Technical Report

Dept. for Pattern Recognition and Image Processing
Institute for Automation
Technical University of Vienna
Treitlstr. 3/1832
A-1040 Vienna AUSTRIA
Phone: +43 (1) 58801-8161
Fax: +43 (1) 569697
E-mail: bis@prip.tuwien.ac.at

PRIP-TR-40 August 10, 1995

Theory and Applications of Digital Image Processing and Pattern Recognition

Horst Bischof and Valentina Filova
(Editors)

Abstract

This report contains the scientific texts of the original proposal of the Austrian research program "Theory and Applications of Digital Image Processing and Pattern Recognition". This program is supported by the Austrian science foundation FWF. This research program shall bring together research groups from Austria that have been, still are and intend to be active in the research of image processing and pattern recognition. The program also includes both theory and application within the individual projects. The program is organized into 5 different projects, and each project contains up to 4 tasks. The cooperative nature of the program is documented by the fact that 11 different institutes will work together. Many institutes take part in more than one project. It will stimulate the research activities for each of the participating groups (and beyond) in a way which has a positive long term effect for activities in this field in Austria.

Contents

Introduction

The human visual system as a functional unit including the eyes, the nervous system, and the corresponding parts of the brain, certainly ranks among the most important means of human information processing. Typical vision tasks that humans perform nearly without any conscious effort are:

- recognition of 'interesting' detail in a complex scene (e.g. a good friend on a busy street);
- fast interpretation of local changes and appropriate reaction (e.g. driving a car);
- visual comparison (e.g. identification of a known human face);
- storing and retrieving of pictures (e.g. the local environment where one lives, a mountain scenery, etc.);

The efficiency of the biological systems in such areas are beyond the capabilities of today's technical systems even with the fastest available computer systems.

However there are areas of application where digital image analysis systems produce acceptable results. Systems in these areas solve very specialized tasks, they operate in a very limited environment, and high speed is often not necessary. Several factors determine the economical application of technical vision systems: cost, speed, flexibility, robustness, functionality, and integration with other system components. Many of the recent developments in digital image processing and pattern recognition show already some of the required achievements. Computer vision enhances the capabilities of computer systems

- in autonomously collecting large amounts of data,
- in extracting information,
- in perceiving its environment, and
- in automatic or semi-automatic operation in this environment.

The development of computer systems in general shows a steadily increasing need in computational power which comes together with decreasing hardware costs. Under these premises it can be foreseen that computer vision systems will find a broad field of application in the very near future.

The development of practical solutions involving digital images requires the **cooperation** of specialists from many different fields. The wide range of fields (scientists from 12 different fields are cooperating) makes sure that the group of applying institutions fulfills this important requirement. Furthermore the often very specialized vocabulary in the different disciplines makes it necessary to have experts in the different areas, which on the other hand are in a close contact and permanent exchange of ideas. For this reason active cooperation between the different groups has been declared an important goal of this proposal. It will stimulate the research activities for each of the participating groups (and beyond) in a way which has a positive long term effect for activities in this field in Austria.

General Problems and Methods

In this section we would like to explain some of the fundamental steps that could be performed to interpret the contents of one or more given digital images.

Many principles from physics can be used for image formation. Natural light reflected by the surface of an object is still among the most frequent sources, however also active sensors like radar systems or laser range finders produce useful digital images. Although the development of new sensors is generally a part of the field of communication techniques, the geometrical and the radiometrical formation principles must be understood to interpret the image data correctly.

The processing of digital images by computer is a new challenge for computer science, because a huge amount of data (a single color image needs close to a megabyte of storage, but even larger images are possible, e.g. 1 Gigabyte for an orbit of the Magellan mission, see Task 1.4) has to be processed and often in extremely short time. New parallel process and operation structures are necessary to achieve acceptable response times.

The tasks of preprocessing steps include the removal of noise, the calibration of the images, and the enhancement of image quality. The resulting images are input to procedures that transform their content into a (compact) form suitable within the scope of the given application.

¿From the many methods for analyzing and processing digital images we just list few typical examples here:

- image filtering, reconstruction and smoothing;
- feature extraction;
- image transformations like Fourier, Gabor or wavelet transforms;
- statistical pattern analysis;
- classification and spatial clustering;
- stochastic methods;
- image compression for efficient transmission, storage, and retrieval;
- massively parallel processing structures like arrays, pyramids, hypercubes;
- correlation for comparing images or image parts;
- methods that integrate information from multiple images over time (multitemporal images) and over space (3-dimensional images, stereo pairs);
- image understanding methods for deriving image descriptions ranging from object-oriented representations up to simple natural language formulations;
- adaptive methods for image analysis (ranging from adaptive filtering up to automatic construction of model databases).

Methodological Classification

This research program covers a large variety of different aspects of digital image processing. In Table 0.1 very global aspects of the methodologies are differentiated in order to roughly classify the focus of the four different scientific projects of the program.

