are estimated by statistics. Each ellipse pair in the figure represents an estimated ellipse.

2.2. Search of the ellipse in the transformed frame

Measure function: To measure each ellipse defined by cells during the following search procedure we need a measure function. The existing AMF measures each cell as though the frame has an ellipse. Hence, the proposed AMF eliminates this bias, we measure the ellipse on the conjunction set of the ellipse and the given frame. Each measured ellipse in the figure represents an estimated ellipse.

A measure function $M(x, y, a, b, \theta)$ is defined to measure for each cell $=(x, y, a, b, \theta)$ how likely the frame has an ellipse defined by $(x, y, a, b, \theta)$. Let $E=E(x, y, a, b, \theta)$ be the ellipse with $(x, y)$ being the center, $a$ and $b$ being the length of major and minor axis and $\theta$ being the degree between the major axis and the $x$ axis. Let $E_p$ be all points on the $E$, including the points in and out of the given frame. We draw two ellipses $E_{in}=E(x, y, a-\delta, b-\delta, \theta)$ and $E_{out}=E(x, y, a+\delta, b+\delta, \theta)$ and $E_p$ be the ellipse ring enclosed by $E_{in}$ and $E_{out}$. Let $(x_0, y_0)$ be any point on $E_p^*=E_p\cap F$. We draw a ray $L$ that starts at $(x, y)$ and passes $(x_0, y_0)$. Let $p(x_0, y_0)$ be the number of the sample points not only on $L$ but also in $E_p$ or $E_{out}$. Then, we define $u(x_0, y_0)$ as follows.

$$u(x_0, y_0) = \begin{cases} 1 & \text{if } n(x_0, y_0) = 0 \& p(x_0, y_0) > 0; \\ 0 & \text{otherwise}. \end{cases}$$

For the ellipse $E=E(x, y, a, b, \theta)$, we define a measure $M(cell) = M(x, y, a, b, \theta)$ as follows.

$$M(cell) = \left( \sum_{(x,y)\in E_p} u(x, y) \right)^2 ||E_p^*||.$$  (4)

where $||E_p^*||$ is the cardinal number of $E_p^*$.

The AMF $\Lambda(cell)$ in the SEHT can be described as follows, where $\Psi(i, j)$ is define in (1).

$$\Lambda(cell) = \sum_{(x,y)\in E_p^*} \Psi(x, y).$$  (5)

To make two measure functions in (4) and (5) comparable we normalize the AMF to become the normalized measure function (NMF) $\overline{\Lambda}(E)$ as follows.

$$\overline{\Lambda}(cell) = \left( \sum_{(x,y)\in E_p^*} \Psi(x, y) \right) \left( ||E_p|| \right)^{-1}.$$  (6)

Searching for the best ellipse: In previous subsection, we have obtained some estimated cells. We search for the best ellipse around each estimated cell in the Hough space. Let $H_4$ be the disjunction of all cell neighborhoods. Then, $H_4 = \{(x, y, a, b, \theta) : |x-x_0| < \delta_x, |y-y_0| < \delta_y, |a-a_0| < \delta_a, |b-b_0| < \delta_b, (x_0, y_0, a_0, b_0, \theta) \text{ is an estimated cell.}\}$. For each cell $\in H_4$, we compute $M(cell)$. During the computation, we keep only the cell with the highest value. The final cell is considered to be the found ellipse in the transformed frame if its measured value is larger than a predefined threshold.

2.3. Refinement of the ellipse found in the original frame

Once we obtain the ellipse in the transformed image, we convert the ellipse in the original image according to the transform we have done into the ellipse in the original image. Let $(x_c, y_c, a_c, b_c, \theta_c)$ be the parameters of the converted ellipse. We define a search space $H_O = \{(x, y, a, b, \theta) : |x-x_0| < \rho_x, |y-y_0| < \rho_y, |a-a_0| < \rho_a, |b-b_0| < \rho_b, |\theta-\theta_0| < \rho_\theta\}$. We search for the ellipse with maximum measure value in $H_O$. We use the same measure defined above, but the search here has several differences from the search in the previous subsection. Firstly, here we need to compute the rotation of the points on the ellipses. Secondly, the search step is smaller than the previous one to find the accurate ellipse. Thirdly, we are searching in the frame with all sample points. During two search procedures, we memorize only the cell with the current highest value using 20 bytes. Unlike the SEHT they do not need an accumulator array. Hence, the proposed algorithm is memory-little contrast, the SEHT needs $4*|H|$ bytes, where $H$ is the Hough space.
3. EXPERIMENTAL RESULTS

