

CHAPTER 7. DYNAMIC VISION FOR PERCEPTION AND CONTROL OF MOTION

E.D.Dickmanns and H.J. Wuensche

Content

7.1 Introduction

7.1.1 Active vision

7.1.2 Dynamic vision

7.2 Simultaneous representations on differential and multiple integral scales

7.3 Application areas discussed

7.4 Sensory information used

7.4.1 Conventional sensors

7.4.2 Vision sensors

7.4.3 Global Positioning System (GPS-) sensor

7.5 Dynamic perception: The 4-D approach exploiting spatio-temporal models

7.5.1 Basic assumptions underlying the 4-D approach

7.5.2 Survey on the cybernetic structure of the 4-D approach

7.5.3 Generic 4-D object classes

 Motion description

 Shape and feature description

7.5.4 Image feature extraction

 Edge detection with CRONOS

 Area-based feature extraction: The 'TRIANGLE'-algorithm

7.5.5 Feature aggregation

7.5.6 State estimation

7.5.7 Situation assessment

7.6 Generation of behavioral capabilities

7.7 Multiple loops in dynamic scene understanding

7.7.1 Flexibility and efficiency from intelligent control

7.7.2 Multiple loop closure on different time scales

7.8 Experimental results

7.8.1 Road vehicles

 Road recognition and relative ego-state

 Relative state to other vehicles

7.8.2 Autonomous visual navigation of air vehicles near the ground

 On-board autonomous landing approach of aircraft

 Landmark navigation for helicopters

7.9 Conclusions

7.10 Literature

7.1 Introduction

Digital image processing started almost four decades ago with rectification and preprocessing of images from remote sensing and from medical applications for human interpretation. Processing time, usually, was in the one-hour time range per image. For a long time, image processing was confined to snapshot images since the data volume of 10^4 to 10^6 pixel per image and the relatively low processing speed of computers up to the 80ies did not allow digital real-time video data processing.

Change detection in images became of interest as data on the same scene but temporally well apart became available; this was of interest both for medical diagnosis and for military intelligence. Starting from this background, the idea of motion recognition and vehicle control by computer vision came up in the context of planetary rover vehicles. Initial test vehicles only had the camera on board, but the computers required were located back in the laboratory; they typically needed a quarter of an hour for image analysis; this was the state in the late 70ies [Stanford cart]. The vehicles moved a short distance and had to wait standing still as long as the image was analyzed. Each image was interpreted from scratch.

At that time, E. D. Dickmanns, coming from the field of control engineering, had the idea to start developing machine vision from a different paradigm. Instead of choosing a visually complex environment and work with very low image frequencies (10^{-1} to 10^{-3} Hz) as usual at that time in 'artificial intelligence', he decided to keep image frequencies rather high (larger than 10 Hz) and visual complexity of the scene to be handled correspondingly low. This low cycle time of 100 or less milliseconds (ms) would allow exploiting temporal and visual continuity in the scene since only small changes had to be accommodated. The change of location of the camera observing the scene was to be modeled so that predictions would allow reducing search space for the same features drastically.

Road scenes, especially roads built for high-speed driving, seemed to be sufficiently simple visually for providing the first field of application. The surfaces are smooth and well kept, and the lanes are marked with bright colors for easy recognition by humans; in addition, in order to keep acceleration levels on human passengers at high speeds low, both horizontal and vertical curvatures are limited in magnitude. The model for this type of road can easily be stated mathematically. Since vehicles are massive bodies which do not jump around but move steadily, also their motion can easily be described by differential equation constraints including the effects of control and perturbation inputs.

All of this has led to a completely different approach to machine vision than those studied by researchers from the field of Artificial Intelligence. Recursive estimation based on incomplete measurement data, developed in the 60ies for trajectory determination in aerospace applications, was well established in the engineering disciplines at that time. These methods have been extended to image sequence processing for full state reconstruction in 3-D space and time at the author's institute (UBM). The resulting method had been dubbed '4-D approach', in contrast to the 2-D, 2.5-D, and 3-D methods under discussion then, disregarding time as the fourth independent variable in the problem domain. The numerical efficiency and compactness in state representation of recursive estimation, which directly allowed control applications for generating behavioral capabilities, finally, led to its widespread acceptance in the community dealing with autonomous vehicles (mobile robots)