Processing stages can be divided into low level and high level. **Low level methods** at the one end include all the preprocessing steps. At the other end are the image understanding methods that are

Table 0.1: Methodological Classification

Type of method (explanation in text)	Project I	Project II	Project III	Project IV
Low level methods	•	∘	∘	•
High level methods	•	•	•	∘
Numerical computation	•	∘		•
Symbolic computation		•	•	∘
2D result	•	•	•	∘
3D result (spatial)	•	∘	•	•
3D result (temporal)	∘	∘	•	
4D result (spatial & temporal)		∘	•	∘
1 scene/few images	•	•	•	•
Large sets of images	•	∘	•	
Time critical	∘	∘	•	•
Detail solutions	•	∘	∘	•
Integrated system	∘	•	•	∘

• denotes main interest,
∘ that the class is also included,
no entry means no clear assignment possible.

normally classified as **high level**. The separation between the two classes is not very precise, but can be drawn approximately by considering the output of a method: all methods producing image-like (array) representations belong to the low level class, whereas all methods producing language-like descriptions form the high level class. A related distinction concerns the type of the data: **numerical** and **symbolic** computation and formats can be distinguished.

The result of the analysis can be either **two-dimensional** ('**flat**') or **three-dimensional (spatial)** (e.g. from stereo, or shape-from-X methods), or **three-dimensional (temporal)** (e.g. motion of 2-D objects), or **four-dimensional (spatial & temporal)** (e.g. motion of 3-D objects).

Furthermore, it is of importance, if a method deals with a **single scene** consisting of only a small number of images, or if **large sets of images** have to be scanned for a specific image property.

Another important distinction is, if the method should be applied in a **time critical** environment (e.g. real-time), or if enough time is available for processing of the images.

Finally there are approaches that solve one problem completely and in all the necessary detail (**detail solutions**). Others attack the problems that are encountered when the many different modules of a system have to function all together for the goal defined by the application (**integrated system**).

Development in Austria

Image processing and pattern recognition are relatively new scientific disciplines that have emerged from several other disciplines. Single research projects in Austria date back to the early 1970.

Austrian Association for Pattern Recognition

In 1981 the Austrian Working Group for Pattern Recognition (*Österreichische Arbeitsgruppe für Mustererkennung*, ÖAGM) was created as an informal platform of communication. Close cooperation exists with the Austrian Computer Society (*Österreichische Computer Gesellschaft*, OCG) and with the Austrian Society for Cybernetics (*Österreichische Studiengesellschaft für Kybernetik*, ÖSGK). The objective of this working group was the organization of regular workshops and the publication of proceedings of these workshops. Most of them appeared at Oldenbourg in OCG-*Schriftenreihe*, Vols. 14, 15, 29, 36, 42, 49, 56, 61, and 62. In 1982 the ÖAGM became a member of the International Association for Pattern Recognition (IAPR), which actually includes member organizations from about 30 countries of the whole world.

A major event for the ÖAGM was the joint organization of the DAGM symposium 1984. The proceedings were published in the *Informatik Fachberichte*, Vol. 87, of Springer. In 1994 another such joint meeting of ÖAGM and DAGM will take place in Vienna. In 1996 the International Conference on Pattern Recognition (ICPR) will be held in Vienna.

The informal working group became an independent non-profit association in 1987: **Austrian Association for Pattern Recognition** (*Österreichische Arbeitsgemeinschaft für Mustererkennung*, ÖAGM). The aim of the new association is to promote pattern recognition in Austria and to represent the Austrian interests within the international community. It continues the successful activities (workshops, proceedings) of the working group. The ÖAGM has 92 individual members both from Universities and from industry. The main research centers are located in the University cities: Wien, Graz, Linz, Innsbruck, Klagenfurt, and Salzburg.

Previous Projects

The Austrian Science Foundation (FWF) funded 4 projects in an early phase of this research program (Coordination, Mathematical Methods and Tools for Digital Image Processing, Image Understanding and Standardized Image Processing, Stereovideometry and Spatial Object Recognition). Ten institutes participated in these projects and these institutes form also the core of the present proposal. During these projects the necessary infrastructure for efficient communication between the participating institutes has been established (e.g. ftp-server, e-mail). Among several institutes a common platform (KBVision) for image analysis has been established. Mutual visits of the cooperating institutes have allowed an efficient use of available hard- and software.

Seven workshops where held during the last two years. Out of these workshops several new ideas for collaboration arose. These ideas are documented by the individual project proposals of this reasearch program. In summary, these previous projects have provided a rich environment for research in the area of pattern recognition, which shall be continued and extended during the proposed research program.

The role of the proposed Research Program

The preparatory work of the ÖAGM (workshops and publications) and the previous projects served as a platform for presentation of individual research projects in the field. The ÖAGM offers its services to the participants of the program, in particular, it offers to jointly organize the annual workshops of the program with the regular ÖAGM-meetings as already done in 1992 and 1993.

In order to combine the existing know-how, an organized cooperation of the individual groups seems to be appropriate at the present time. Image processing is still a growing field, in which new concepts are welcome, and where a big practical success can be achieved by intelligent approaches with relatively modest funds. The growth in the interest in image processing can be seen in the increasing number of participants at conferences (e.g. CVPR 1993 in New York, had 450 participants, whereas in previous years it had about 300 participants).

Austria has the chance to take advantage of this favorable opportunity and to participate in (and contribute to) the international development of this field during its expansion phase by intensifying and coordinating the research efforts. This effort would also offer the possibility to produce and to export the intelligent products that emerge of this initiative. Without such efforts Austria risks depending more than necessary from foreign companies and institutions in the future. In this case the missing know-how would have to be imported with great financial effort. In this sense investments into research in this area will certainly prove to be extremely benefitial, even from an economical point of view.

