Pedestrian Tracking from a Moving Host Using Corner Points

Mirko Meuter[1], Dennis Müller[2], Stefan Müller-Schneiders[2], Uri Iurgel[2], Su-Birm Park[2], and Anton Kummert[1]

[1] Faculty of Electrical Information and Media Engineering
University of Wuppertal
D-42119 Wuppertal, Germany
[2] Delphi Delco Electronics Europe
Advanced Engineering
D-42119 Wuppertal, Germany

Abstract. We present a new camera based algorithm to track pedestrians from a moving host using corner points. The algorithm can handle partial shape variations and the set of point movement vectors allows to estimate not only translation but also scaling. The algorithm works as follows: Corner points are extracted within a bounding box, where the pedestrian is detected in the current frame and in a search region in the next frame. We compare the local neighbourhood of points to find point correspondences using an improved method. The point correspondences are used to estimate the object movement using a translation scale model. A fast iterative outlier removal strategy is employed to remove single false point matches. A correction step is presented to correct the position estimate. The step uses the accumulated movement of each point over time to detect outliers that can not be found using inter-frame motion vectors. First tests indicate a good performance of the presented tracking algorithm, which is improved by the presented correction step.

1 Introduction

Camera-based optical tracking of objects plays an important role in visual surveillance and sensing applications, and a wide variety of methods has been developed. Simple methods for optical tracking like correlation work straightforwardly by extracting a template region from a source image or edge image [CSR94]. This region is convolved with a search region in the most recent frame to find the area with the maximum matching. The position where the maximum matching was found is used to calculate the displacement vector from the previous to the current frame. More advanced methods work based on the selection and extraction of features like edges, contours or corner points [LF05]. Especially corners are an obvious choice for tracking. They provide accurate localization and usually, the image has a high information content in the image area surrounding these points [SMB00]. There are two major methods to track objects based on feature points. One method is the one time detection of features and their subsequent

G. Bebis et al. (Eds.): ISVC 2007, Part II, LNCS 4842, pp. 367–376, 2007.

tracking by relying on the optical flow e.g. the Kanade-Lucas-Tomasi (KLT) tracker. Such an approach has trouble with the appearance or disappearance of feature points due to partial occlusion or aspect changes which can lead to subsequent tracking failures [LF05]. The other method is the detection of feature points in each frame and their subsequent matching. To track an object, the points are extracted in the region where the object was detected or estimated and additionally in a search region in the next frame. The image region in the neighbourhood of each point is compared to the neighbourhood of points in the next frame to generate a list of point matches. This method yields a set of point movement vectors that can contain outliers caused by false matches and points on the background together with movement vectors from the object of interest. [LF05] reported about the application of interest point tracking to estimate the pose change of a plane. They employ the RANSAC algorithm to handle the outlier problem. In [AB01], interest points have been used to track vehicles from a moving host. They proposed to use a robust estimator to solve the outlier problem. From our knowledge, both techniques give good results, but can have high requirements regarding processing time. Smith [Smi95] presented the ASSET-2 tracking system for vehicles. They estimate the point flow using a Kalman filter for each corner to reduce the outlier problem. Then they use a motion segmentation algorithm to assign points to the object of interest or to the background. Their example videos show a good performance, but in these examples, the motion is very simple separable. [PGV05] presented the application of interest point tracking for pedestrians, but this application considered translational movement only which will lead to problems under scale changes.

We assume that we have a detection algorithm that is able to deliver a good bounding box from the object of interest in every n^{th} frame. In practice, we simulate this behaviour by hand labelling our object of interest and feeding the box into the tracker as initialization. Then we proceed as follows. We extract corner points for the object in a region of interest and in a search region in the next frame using the well known Harris corner detector [HS88]. The local neighbourhood of each point in the current frame is matched against the local neighbourhood of close corner points in the next frame to find point correspondences. A modification of the method described in [Smi95] is used to generate the matching results. This method has low run time requirements, gives good results and needs to compute less comparisons than a 3×3 SSD comparison. For movement estimation, we employ a simple model for pedestrian tracking using only translation and scale. This model seems to be sufficient for our needs, and has to our best knowledge not previously been applied for pedestrian tracking using interest points. Outliers in the set of correspondences are rejected by iteratively estimating the model coefficients using an MMSE estimation and removing those points where the residual between the matched and the expected position exceeds a certain threshold. Even if this is not a robust estimation technique, the algorithm works in most cases for a minority of outliers and is still very time efficient.