Results from L1 optimal camera paths method

Ishank Juneja

July 14, 2020

1 2D Motion model

The motion of certain tracked feature points - tracked using the KLT tracking algorithm - are modeled using a discrete time 2D motion model. The choice of model is taken to be a 6 DOF affine transformation model. The raw unfiltered camera trajectory is modeled using a 6 DOF 2D motion model. Let F_t be the 2D 6 DOF transition model between the frames I_{t-1} and I_t. For a 2D affine model, with 6 DOF will be

$$F_t = \begin{pmatrix} a_t & b_t \\ c_t & d_t \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} \Delta x_t \\ \Delta y_t \end{pmatrix}. \tag{1}$$

As a transform in homogeneous coordinates, F_t would look like,

$$F_t = \begin{pmatrix} a_t & b_t & \Delta x_t \\ c_t & d_t & \Delta y_t \\ 0 & 0 & 1 \end{pmatrix}^T , \tag{2}$$

So at the next time instant a feature point (or any point once we apply the transform to all points) would be given by,

$$\begin{pmatrix} x_t \\ y_t \\ 1 \end{pmatrix}^T = \begin{pmatrix} x_{t-1} \\ y_{t-1} \\ 1 \end{pmatrix}^T F_t. \tag{3}$$

In 2, the transpose operation is needed since we use right multiplication of transforms instead of the usual left multiplication. The camera trajectory C_t (called ‘C_trajectory‘ in the code) is defined as the accumulation of these transitions F_t,

$$C_{t+1} = C_t F_{t+1} \implies C_t = F_1 F_2 \dots F_t. \tag{4}$$

So to track the motion of a point (x_0, y_0) in pixel coordinates under this motion model, we can perform the operation

$$\begin{pmatrix} x_t \\ y_t \\ 1 \end{pmatrix}^T = \begin{pmatrix} x_0 \\ y_0 \\ 1 \end{pmatrix}^T C_t. \tag{5}$$

This was for the raw un-stabilized trajectory, for the stabilized camera trajectory we can perform the same computation with the stabilized camera trajectory P_t, which is called ‘P_trajectory‘ in the code.

$$\begin{pmatrix} x_t \\ y_t \\ 1 \end{pmatrix}^T = \begin{pmatrix} x_0 \\ y_0 \\ 1 \end{pmatrix}^T P_t. \tag{6}$$

2 Examples

The examples are from the dataset linked here Regular.zip. The trajectories are generated by tracking the motion of the homogeneous coordinate $x_0 = (0, 0, 1)^T$ over time. In this case we plot the y coordinates over time, i.e the plot is for $x_t(2)$ vs t for the original and stabilized trajectories.

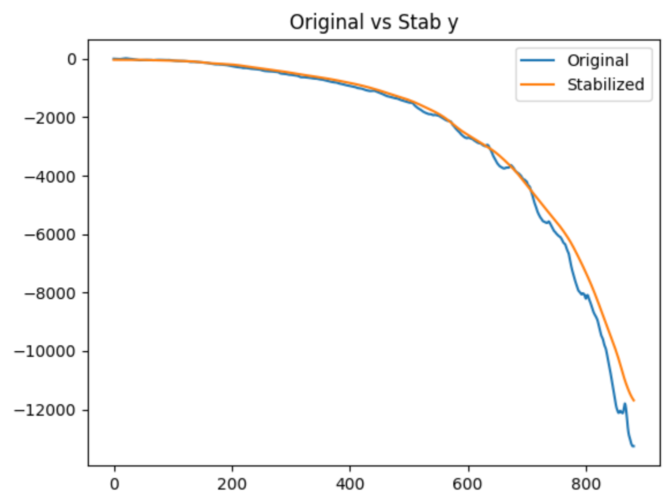

Figure 1: Example 22.avi. x axis is the frame time step which relates to the original fps rate of the video file, the y axis is the motion of the point $(0, 0)$ in pixel coordinates. In this particular example, the range in the y axis of displacement appears to be large because of the slightly larger length of the video (900 frames) and the fact that the video is a person walking continuously in a single direction with camera in hand.

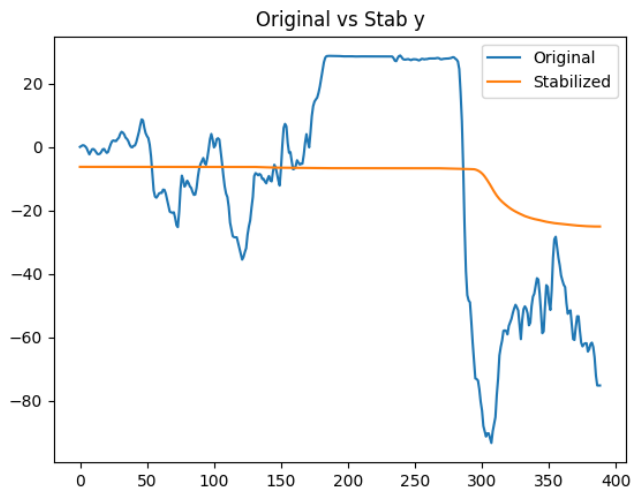

Figure 2: Example 13.avi. On the Y axis we have purchase probability and income brackets on the x axis