Background and Foreground Estimation using Median Filtering

Basheer Ahammad Ragimanu

2022

1 Background Estimation using Median Filtering

Given a video sequence consisting of $K+1$ frames, we perform background estimation up to the Kth frame using a median filtering approach. The steps are as follows:

1. **Initialize:** Acquire K frames. For each pixel p and color channel c, calculate the median intensity $B_{p,c}$ of these K frames. This represents the current background estimation:

$$B_{p,c} = \text{median}(\text{frames}_{1:K,p,c})$$

2. **Acquire Frame $K+1$:** Obtain the $(K+1)$th frame of the video denoted as F_{K+1}.

3. **Compute Difference:** Calculate the absolute difference between the $(K+1)$th frame and the current background estimation at each pixel and color channel:

$$D_{p,c} = |F_{K+1,p,c} - B_{p,c}|$$

4. **Thresholding:** Apply a threshold τ to the difference image D to remove/reduce noise. The thresholded image T will be binary, with pixels above the threshold indicating potential foreground regions:

$$T_{p,c} = \begin{cases} 255 & \text{if } D_{p,c} > \tau \\ 0 & \text{otherwise} \end{cases}$$

5. **Region Filtering:** Use morphological operations to remove small regions and fill gaps in larger regions, resulting in a refined binary mask M.

6. **Update Background:** Incorporate the $(K+1)$th frame into the background estimation and discard the earliest frame from the background estimation:

$$B_{p,c} \leftarrow F_{K+1,p,c}$$

2 Foreground Binary Image at Frame $K+1$

To obtain the foreground binary image at frame $K+1$, we perform the following steps:

1. Load the updated background image obtained after K iterations.

2. Load the $(K+1)$th frame of the video.

3. Calculate the absolute difference between the $(K+1)$th frame and the updated background image:

$$D_{p,c} = |F_{K+1,p,c} - B_{p,c}|$$

4. Create a binary mask M by thresholding the absolute difference image D. Pixels above a specified threshold τ are set to white (foreground), while pixels below the threshold are set to black (background):

$$M_{p,c} = \begin{cases} 255 & \text{if } D_{p,c} > \tau \\ 0 & \text{otherwise} \end{cases}$$

5. Display the resulting binary mask M, where white regions represent potential foreground objects.

3 Results

Figure 1: Background Estimated Image

Figure 2: Foreground Image at K+1 Frame