The test was conducted on one whole game of FIFA2002 Quarter Final (Senegal vs. Turkey). The frames of the segments are grabbed from MPEG1 video by DirectX 9.0 and the video was recorded by a WinTV card connected to a TV antenna. Our algorithm takes 194 minutes to obtain all ellipses on a PC with 2.4 GHz CPU. In our evaluation, a frame is considered to have an ellipse if the ellipse that is the projection of the center circle of the field has more than its half appearing in the frame. In the algorithm, we use template matching and sample point statistics to estimate the ellipse parameters. Two methods are effective on different frames. Hence, their complementary strengths can be combined to achieve better result. The experimental results of ellipse detection by the proposed algorithm on the whole video are shown in Table 1. In the current implementation, \( \delta \) in \( E_{\text{in}} \) and \( E_{\text{out}} \) in subsection 2.2 is set to 5; \( \rho_x, \rho_y, \rho_z, \rho_b \) are set to 7 with the search step being one pixel; \( \rho_s \) is set to be 20 with the search step being \( \pi / 1000 \).

To show that the proposed measure function is better than the AMF, we construct an algorithm, called the SEHT-Based algorithm or short as the SEHT-based, which is the same with the proposed algorithm except it uses the AMF used by the standard ellipse Hough transform (SEHT) to measure cells in Step 3 and 4. That is, the only difference between them lies in the procedure to search the best ellipse around the estimated ellipses and the converted ellipses. The experimental results by using the proposed algorithm and the SEHT-based are presented in Table 2. Results show that the proposed algorithm can achieve the perfect ellipse detection performance by setting a threshold to the measure function. But the SEHT-based cannot achieve the perfect detection result by defining a threshold to the absolute measure function. To our best knowledge, the existing modified ellipse Hough transforms modified the SEHT only in voting procedure or in reducing the memory requirement but not in improving the measure function. Hence, they should have the same robustness with the SEHT. In conclusion, the proposed algorithm is much more robust than the EHTs using AMF. In other word, the proposed algorithm overcomes the two main problems of the SEHT. The proposed algorithm has several contributions. Firstly it is a very fast algorithm by exploiting the symmetry property of ellipse and the domain knowledge of soccer video. Secondly, it is much more robust than the SEHT and the modified EHTs using AMF as it uses the improved measure function. The SEHT and the modified EHTs using AMF measure each cell as though the ellipse defined by the cell were a complete ellipse so that this measure has a bias against the partial ellipse. We have proposed an improved measure function to fairly measure the whole and partial ellipses, which considers the partiality and the obliqueness of target ellipses. This intelligent measurement function makes our proposed algorithm robust. Thirdly, it is memory-little as it memorizes only one cell when it searches the best ellipse in the Hough space. To save the memory the algorithm measures the cells in the Hough space one by one and it only memorizes the cell with the highest value since each frame of broadcast soccer video has at most one target ellipse. The proposed algorithm has achieved above 96% recall and 100% precision in ellipse detection. On the contrast, the SEHT cannot achieve the high recall when we set a high threshold; or it cannot achieve high precision when we set a low threshold.

The promising results of this paper encourage us to study our method further to apply it to other ellipse detection problems. We also want to explore how the measure function can be used in the general Hough transforms. In addition, we will use the results of ellipse detection to reconstruct 3D images of broadcast soccer video and to estimate the ball size in each frame.

5. REFERENCES


4. CONCLUSIONS AND FUTURE WORK

We have presented a Hough-based algorithm for ellipse detection in broadcast soccer video. This algorithm is memory-little and fast. More importantly, it is more robust than the EHTs using AMF. In other word, the proposed algorithm overcomes the two main problems of the SEHT. The proposed algorithm has several contributions. Firstly it is a very fast algorithm by exploiting the symmetry property of ellipse and the domain knowledge of soccer video. Secondly, it is much more robust than the SEHT and the modified EHTs using AMF as it uses the improved measure function. The SEHT and the modified EHTs using AMF measure each cell as though the ellipse defined by the cell were a complete ellipse so that this measure has a bias against the partial ellipse. We have proposed an improved measure function to fairly measure the whole and partial ellipses, which considers the partiality and the obliqueness of target ellipses. This intelligent measurement function makes our proposed algorithm robust. Thirdly, it is memory-little as it memorizes only one cell when it searches the best ellipse in the Hough space. To save the memory the algorithm measures the cells in the Hough space one by one and it only memorizes the cell with the highest value since each frame of broadcast soccer video has at most one target ellipse. The proposed algorithm has achieved above 96% recall and 100% precision in ellipse detection. On the contrast, the SEHT cannot achieve the high recall when we set a high threshold; or it cannot achieve high precision when we set a low threshold.

The promising results of this paper encourage us to study our method further to apply it to other ellipse detection problems. We also want to explore how the measure function can be used in the general Hough transforms. In addition, we will use the results of ellipse detection to reconstruct 3D images of broadcast soccer video and to estimate the ball size in each frame.