Motion Estimation and Interpretation for In-Car Systems

van Leeuwen, M.B.

Publication date
2002

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 6

Summary and Conclusion

In this thesis methods and algorithms are investigated for vision based motion estimation and interpretation. We consider the application where a single camera is attached to a moving vehicle, typically driving on a highway. Four characteristic aspects for motion interpretation are distinguished.

In the first place we consider the impact of egomotion on the interpretation of the observations. We estimate the mean viewing direction of the camera (measured relative to the driving direction of the vehicle) and the inter-frame camera rotations. Three observations are important when the issue of egomotion estimation is addressed. (i) The dynamical characteristics of these two parts differ considerably, asking for different estimation techniques. (ii) Both estimation techniques employ data observed in the background. In order to be robust to the large diversity in appearance of the background only moderate assumptions are made on the appearance of this data. (iii) The third observation only concerns the inter-frame rotations. We exploit the redundancy in the information provided by the optical flow field of background points. It is crucial to realize that a re-weighted orthogonal least squares optimization technique is required for the estimation of the rotation parameters from the measurements. The information about viewing direction and camera rotations makes accurate stabilization of the video stream possible. The high accuracy, needed to enable motion interpretation, is obtained.

Vehicle detection is the next aspect of the interpretation process. It concerns separation of visual data belonging to objects (i.e. vehicles) from background data and represents the initiation phase for vehicle tracking. It is important that vehicles are being detected in an early stage and with a high probability. Not recognizing the presence of a vehicle can have disastrous consequences in practical applications. False detection of objects as vehicle is relatively easy taken care of in the tracking phase. We selected clues for vehicle detection, based on their computational complexity and advantages and disadvantages regarding our application. Robustness to the large diversity in the appearance of vehicles is of course an important factor. Therefore, different clues are assigned to detect passing vehicles (observed 'in the corner of the eye') and vehicles observed in the distant or at mid-range. For passing vehicles, we combine temporal differencing, shadow and motion clues into a single detection scheme by analyzing projected motion. Analyzing the projected motion appears to be a new approach. For mid-range and distant vehicles we
developed an efficient strategy that employs shadow, entropy and horizontal symmetry clues for detection. Each of these clues has been applied for vehicle detection before. Combining them distinguishes our method from earlier proposed approaches. This ultimate result consists of two algorithms, both capable to detect vehicles in practice with a high probability. In experiments, all passing vehicles are correctly detected. Mid-range and distant vehicles are detected at high detection rates (within two or three frames).

Once a vehicle is detected its position in the image plane is tracked through the image sequence until it moves from the field of view. Also with vehicle tracking the issue of robustness plays an important role. The algorithm should stick to the vehicle under consideration and is not allowed to be distracted by the presence of other vehicles or changes in environmental conditions. High flexibility is demanded. This points namely to the diversity in appearance of both vehicle and background (and their interaction). The potential rapid fluctuations in illumination of the scene adds to this demand for flexibility. We designed a platform suitable for vehicle tracking in practice. The platform is based on matching of both static and dynamic 2D templates to the visual data. To avoid distraction of the matching procedure by similar objects in the scene a combination of a Kalman filter and multiple hypotheses reasoning is applied. The dynamic structure of the platform, together with the flexibility in the design of these templates makes flexible, robust, efficient and accurate performance possible in practice.

This leaves us with the translation of the vehicle motion observed in the image plane into the motion or behavior of the vehicle relatively to the camera. Two motion parameters that characterize motion interpretation for our application are discussed. The first one concerns the lane in which the vehicle is observed and the direction its heading to (parallel to our lane or not). This information is directly obtained from the output of the vehicle tracking algorithm. More complicated is the derivation of the second parameter: the time that elapses until a vehicle reaches our vehicle. In literature, this parameter is usually referred to as the time-to-contact. After reviewing different techniques developed to estimate this parameter, three approaches that appear to be most appropriate for our application are considered in detail. The measurement input for these approaches respectively consists of the horizontal position in the image plane of a point observed on a vehicle, the vertical position and the observed vehicle width. Tracks in the image plane of (for example) only the lower left and right 'corner of a vehicle' provide enough data to freely switch between the three approaches. These tracks can robustly be obtained in practice. Another important aspect of these approaches is that they provide good means for temporal smoothing of the measurement input. This is important to suppress the influence of disturbances in the observations in practice. The similar structure of the three approaches enabled us to design a single (new) temporal filtering technique. From the observations one can determine which approach will lead to the most accurate estimate for the time-to-contact for the given situation. Both lane-shifts and the time-to-contact are estimated from real image sequences. The main contribution of the chapter about motion interpretation is the insight it provides in different approaches for the estimation of the time-to-contact, in relation to our application.
The practical value of the methods and algorithms proposed in this thesis is shown by means of real image sequences. These sequences were obtained with a camera placed in a vehicle that was mainly moving on the highway. Observations were made through the windshield as well as through the rear-window of the vehicle.