Tracking a the Speed of a Ball through a Scene

The goal of this project is to determine the speed of a ball at each frame in a video. In order to determine the speed of the ball at each frame the position of the ball in real space, the (x,y,z) coordinate values, needed to be found. To find the position of the ball in each frame, a pipeline of techniques were used to process each frame independent from each other. The techniques are color space conversion, background subtraction, connected components, region analysis, and stereo vision with one camera. Each frame went through the following pipeline.

First the raw RGB frame was converted to the YIQ color space. This was done to deal with shadows are reflections. The Y color channel is dropped, and only the I and Q channels are used for now on.

Next, image difference is done to the frame and the first frame in the video. It is assumed that the first frame does not have the ball in it and that the the only difference between the first frame and this frame is that the ball may be in it. This means the camera must stay very still and the background of the video cannot change at all. This is one of the “General Cases of Motion”; “Still camera, single moving object, constant background” (Davis, Lecture slide on Motion). The difference value is calculated from Equation 1.

Equation 1:

This is the Euclidean distance between the first frame’s IQ channels and the current frame’s IQ channels. This operation was performed pixel-wise. This operation produces a difference matrix that is the same size as the original image. The absolute value of the difference matrix is calculated. A threshold with a value of T1 is used to convert the difference matrix into a binary matrix. A zero represents background. A one represents a large change between the first and current frame.

All the connected components are found in the binary matrix. The connected component with the largest area is marked as the ball. The rest of the connected components are set to zero. This is done because there could be some noise in the image that got through to the binary matrix. Only analysis on the ball wants to be done. It is assumed that the largest connected component represents the ball. This means no objects larger than the ball can be moving in the scene. Some frames in the scene do not have a ball. To handle this if the largest connected component’s area is not greater than T2 the frame is marked as not having a ball and is discarded.

Then, the X-centroids and Y-centroids are calculated for the connected component. These points represent the center pixel of the ball in the frame. The midpoint of the heigh and the width of the connected component were not used, because noise around the edge of the ball could skew the real position of the ball much more than using the centroids.

The height of the connected component was calculated to determine the diameter of the ball in pixels. The height was used because the ball mainly travels horizontally through out the video. Therefore, the height of the connected component in not affected by motion blur like the width of the connected component was.

Using the X-centroid, Y-centroid, and diameter, the real (x,y,z) values of the ball in 3-D space can be aproximated.

First, the distance of the ball from the camera is calculated using the diameter of the ball in pixels (previously calculated), the diameter of the ball in inches (known variable), the width of the frame in pixels (known variable), and the camera’s horizontal field of view (known variable). Using these three variables the distance can be calculated using equation 2 and 3, where c is circumference of the horizontal field of view at the position of the ball in inches, w is the width of the frame in pixels, dp is the diameter of the ball in pixels in the frame, din is the diameter of the ball in inches, d is the distance of the ball from the camera in inches, and FOVx is the horizontal field of view in degrees.

Equation 2:

Equation 3:

This distance can also be calculated using the vertical field of view and the heigh of the frame in pixels.

Next, the horizontal angle of the position of the ball with respect to the camera was calculated. Let this value be . This can be calculated by using equation 4, where x is the center of the ball in pixels from the left side of the image, and w is the width of the frame in pixels.

Equation 4:

Similarly, the vertical angle can be calculated using the vertical center of the ball the vertical field of view, and the height of the frame in pixels. Let this value be Θ.

d, , Θ, are the polar coordinates in the 3D coordinate system. It is trivial to convert these values to (x,y,z) cartesian value. The frame of reference is the camera. The camera’s position is (0,0,0).

These cartesian coordinates represent the position of the ball in real 3D space. Using these positions, it is trivial to calculate the distance of the ball between frames and the speed of the ball from frame-to-frame.

The speed of the ball between each frame was calculated and is show in figure 1.

methods used, problems encountered, outcomes, lessons learned, allocation of work for each team member, possible future work, etc.

Break up the story more, add problems in story, give good introduction