Tracking Objects in Videos

Computer Vision

—

Abdallah Yasser Youssef	18015026
Ahmed Bahgat Elsherif	18010078
Mohamed Metwalli Noureldin	18011587

Overview

Lucas-Kanade alignment is an algorithm which aligns a template to be found in an image based on a geometric transformation between the template image and the test image. The model of transformation can be chosen based on the nature of the dataset at hand. For example, if we only have translation between the template and the test image, we can have a model with only two degrees of freedom. For our problem, we used an affine transformation model which allows for translation, rotation, shearing, and scaling in both axes.

Assumptions

- An initial guess of the transformation parameters between the template and the model to be found.
- For the tracking problem, it is assumed that the time step between every two consecutive frames to be small such that no apparent changes between the two frames are found.
- We assume the differences are mostly geometrical and the differences in brightness or intensity values between the frames are small or negligible.

Algorithm

Parameters

- Bounding box coordinates
- Tolerance value for dp
- Maximum number of iterations
- Template and video frames

Steps

- Load the template image and set the parameters of the bounding box of the object to be tracked
- Load video images one by one
- To maintain brightness constancy, it was empirically found that applying histogram equalization algorithms produced slightly better bounding boxes
- The jacobian is calculated once at the beginning from the template points to be warped
- The warping matrix is initially set to equal the identity matrix

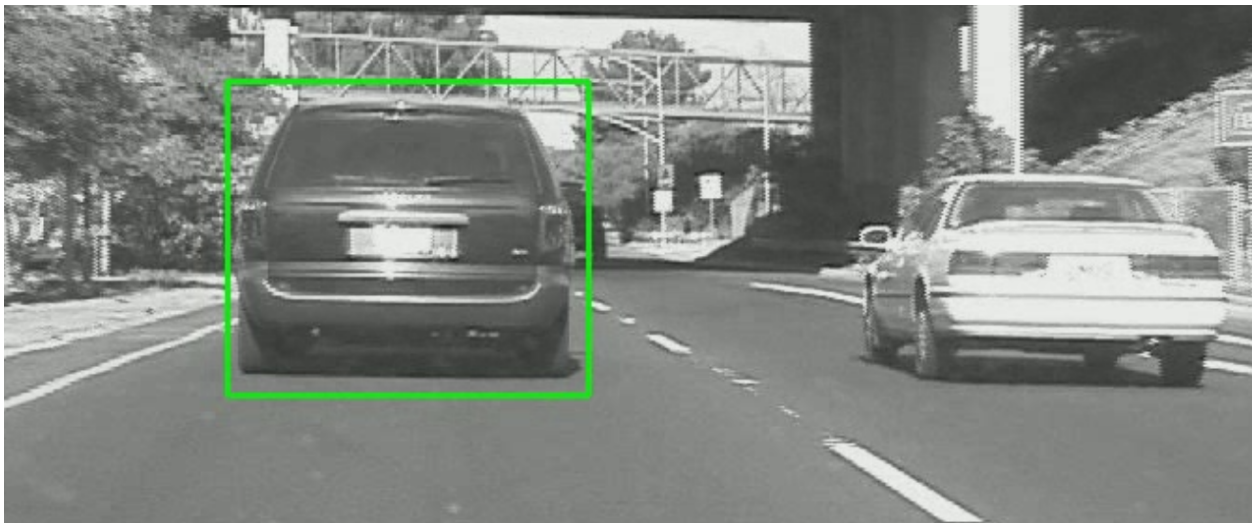
- Image gradients are calculated
- The warped image and gradient intensity values are extracted
- Hessian matrix is calculated
- The inverse of the Hessian is multiplied by the terms containing the error of prediction
- This process is repeated multiple times until the dp is smaller than a predefined tolerance or until a maximum number of iterations has been executed which are algorithms parameters

