Development of ball trajectory visualization system for live baseball broadcasts and application to other sports

We have developed a system called B-Motion which is able to display baseball pitch trajectories during live baseball broadcasts. The system is able to follow the ball accurately using object-extraction processing based on motion and image characteristics of the ball, and to draw the trajectory in semi-real-time. It also has a function to interpolate ball-location, so it can still display smooth, continuous trajectories, even when there are sections of the trajectory that are obstructed, as with a left-handed batter. The system requires only video from a regular broadcast camera, so no additional camera equipment is required, and it needs no preliminary calibration, so it is very easy to use. The system has been used constantly through six seasons of NHK professional baseball broadcasts. Its functionality has also been extended to track bowling balls in live bowling coverage, and to display tee-shot and putt trajectories for golf coverage.

1. Introduction

It has become common during live sports broadcasts, to use techniques such as computer-graphic overlays on the camera image to show the state of the game and explain rules and tactics in an intuitive, visual way. In particular, there have been significant improvements in usability of virtual technologies allowing computer graphic overlays linked with camera work, allowing for much more intuitive production of sports events, using virtual display to show aspects such as the K-point in ski jumping or the qualifying record line for swimming. However, most virtual displays used in live broadcasts are limited by what information can be entered manually at the broadcast site. Most aspects that change dynamically as the competition progresses are produced manually in post production, and have been difficult to use in live broadcasts.

Visualization of the ball trajectory in ball sports is often useful in understanding strategy or state of the game. For baseball, other equipment has already been developed, including a device that uses inter-frame differences to display a ball after image. However, this technique shows afterimages of everything that moves within the scope of processing, resulting in problems. Afterimages of moving objects other than the ball, like the arms of the pitcher or batter, are displayed, and the position of the shot can cause difficulties, such as when the ball overlaps with the batter for left-handed batters.

Object extraction technology can be useful in resolving these types of problems, by extracting the ball region from images. Although there has been much research in object extraction technology in areas of robotic vision and object recognition, it has been difficult to create a generally applicable theory or apparatus. Techniques generally also require a huge amount of computation, making them even more difficult to apply to live broadcasting.

In light of the above, we have devised a method for extracting and tracking the ball in live video of baseball, and have developed a system called B-Motion, which is able to display the trajectory in real time. The system automatically tracks the ball position within the broadcast camera image and draws the trajectory based on the extracted positions. The trajectory can be drawn in semi-real time, so it can be used for replays immediately after a pitch. The system has been used through five seasons of live NHK professional baseball broadcasts due to its excellent versatility and ease-of-use, including that it operates with conventional camera equipment and is effective when applied for left-handed batters as well.

We have also developed extensions to this system for visualization of ball trajectories in bowling and for tee shots and putts in golf. In this article, we describe our method for high-speed, accurate extraction and tracking of ball position, and introduce some examples of application in live broadcasts of other sports.

2. System organization

A block diagram of the system is shown in Figure 1. For input, the system uses only the signal from the center camera of a baseball broadcast, and no new equipment, such as additional fixed cameras or range finding devices, is required. Also, both the detection and drawing processes use the two-dimensional coordinates of the center-camera image, which is fixed during the pitch
shot, for ball-location coordinates. No calibration for coordinate transformations is required, and no operational changes to conventional broadcast procedures need to be made.

We use regular PCs for ball detection and trajectory drawing processes, but to improve performance, we distribute processing over two PCs. The extraction and tracking PC detects and tracks the position of the ball in the input video, outputting a ball position for each frame. The interpolation and drawing PC takes the ball positions, applies interpolation to them, and creates a computer graphic of the ball trajectory. The trajectory graphic is converted to high-definition (HD) broadcast format by a scan converter\(^1\), and composited with the camera image using an inserter. In order to match the playback timing of the camera and computer graphic images, the camera image is delayed using a frame delay for approximately 15 frames, producing a semi-real-time

\( ^1\) A device which changes video signals from devices like PCs and TVs that have different sync signals to have the same sync signal.
trajectory composite image. Figure 2 shows an external view of the system.

3. Ball extraction and tracking

3.1 Ball extraction

We now describe processing by the extraction and tracking PC. The video input to the system is HD resolution, so processing the entire image would be computationally intensive. Accordingly, the ball-detection search area is limited to a rectangular area as shown in Figure 3. Once the ball has been detected, the size and position of the search region is updated successively according to the motion of the ball in subsequent frames, tracking the ball automatically as in Figure 4 until it arrives in the catcher's mitt.

With a fixed camera, it would be possible to set the position of the search area for each pitcher ahead of time, but broadcast cameras are not fixed and the shot direction changes slightly for each pitch. Thus, in actual operation, the operator must correct the initial position manually before each pitch.