Relevance for research policy

Image Processing is one of the fastest growing fields of microelectronics and computer science. The requirements for multifunctional "image"-supported workstations and their fields of application in economy has been the topic of a dissertation [1]. A prognosis predicts for the American market an increase of low-level image supported systems from 12% in 1986 to 45% in 1995. The volume of the image processing market in Europe is expected to increase from 160 Million Dollars in 1989 to 500 Million Dollars in 1993 [2]. In the same study it was stated that only 5% of possible market for image processing has already been exploited.

This trend has already been noticed by several **international institutions**:

- In France a resrarch program (called GDR-TDSI, Research Group Signal and Image Processing) similar to this proposal has already been installed. There exist close contacts to several researchers of the French group, so that a cooperation is possible.
- In Germany a project entitled "Semantische Modellierung und Extraktion räumlicher Objekte aus Bildern und Karten" has recently been set up. In this program 10 projects will be funded. The program coordinator has expressed strong interests in a possible cooperation with this program.
- Similar projects will be or are already installed in the U.S.. For example DARPA and NSF have evaluated in various workshops the need for funding in image analysis.[3] This has also been expressed in a panel held recently at CVPR'93 in New York.

The research concept for microelectronics and information processing (*Forschungskonzept Mikroelektronik und Informationsverarbeitung*, ME + IV) of the Austrian *Bundesministerium für Wissenschaft und Forschung*, 1988, comes to a similar conclusion. In this study scientific proposals for research programs in Austria have been commented for their political and economical relevance by experts from industry. General conclusions of this committee are the need for *close cooperation*, a better coordination and communication, and the concentration on several important research initiatives. Pattern Recognition and Image Processing is among the proposed topics. Efforts are suggested in three areas:

1. conservation of the standard of research (basic and advanced);
2. short-term concrete joint projects (1-2 years); and
3. mid-term research areas, the **"topics of tomorrow"** (5 years).

Personal contacts with representatives from industry clearly indicate serious interest in the research proposed in the present program.

The proposed research program fulfills all the above requirements for the scientific side. There exists already interest of and contacts with representatives from industry to build an **accompanying industry commission** as a vital link that should make sure the transfer of technology developed within the program. In particular the program shall contribute:

- to basic and advanced research by
- top level scientific contributions (to participate in the international scientific communication);

[1] Peter Seidinger, *Anforderungen an multifunktionale, "Bild"-unterstützte Arbeitsplatzsysteme und deren Einsatzbereiche in der Wirtschaft.* Technische Universität Wien, 1988.

[2] K. Singer, *Nachholbedarf bei der Bildverarbeitung*, Bild & Ton, Bd. 45, p. 101, 1992

[3] Challenges in Computer Vision Research and Future Directions of Research, Final Report of the NSF Workshop, November 1991; Report on workshop on high performance computing and communications for grand challenge applications: Computer Vision, speech and natural language processing, and artificial intelligence. Report of NSF Workshop, November 1992

- to the dissemination of advanced know-how in image processing and pattern recognition and
- to the training of highly qualified graduates that constitute the main source of know-how that should enable Austria's industry to produce high level (intelligent) products.

The realization of the know-how to be developed within this research program is not limited to big reseach laboratories nor to big industry, also mid-range enterprises can successfully apply the new technologies. With the decreasing costs of advanced hardware for acquisition and processing of image data, the need for more than superficial knowledge of existing image analysis software packages will drastically increase. Therefore it can be expected that research projects in the area of image processing and pattern recognition can have a positive effect on the development of the Austrian economy.

Structure of the Program

Program goals

The two general goals of the proposed program are:

1. Advancement in Methology
2. Cooperation

The methodological classification of the various projects can be seen in Table 0.1. The types of methods to explore are documented in the individual project proposals. The general goal in this respect is that after the five years period of this program a rich set of methods for image analysis will be available. For each method it should be documented for what type of problem it is applicable and which group has the relevant knowledge about it. The compilation of the individual methods will be done by the coordination project. Having such a set of methods will favour applications in different areas. As a side effect of working in this program highly qualified personnel will be available to the industry.

The cooperative nature of the program is documented by the fact that 11 different institutes will work together. Many institutes take part in more than one project. From this cooperation several interesting issues arise, for example: a comparison of the developed methods is possible; resources (hard/software, data) can be shared among different institutes; new areas of application can be exploited. In order to achieve these goals, regular workshops will be held on specific topics, where problems can be discussed in depth. There will also be an annual workshop where all participating groups meet. This cooperation should guarantee that small groups can also perform competitive international research due to the synergistic effect of cooperation.

Projects and participating groups

The proposed research program contains four scientific research projects (Table 0.2).