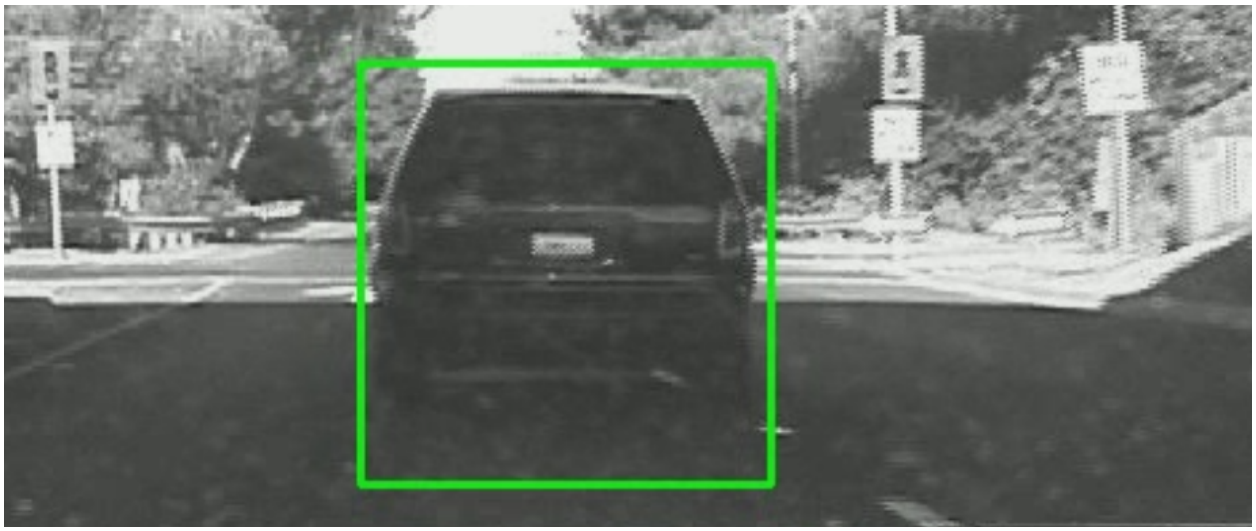
Output

- Video frames annotated with a bounding box around the tracked object

Results

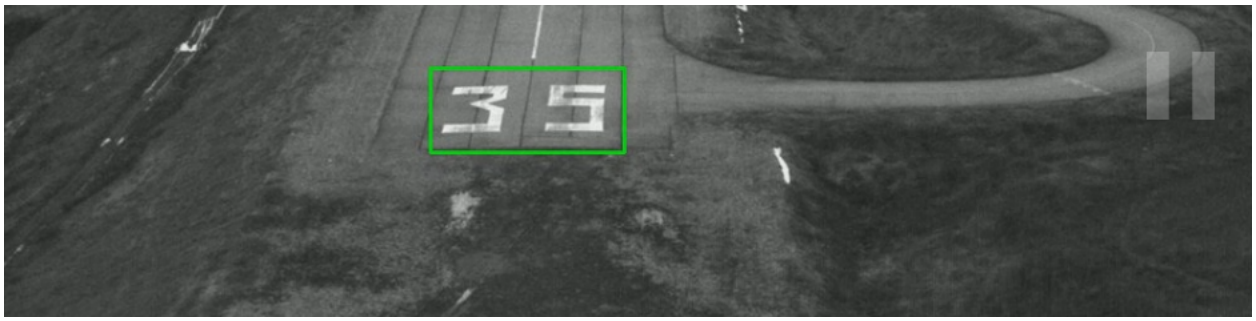
Car dataset

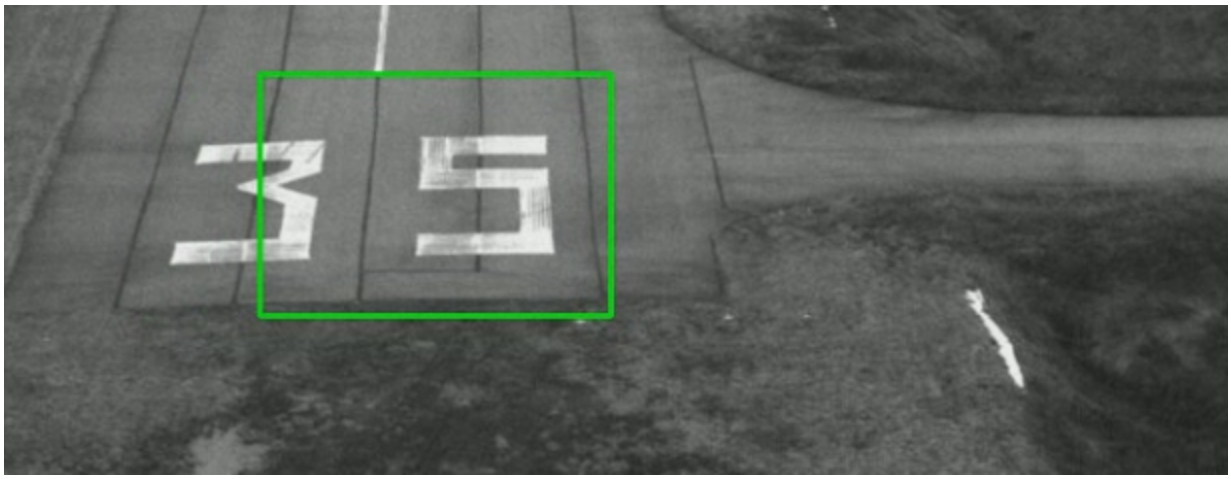




In this frame we can see that the bounding box is enlarged due to the fact that brightness constancy has been violated, but it is restored in the next few frames

Landing dataset





We can see here that the bounding box in the end is not quite accurate due to the fact that the time step between consecutive frames is large (we can see that frame numbers in files sometimes have increments of four frames). This is why some frames have been excluded in the end.