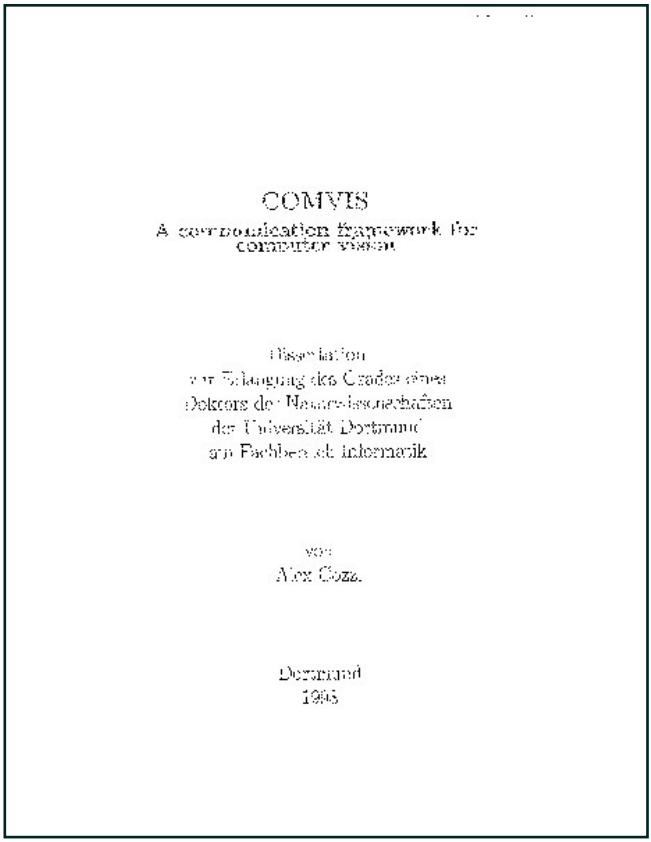
Alex Cozzi (Autor)

Comvis a communication framework for computer vision



Introduction

The visual sense is our principal source of information about the external world. Thanks to our visual capabilities we can deal with a complex and quickly changing three-dimensional environment in a competent and efficient manner. We are so accustomed to the performances of our effortless seeing ability that the great difficulties of science in understanding the visual sense and of technology to build machines able to see is a source of continuous surprise and frustration.

This work is a contribution to the advancement of the discipline of computer vision. Its main goal is to explore the feasibility of building advanced computer vision systems on the basis of the principle of the diversification of the information sources.

I will show that building a computer vision system which relies on the communication of several different vision strategies is a fertile approach, leading to better and more reliable systems.

The motivation of the communication approach is derived from the failure of computer vision systems based on a single kind of image cue to cope with the complexity of real images.

This thesis is built on the experience of the computer vision community, adopting the different techniques that have been devised to extract useful information from images, and combining them in a way to compensate for their weaknesses.

The communication approach can be synthesized in improving the performances of current computer vision systems not through more complicated algorithms but through the interaction of several simple algorithms.

This study is divided into three parts: the theoretical background, the implementation and the validation of the approach.

In the theoretical part the problem is formally described and possible solutions are derived. In the implementation it is described how the solutions are embodied in a computing architecture. Finally, in the validation, the algorithms are tested on synthetic and real data and their performances characterized.

A guiding principle of this study has been to minimize the use of heuristics as much as possible, proceeding in a rigorous and mathematically well-grounded manner.

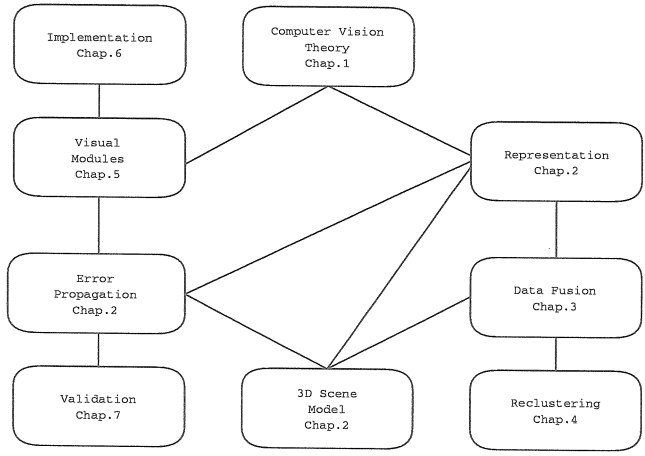


Figure 1: The connections between the main topics of the thesis.

Structure of the Exposition

Figure 1 shows the connections between the principal topics defining the chapters of this thesis. The theory of computer vision represents the background and the context needed to understand the objectives and the motivations of this work.

This theory is connected with the discussion of the specific visual modules and with the problem of how to represent scene knowledge. The representation is connected with the problem of representing the error of the collected information, how to estimate this error and how to study its propagation through the several processing stages.

The solution of the representation problem leads to the study of how to fuse and communicate the information. The communication and data fusion mechanism is the first objective and the major contribution of this study. The proposed solution builds on the theory of stochastic system modeling and Markov random fields to propose a coherent and original procedure.

This procedure has been embodied into an actual complete computer vision system, the COMVIS (**COM**municating **VIS**ual modules) system, whose structure is described in Chapter 5; its implementation is the topic of the Chapter 6, where the object-oriented class library developed in this thesis is detailed.

Chapter 7 deals with the experimental assessment of the approach's validity. The software system and the theoretical approach are validated on a set of simulated and true scenes and their performances measured. In this chapter the improvements achievable by the communication process are quantified. The final chapter is dedicated to summarizing the achievements of this work and to

the discussion of its limitations and possible extensions.