Figure 5 shows the process for extraction and tracking. First, the search area is isolated from the camera video, and fast-moving objects are detected within it, as candidates for the ball area. Camera operations such as pan and zoom are not done during the pitch, so high-speed objects can be detected using inter-frame differences. Per-pixel differences are computed for three successive frames ($t-2$, $t-1$, $t$), creating two difference frames, (frame $t-1$) - (frame $t-2$) and (frame $t$) - (frame $t-1$), as shown in Figure 6. If there are fast-moving objects in the frames, areas of positive-negative reflection will appear in the two difference images. We select these regions as the candidate ball regions for frame $t-1$.

Nearer to home plate, there are more moving objects, such as the bat and the catcher's arms, so in some cases, more candidate ball regions are detected. In these cases, an image-feature filter, as shown in Figure 5, selects a ball region from among candidate regions based on certain image features, which include area, circularity,
luminance, chrominance, and distance from the estimated trajectory. Each feature value is calculated as described below.

- **Area**
  Total number of pixels in the object. The average ball area in the center camera image for professional baseball broadcasts is about 200 pixels.

- **Circularity**
  Object circularity is expressed by equation (1).

\[
e = \frac{4\pi S}{l^2}
\]

Here, \( S \) is the object area, \( l \) is the perimeter, and \( \pi \) is the circle ratio. A perfect circle has circularity of 1.0, but due to its motion, the ball image is not a perfect circle, so we use a target value of about 0.8.

- **Luminance, chrominance**
  The brightness and average of \( C_b, C_r \) * values (in the range 0 to 255). The ball is white, so we expect a luminance value in the range 160 to 220. Also, white is a neutral color, so we expect the average of \( C_b \) and \( C_r \) to be about 128. Note that if there is an object of equivalent luminance or chrominance in the background (left-handed batter, advertisement, etc.), the ball can be difficult to distinguish.

- **Distance from estimated trajectory**
  The distance of the centroid of the candidate region to the estimated trajectory computed by the position estimation process. The estimated trajectory is discussed in section 3.2.

The above feature values are computed for each region and the region with feature values closest to the target values is selected as the ball. The centroid of the selected region is computed and used as the ball position. Note that the values obtained for each pitch are used to compute new averages, and these are used to set new target values which are automatically updated with each pitch.

### 3.2 Position estimation

The search area is small, so it is necessary to update its position to match the motion of the ball. The position of the ball in the next frame is estimated based on its position in past frames, and the search region is moved to the new position.

A Kalman filter is used to estimation the position. A Kalman filter is recursive, so it is not computationally expensive, and is suitable for high-speed processing. The position in image coordinates, \((p_x, p_y)\), and velocity, \((v_x, v_y)\) are used as the state value, \(x = (p_x, p_y, v_x, v_y)^T\) (where matrix \(A^T\) is the transpose of matrix \(A\)).

For each frame, the state is estimated by:

\[
x_{t+1} = Fx_t
\]

\[
P_{t+1} = FP_tF^T + Q
\]

With \( t \) being the frame number, \( x_{t+1} \) being the estimate of the state value for the next frame, and \( P_{t+1} \) estimating the error covariance. Here, \( F \) is the state transition matrix for uniform linear motion, and \( Q \) is the covariance matrix for a Gaussian noise distribution with mean of zero added to the state transition.

The \( x_{t+1} \) obtained from the state estimation equations gives an estimation of the ball position for the next frame, but if the ball position is extracted successfully from the frame, the extracted position (observed value), \( y(o_x, o_y) \), is applied using the update equation:

\[
x_{t+1}' = x_{t+1} + K(y - Hx_{t+1})
\]

\[
P_{t+1}' = (I - KH)P_{t+1}
\]

\(^2\) \( C_b \) and \( C_r \) are the blue and red color difference signal.

\(^3\) A matrix representing the accuracy of the state values.

---

![Figure 6: Selection of ball candidate regions](image_url)
Making $x_{t+1}'$ the estimated ball position for frame $t+1$ as well as updating the Kalman gain, $K$. $H$ is the transformation matrix from state values to observed values, and $R$ is the error covariance matrix for observed values.

In this way, the ball position in successive frames is estimated, while applying corrections for frames where extraction is successful. Also, small values for the diagonal elements of error covariance matrix, $P$, are used for frames when extraction is successful, and large values are used when it fails, setting an optimal search region size updated from one frame to the next with these values.

Also, when there have been $n$ or more earlier extracted points, a second-order best-fit curve is computed using a least-squares method, and this is used as the estimated trajectory. Short segments of a ball trajectory can be very closely approximated by a second-order curve, so this improves the accuracy and reliability of ball extraction, and tracking.

The estimated trajectory, a second-order curve using the lower left corner of the input screen as the origin, is given by the following equation:

$$y = ax^2 + bx + c$$

Here $a$, $b$, and $c$ are arbitrary constants.