Table 0.2: Project list

Proj.	Title	Coordinator	Inst.
O	Coordination	Prof. Dr. Walter G. Kropatsch	1
I	Mathematical and Algorithmic Tools for Digital Image Processing	Prof. Dr. Franz Pichler	5
II	Robust and Adaptive Methods for Image Understanding	Prof. Dr. Walter G. Kropatsch	2
III	Information Fusion and Physical Models for Image Understanding and Navigation	Doz. Dr. Werner Schneider	3
IV	Stereovideometry and Spatial Object Recognition	Prof. Dr. Heribert Kahmen	4

Each of these projects deals with a specific area of research which is harmonized among the partners within the project and among the projects of the program. The many interrelationships are pointed out in Table 0.1 and, in more detail, in the sections 'cooperation' in the respective projects. Almost all the active research topics are represented in one of these five projects. The strong cooperative nature of this program is highlighted by the different institutions that collaborate within each of the projects (Table 0.3). A complete list of participants can be found in section **??**.

Table 0.3: List of participating institutions

Research institution	located at
Dept. f. Pattern Recognition and Image Processing	TU Wien
Dept. of Mathematics	Univ. Wien
Dept. of Systems Science	Univ. Linz
Dept. of Statistics and Computer Science	Univ. Wien
Inst. of Surveying and Remote Sensing	Univ. of Natural Resources, Wien
Inst.f. Photogrammetry and Remote Sensing	TU Wien
Inst. of National Surveying and Enginieering Geodesy	TU Wien
Inst.f. Dig. Image Processing	Joanneum Research, Graz
Inst.f. Compter Graphics	TU Graz
Research Inst.f. Symbolic Computation	RISC Linz
Inst.f. Softwaretechnology and Parallel Systems	Univ. Wien

This research program shall bring together research groups from Austria that have been, still are and intend to be active in the research of image processing and pattern recognition. The program also includes both theory and application (see Table 0.4) within the individual projects. We have experienced the difficulties of cooperation e.g. during the preparation of this program, but we also know (now) that the synergistic effort enhances the potential of the whole group.

Table 0.4: Applications of the Research Program

• . . . will be realized,
○ . . . is possible.

Application	Project			
	I	II	III	IV
Medicine	○		•	
Office Automation	○	•		
Industrial Automation	○	○	○	○
Industrial Manufacturing				○
Quality Control	○			○
Surveying			○	•
Remote Sensing	•	○	•	•
Astronomy/Geophysics	•			
Navigation			•	

Expertise

The following list displays the interdisciplinary character of the research program. It includes all the scientific disciplines from which expertise is represented in at least one of the participating groups:

- Computer Graphics
- Computer Science
- Image Processing
- Image Understanding

- Medical Image Analysis
- Operations Research
- Mathematics
- Pattern Recognition
- Photogrammetry
- Remote Sensing
- Statistics
- Surveying
- System Sciences

Potential applications and relations to industry

Table 0.4 summarizes those applications that are directly related to the proposed projects. But there are many other areas where digital images may facilitate a production process or even where they make a production possible (e.g. in dangerous environments). It should be mentioned that the major goal of the research program is the advancement in methodology, and that most applications in this program should validate the proposed methods, they should point out the range and the limits of applicability rather than a single solution to a specific problem.

The table distinguishes between applications that are planned to be realized in a prototype system (•) and applications to which the results could be easily transferred (◦).

It should be mentioned here that close contacts with industry already exist. They are documented for each individual project in the sections 'cooperation with industry'. Representatives from industry were also present at the FWF-workshop on December 11, 1989 and have expressed their support for the program. Progress meetings and the planned workshops of the research program shall encourage the mutual information exchange with industry (like it is done during the "journées de travail" at GDR TDSI in France). Hence an accompanying industry commission should be formed. Such a commission could provide information about the actual needs of industry and, on the other hand, it could match the solutions offered by the research program with the problems at hand.

Chapter 0

Coordination

Project coordinator:

o. Univ. Prof. D.I. Dr. Walter G. Kropatsch,
Department of Pattern Recognition and Image Processing (PRIP),
Technical University, Vienna.

0.1 Problem Statement

12 research groups in the 3 major University cities of Austria propose to cooperate in this research program. The general management of such a large and geographically disjoint group has to provide first of all a smooth flow of information among the four projects of the program. But on the other hand it must coordinate the contacts to the research foundation (FWF), to industry, to the media, and to the international research community.

During this project a lot of algorithms will be developed and a lot of different data (images) will be used. It is therefore an important part of this project to provide the fundamental organizational and technical prerequisites to allow for this multi-directional flow of information, including conventions for the exchange of data and software. Consequently, a considerable part of this project will be devoted to an informal *standardization effort* as a separate project focus.

Standardized Image Processing

Many solutions in Computer Vision make use of a limited set of similar functional components. A basic system that efficiently supports this common functionality would dramatically reduce the individual development efforts and allow to put more emphasis on innovative research.

There are several groups in Austria that have developed in the past nearly independent of each other knowhow and software in the areas of image acquisition, image processing, and image interpretation. Two reasons motivate the proposed standardization in the area of image processing: Image analysis receives increasing interest from the commercial applications while more and more software is produced on special systems and hardware (not only in Austria). Hence the need to search for possibilities that would allow to build the new programs on existing modular systems and on well defined software interfaces, and to exchange diverse software components among different hardware environments. This could produce a much greater efficiency in research and development because time consuming redesign and rewriting of software can be replaced by copying and adapting the transportable components. Since real images are very noisy data in general with many sources for disturbances the selection of methods is often based on trial and error. Therefore the demand to build a large toolbox of programs from which the appropriate method can be selected (ref. to task I.1).

