1   Problem Statement

The problem focuses on estimating the optimal radius for the ball in a given video frame. The video footage has a pitcher throwing the ball and the goal is to estimate the spin of the ball. We have developed an interface that shows the ball to a user and asks them to click on the centre of the ball, denoted by $\tilde{c}$. We predict the trajectory of the ball based on these points that are clicked by the user

2   Assumptions

We make the following assumptions about the problem.
1. Initially we assume that the we already know the centre of the ball. Formally,
$|c - \tilde{c}| = 0$
where c is the centre of the ball and $\tilde{c}$ is the pixel in the image which is clicked by the user.
2. The radius of the ball at a given frame t, denoted by r_t grows linearly with respect to the frames. Formally,
$r_t = \alpha r_{t-1}$, where $\alpha \geq 1$
3. The pixel intensity of the ball remains roughly the same for different frames. For a given frame t with centre and radius (c_t, r_t). We assume that that pixel intensity of the ball will be similar to its previous frames $\{t-1, t-2, t-3...., 0\}$ with centres and radii at $\{(c_{t-1}, r_{t-1}), (c_{t-2}, r_{t-2}), ..., (c_0, r_0)\}$. We use image histograms to compare pixel similarity between . The notation $H(c_t, r_t)$ denotes the histogram of all pixels in frame t that lie inside a circle with radius r and centre c.

3   Decision Variables

Since we assume that the centre of the ball is already known. Our, decision variable in this case will only be the radius of the ball r_t at a given frame t. The objective function can either maximize or minimize the decision variable. In our case we would like to maximize the radius of the ball. The reason behind choosing this objective is that minimizing the radius of the ball can lead to detecting a circle that is much smaller than the size of the ball with the aforementioned assumptions being true. Our decision objective should also take assumption 3 into consideration and ensure that the histogram of pixel intensity of the ball for a given frame is similar to the histogram of pixel intensity of the ball in the previous frame. We can use euclidean distance as a metric to measure the similarity between the two histograms.

4   Objective Function

The objective function is given as
$$\min_{r_0} \sum_{i=0}^{n} \{\lambda_1 ||H(\alpha^i r_0, c_i) - S_i||_2 - \lambda_2 \alpha^i r_0\}$$

subject to
$r_0 > 0$
$r_t = \alpha r_{t-1}$ where $\alpha > 1$
Here, $S_i = \frac{1}{i-1} \sum_{k=0}^{i-1} H_k(\alpha^k r_0, c_k)$
S_i is the mean value of all color histgograms prior to frame i. The objective function is defined for all set of frames till the present frame n. Here, λ_1 and λ_2 are the associated weights, that can be tuned manually. The the objective function associated with λ_1 tries to minimize the euclidean distances of histograms of the current frame and the mean color histogram of all previous frames. The optimization objective ensures that the pixel intensity of the ball remains the same and the offset while tracking the ball is minimized and also the objective function associated with λ_2 tries to ensure that the radius is not zero.

5 Extensions

We can also add a texture based descriptor to denote the shape of the ball.