The sum of square errors, $E$, of extracted points $(x_i, y_i)$ to the curve is given by:

$$E = \sum_{i=1}^{n} (y_i - ax_i^2 - bx_i - c)^2$$

In some cases when there are regions with luminance and chrominance similar to the ball, such as when the pitcher is wearing a white uniform, or when there is advertising with white lettering, ball extraction can fail. However, tracking continues, estimating successive positions, which allows detection to resume after passing through such regions. Figure 4 shows an example of tracking the ball through a scene with a left-handed batter wearing a white uniform. Past detected points are shown with white circles, the second-order trajectory estimate with a curved line, and the current search area with a rectangle. Ball extraction failed when the ball overlapped the batter, but resumed correctly after passing the batter.

When extraction has failed for a set number of successive frames the tracking process ends, determining that the ball has either been caught by the catcher or struck by the batter.

4. Trajectory drawing

The interpolation and drawing PC draws a trajectory computer graphic based on the ball positions output by the extraction and tracking PC. It performs interpolation and error correction using a second-order curve from the input ball positions, taking segments where extraction failed and incorrect extractions into consideration. This allows it to draw a smooth, continuous trajectory. Interpolation and extrapolation is also provided, so the trajectory can be extended into the past and future.

After performing interpolation and correction, a computer graphic is drawn over the trajectory for a fixed time interval. The interval, shape, color, size and brightness can be configured freely.

Figure 7 shows an example of a broadcast image, showing a continuous trajectory in spite of the left-handed batter region which made extraction difficult.

5. Application to TV programming

This system has been used since 2004 for instant replays of pitches in NHK professional baseball broadcasts. Display of the trajectory overlaid on the broadcast camera image has been well received by viewers, who have commented that it makes the ball path and amount of drop easy to see.

The system has also been used on educational or entertainment programs other than baseball broadcasts to show ball trajectories. For example, it was used to visualize softball trajectories on the "TV Sports Class" program on the NHK Education channel. It has also been used for pitching scenes from various directions other than that of the center camera, such as directly horizontal to the battery, or from the batter's viewpoint. Figure 8 shows such an example for softball.
6. Applications for other sports

6.1 Bowling broadcasts

The B-Motion baseball pitch tracking system has also been applied to bowling broadcasts. Bowling broadcasts had not been using a broadcast camera able to view the entire lane from a high position, so a new fixed camera was introduced and used to extract and display the ball trajectory.

Figure 9 shows a broadcast image from coverage of the 2008 NHK Cup Bowling Championship. The system was able to draw smooth and stable trajectories, regardless of bowler handedness or ball color.

6.2 Golf broadcasts

We have also extended this system for golf broadcasts, developing a system able to display trajectories for tee shots and putting scenes. Since camera operations such as pan and zoom are used in golf coverage for shots and putts, the trajectory cannot be drawn using only the ball position in two-dimensional screen coordinates. Thus, we used an encoder to record the broadcast camera
parameters (pan, zoom and focus values) for each frame. The ball trajectory is computed using these camera parameters and extracted ball positions, and superimposed on the broadcast camera image. Figure 10 shows the system organization used for golf broadcasts.

Figure 11 shows a broadcast image with composite tee-shot trajectory from the broadcast of the 2005 Japan Open golf coverage. In 2005, the system required stroke, landing and resting points to be specified manually, but it was able to display the ball trajectory overlaid on the camera image from tee to green. The system used for the 2006 Japan Open and Japan Women’s Open was able to display putt trajectories. Highly accurate trajectories could be drawn in semi-real-time, allowing very effective production with replays immediately after each putt. Figure 12 shows an example of a putt trajectory.

In all cases, the system could be operated using a single camera, allowing for live golf broadcast operation in the same style as earlier broadcasts, without the need for new additional equipment.

7. Conclusion

We have developed a baseball pitch trajectory display system called B-Motion that can be used in live baseball broadcasts. The system is able to display the trajectory of just the ball in semi-real-time, even if there are other moving objects in the scene, by using object extraction that takes the dynamic and visual characteristics of the ball into consideration as well as position estimation. The system uses only the image from the broadcast camera for input, so no additional photography equipment is necessary. It also requires no calibration beforehand, so it is very easy to use. The system has been used for live NHK professional baseball broadcasts through all seasons since 2004.

Extensions to the baseball system have been developed and used for live broadcasts including for ball trajectories in bowling, and tee shots and putts in golf. In the future we hope to devise even more-reliable and general object extraction and tracking methods, and expand the range of applications.

(Masaki TAKAHASHI, Toshie MISU, Mahito FUJII, Seiichi GOHSHI*, Nobuyuki YAGI, Yoshihiro FUJITA
* SHARP CORPORATION)