Recently progress has been made to standardize low-level image processing and interchange (c.f. ISO/IEC JTC1/SC24/WG1 DIS 12087-2) Image Processing and Interchange Part 1, 2 & 3). It is forseeable that in the next one or two years an international standard for image processing and interchange will be established. For Austria it is very importnat to actively participate in this international standard.

0.2 Proposed Tasks of the Coordination Project

In particular, the coordination has to take over three major tasks:

- Standardized Image processing
- Coordination within research program
- Representing the program

0.2.1 Standardized Image Processing

The tasks in this part are tightly coupled to the part of the coordination within the research program (0.2.2). The idea is to document the software developed in this research program in accordance with the international standard. It is therefore necessary to continue the work on the standard to keep in touch with new developments. As a major goal of the coordination project we will provide the necessary infrastructure to exchange data and software between the researchers of this research program. For the definition of the interfaces we will stick as closely as possible to the international standard. This gives us the opportunity to extend the standard with new functions with a minimal effort. If software(hardware) implementations of the standard become available, the new routines can be easily included.

The main activities and goals within this part are:

1. Continue the work on the international standard (ISO/IEC JTC1/SC24) in the ÖNORM-working group
2. Provide the partners of the reasearch program with the necessary information about developments of the standard
3. Create and maintain an ftp-server for the exchange of software and data
4. Define (and use) a common image format and provide the necessary conversion routines in order to build an image libary
5. Building-up and maintaining a catalog of:
    - individual problems,
    - methods for solutions,
    - available equipment,
    - other research groups etc.
6. A literature data base related to the topic of the program would be beneficial to all the interested groups in Austria.

All this information will be made available by the ftp-server to all partners of the research program.

0.2.2 Coordination within Research Program

Task 1: Provide the means for efficient e-mail communication

Task 2: Correspondence and circular letters.

Task 3: Organisation of visits at the participating laboratories.

Task 4: Organisation of meetings of the project groups.

Task 5: Organisation and preparation of progress meetings of all participants in the research program.

Task 6: Support for exchange of experiences between the three scientific projects in general.

0.2.3 Representing the Program

Task 7: General coordination with and reporting to the FWF.

Task 8: Contacts with industry.

Task 9: Public relations.

Task 10: Coordination with international projects and research programs.

Chapter 1

Mathematical Methods and Tools for Digital Image Processing and Pattern Recognition

Project coordinator:

o.Prof. Dr. Ing. Franz Pichler,
Department of Systems Science,
University of Linz.

1.0 General Problem Statement

Mathematical methods have a long tradition in the field of Signal Processing, especially in Digital Image Processing and Pattern Recognition. The advancement of computer technology in hard- and software gives the possibility to extend standard methods and to develop new methods based on new mathematical concepts. It also permits the use of open computing architectures where previously only especially built hardware could do a task. The cooperating groups which perform Project I cover by their specific expertise a rather broad area of mathematical methods and algorithmic skills. Each group has selected a specific research area in which it has already achieved competitive expertise.

However, since all groups are concerned with applications in DIP it will be very natural that there is a close cooperation among them to exchange knowledge and to stimulate individual research. Furthermore they can give help to other groups of the overall research in DSP in specific problems dealing with mathematical methods and their tool integration.

Since each individual group is also engaged in graduate teaching the research project will also stimulate teaching and finally contribute to the improvement of courses in Methods of Digital Image Processing and pattern Recognition as they are offered to students in Applied Mathematics and Computer Science at Austrian Universities.

1.1 CAST.FOURIER and Applications

Task leader:

o.Prof. Dr. Ing. Franz Pichler,
Department of Systems Science,
University of Linz.

1.1.1 Problem Statement

Introduction

Computer Aided Systems Theory (CAST) is an effort of systems scientists to provide various engineering disciplines with software implementations of concepts and methods based on systems theory. One goal of CAST research is to make results of systems theory applicable to modern engineering. Systems theory approaches and techniques are best being offered as convenient and versatile interactive method banks, preferably implemented on the same high-performance graphics workstations which are used by the engineer's CAD tools .

CAST systems have to support modelling activities (in terms of systems theory concepts), to offer a collection of methods, and to enable users to interactively and, to some extent, exploratively navigate through the problem space, i.e., to apply and to concatenate methods until acceptable solutions are found. In this setting, it is up to the user to decide which methods are applied and how they are combined to finally yield a solution for a given problem.

In the past a powerful prototype of a CAST system (CAST.FSM, CAST for the case of Finite State Machines) has been developed. CAST.FSM is a method bank which allows to experiment with concepts and methods from automata theory. Areas of application which can benefit from this system are digital circuit design, in particular design for testability, cryptology, and systems theory education.

It is subject of this project to extend the method bank of CAST.FOURIER and to evaluate potential applications of new algorithms. CAST.FOURIER will serve as a tool for investigating and using (generalized) spectral techniques in the areas of multidimensional signal processing. Motivation, goals, concepts, state of implementation, and the research problems of the project will be described below. We also will point out the state of the art as known to us by our international contacts, promoted by special workshops which we organized (EUROCAST 89 Las Palmas, EUROCAST91 Krems, EUROCAST 93 Las Palmas) and by exchange of research results between experts in the field.

Motivation

Two principal reasons motivate the development of a CAST tool for digital image processing (DIP).

1. In DIP, many applications require a highly interactive and experimental problem-solving strategy. One basic problem in DIP is to improve pictorial information for human inspection and interpretation. Human guidance is essential in selecting the appropriate enhancement and restoration techniques to obtain images of satisfying quality. A second class of problems deal with preprocessing data for autonomous intelligent machine perception (computer vision). The results of this processing step are often crucial for the success of the entire vision task. Therefor, the designer of a pre- processing algorithm has to carefully select, analyze and combine suitable methods. In this process, he may have to go through a number of experiments and iterations until the right computational subtasks are founded. This interactive style of tackling problems also calls for powerful and versatile software support. It all corresponds to the methodology used in CAST systems indicated above.

2. DIP has in its mathematical foundations a long tradition with systems theory. Examples are design and analysis of digital filters, spectral transforms, and related frequency-domain techniques. There is also a number of tools which make these well-established areas available for practical applications. Generalized spectral techniques, however, can be considered a new branch of systems theory (or applied mathematics) which has not yet generally been introduced as a method to solve DIP problems (although back in the 70s rather wellknown scientists like Andrews and Welch showed strong interests). This is partially due to the lack of convincing demonstrations of the usefulness of these approaches, but certainly also due to the lack of appropriate software tools to develop appropriate algorithms and to apply these techniques in practice. CAST.FOURIER is designated to fill this gap.

It should be noted that our system is not only intended to enhance the toolbox of a scientist faced with DIP problems, i.e., to "instrument" problem-solving, but also to serve as a research tool which will help to gain insight into and experience about (generalized) spectral methods. As will be discussed in more detail below, knowledge about these novel approaches is limited, and a researcher exploring possible applications must have an appropriate tool at hand. Our research has the goal to show how CAST.FOURIER is successfully applied to reach new algorithms in DIP.

State of the Art and available installation in CAST.FOURIER

Most of the spectral techniques developed in DIP are based on the "classical" discrete Fourier transform (DFT). Frequency-domain methods are well-established in DIP due to their well-known theoretical background, the obvious interpretation of the spectral domain in terms of spatial frequencies, and the availability of the fast Fourier transform (FFT). Examples of applications are filter design and filtering, image restoration, and image encoding (data compression) methods.

Another significant transform is the Walsh-Hadamard transform (WHT) which is perhaps the most well-known of the nonsinusoidal orthogonal transforms. Many theoretical results have been derived, fast algorithms have been developed, and the concept of sequency has been suggested for interpreting signals in the spectral domain. Since, in addition, the WHT can essentially be computed using additions and subtractions only and, hence, is more desirable than the DFT from a computational point of view, it has gained some prominence in various applications. Examples include image compression, image enhancement as well as pattern recognition tasks. Various modifications of the WHT are known.

Besides these two notable classes of spectral techniques, a variety of more specific transforms and related "spectral-domain" methods are implemented in CAST.FOURIER. They are listed in the sequel. For theoretical details, we refer to the existing literature [AR75, Bea84, GW87, W.K78].

- Discrete cosine transform (DCT). The DCT closely approximates the KLT in a statistical sense and, hence, is often used to replace the KLT.
- Discrete sine transform (DST)
- Haar transform (HT)
- Slant transform (ST).
- Discrete Hartley transform (DHT), which has been proposed as a substitute for the DFT for some purposes [Bra84].

The implementation of these transforms in CAST.FOURIER is described in detail by the master thesis of Kastner [Kas90].

Furthermore it should be mentioned, that the DFT has been generalized in different respects. This effort has yielded, for instance, number-theoretic transforms and polynomial transforms [Nus82]. The generalized discrete Fourier transform GDFT as already implemented in CAST.FOURIER [Hel89, Kun77, PS90] is introduced in the following in more detail.

The basic notion for the GDFT is to consider complex-valued (or real-valued) signals x defined on a finite Abelian group G, that is,

$$\mathrm{x} : G \rightarrow \mathbf{C}(\mathrm{or}\,\mathbf{R}).$$

It is well-known from algebra that any finite Abelian group can isomorphically be represented by a direct product of cyclic groups $\mathbf{Z}_{\mathrm{N_k}}$ of prime order power $\mathrm{N_k} = \mathrm{p_k}^{\mathrm{a_k}}$. Hence, the discussion can be restricted to the following case:

$$G = \mathbf{Z}_{\mathrm{N}_1} \times \mathbf{Z}_{\mathrm{N}_2} \times \cdots \times \mathbf{Z}_{\mathrm{N}_m}, \text{ and, } \mathrm{N_1 N_2 ... N_m} = \mathrm{N} = \mid G \mid$$

Then, the GDFT, in onedimensional representation, is a linear, orthogonal transform defined by the matrix operation

$$\mathbf{y} = \mathbf{V}_{\mathrm{N_1 N_2 ... N_m}}\, \mathbf{x}.$$

where $\mathbf{x}$ and $\mathbf{y}$ denote the signal and spectral vectors of length N, respectively, and the N $\times$ N-GDFT matrix $\mathbf{V}_{\mathrm{N_1 N_2 ... N_m}}$ is obtained as the Kronecker product of the individual Fourier matrices $\mathbf{U}_{\mathrm{N_k}}$ belonging to the cyclic groups $\mathbf{Z}_{\mathrm{N_k}}$:

$$\mathbf{V}_{\mathrm{N_1 N_2 ... N_m}} = \mathbf{U}_{\mathrm{N_1}} \otimes \mathbf{U}_{\mathrm{N_2}} \otimes ... \otimes \mathbf{U}_{\mathrm{N_m}}.$$

It is easily seen that the GDFT denotes a *family* of orthogonal transforms where an individual transform is determined by selecting a specific group G. There are two notable special cases. If the group G is chosen to be the cyclic group of order N, $G = \mathbf{Z}_{\mathrm{N}}$, the transform obtained coincides with the standard (cyclic) DFT. On the other hand, if $\mathrm{N} = 2^{\mathrm{n}}$ then the underlying group is the so-called dyadic group of order n, i.e., it is the direct product of n cyclic groups of order two, $G = \mathbf{Z}_2 \times \mathbf{Z}_2 \times ... \times \mathbf{Z}_2$ (n times), and we end up with the Walsh- Hadamard transform. It has been shown that, if $\mathrm{N} = 2^{\mathrm{n}}$, the GDFT comprises N/2 different transforms [Hel89] where the limiting cases are the DFT and WHT, respectively, and the other transforms enable a systematic transition from the DFT to the WHT. This situation is depicted in Fig. 1.1.

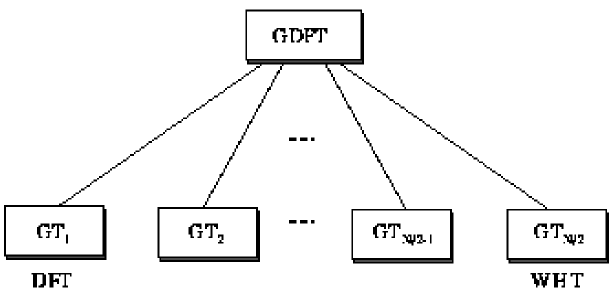


Figure 1.1: GDFT - a family of "general" orthogonal transforms

Generalizations of the DFT have also been proposed by [AR75] and by [BCF76]. The approaches chosen by these researchers are quite different from the one outlined above. However, they also reach a family or class of linear transforms where the DFT or both the DFT and the WHT occur as special cases.

A rigorous algebraic approach, even more general than the one adopted by the GDFT above, was first taken by [Nic71]. Generalizations of the standard DFT both with respect to the underlying group and with respect to the number system which is used for the signal values were developed. The underlying group G is there assumed to be an arbitrary finite Abelian group. The number system is no more restricted to the complex numbers $\mathbf{C}$ but is assumed to be a commutative integral domain.

This direction of generalization was further explored by other researchers, notably by [Bet84], who proposed the most general Fourier transforms in that, where, for example, also non-Abelian finite groups G are considered.

In summary, there is a remarkable variety of (generalized) spectral techniques and, correspondingly, an extensive body of theoretical knowledge. However, it is our belief that know-how about the practical significance, benefits and drawbacks, and potential applications of these transforms is not equally well developed or, at least, not wide-spread.

As an example, consider the GDFT on finite Abelian groups in the sense outlined above. Many of the desirable properties of the DFT also hold for the other transforms of the class, e.g., linearity, (generalized) convolution property, straightforward inverse transform, fast algorithm. In addition, a "general" transform can offer advantages over the DFT in that, for example, it may be computationally less expensive and, simultaneously, better approximate the optimum transform (KLT) [Kun77]. It was proposed, therefore, that a "general" transform can advantageously replace the DFT or WHT in the domains of image compression and generalized filtering [Pic80, Pic81].

However, this approach has, up to now, never been shown to operate in practice, nor have definite guidelines been worked out which help to select the transform best suited for a given problem.

It is therefor the principal purpose of CAST.FOURIER to improve this situation, that is, to provide users as well as researchers with a tool which enables them to apply and investigate (generalized)spectral methods. In the above setting of the GDFT, the CAST system should, for instance, aid in defining criteria and in finding a systematic way to determine the "best" transform for a specific application problem.

State of Development of CAST.FOURIER

The implementation of CAST.FOURIER has already reached a certain stage of development. The system in its current version heavily relies on the Vision Kernel System (VKS), a research tool which has been developed at our department to support interactive machine vision programming [Bur88, Bur89]. VKS features basic functions for archiving and displaying images on a general-purpose color graphics workstation, for manipulating images and for performing primitive image operations. VKS is being implemented on UNIX workstations. It provides a versatile programming interface based upon Common Lisp (CL) and the object-oriented programming system CLOS (Common Lisp Object System). Computationally intensive routines are coded in C and integrated into the CL environment. The underlying concepts and details of VKS are reported elsewhere [Bur88, Bur89].

CAST.FOURIER can be considered as modular extension of VKS; it is based on the same concepts and makes use primarily of the VKS display and archive facilities. It comprises all the methods which are usually associated with (generalized) spectral- domain approaches and which not necessarily have to be seen by the VKS users or developers. Examples are the GDFT (GFFT), various scale and display functions for (generalized) spectral images, miscellaneous other transforms, and the methods that are currently being developed to test the GDFT in an actual DIP task. The overall structure of the software is conceptually depicted in Fig. 1.2.

CAST systems are preferably implemented using the object-oriented programming paradigm. Systems types and systems, which in our case are several sorts of images, for instance, map naturally to classes and instances, respectively. System transformations, e.g., spectral transforms, are described by generic functions and accompanying methods.

The CL-based OOPS represents a (nearly-)persistent interactive programming environment that meets most of the requirements of an "interactive systems theory method bank" as proposed by [PS90]. Objects (e.g., images, filters, windows) are permanently available and constitute the data base of the CAST system. A number of methods (e.g., spectral transforms) is at the user's disposal and can be applied and combined arbitrarily (as long as this makes sense from a mathematical or application point of view). In this way, the user can perform experiments and exploratively compose a solution for his problem. If

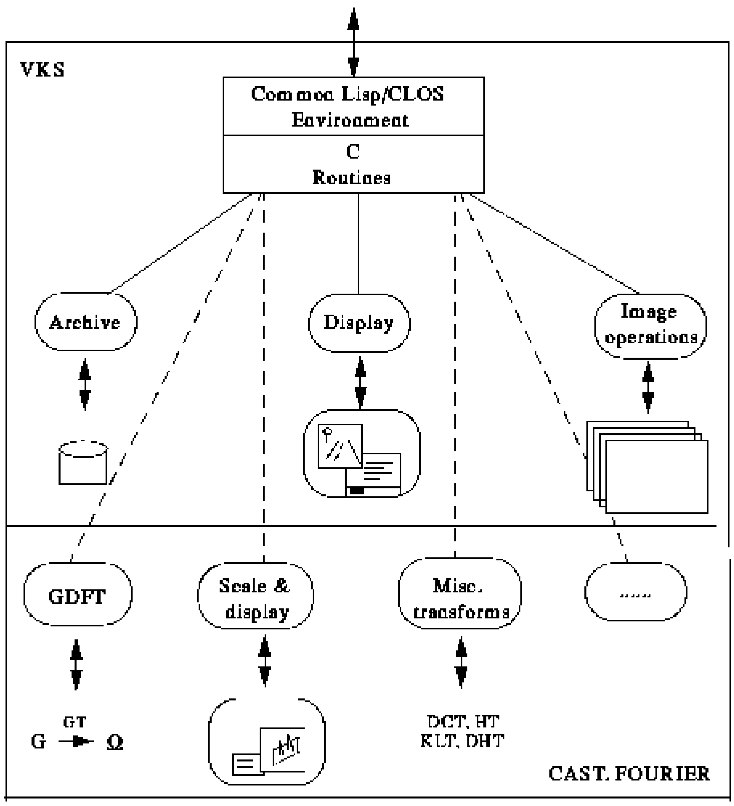


Figure 1.2: VKS/CAST.FOURIER Structure

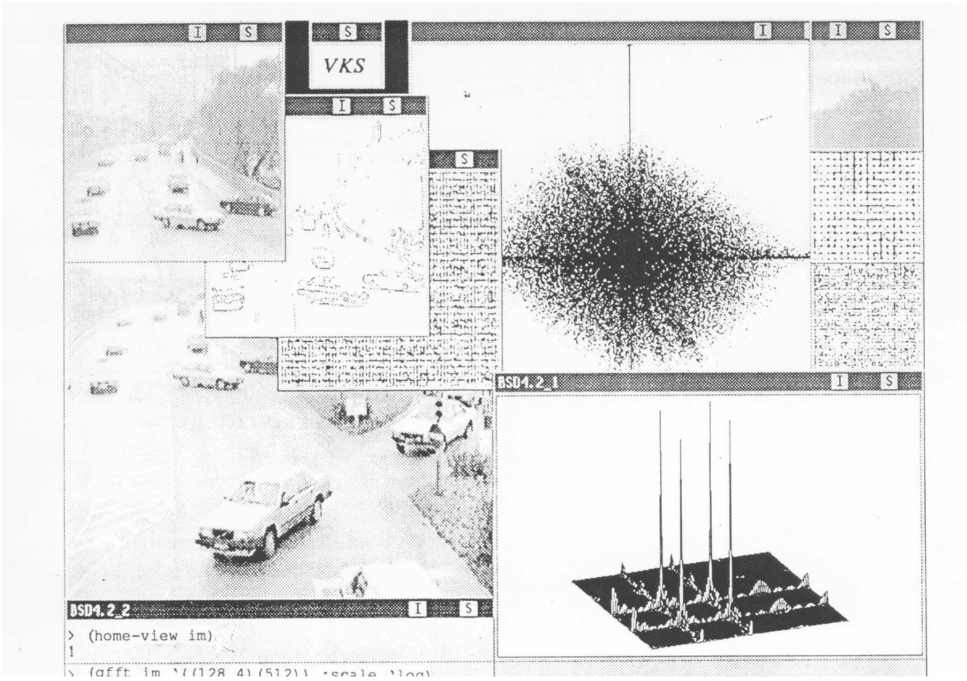


Figure 1.3: Typical VKS/CAST.FOURIER Screen

required, object definitions and objects as well as functions (methods) can be added at any time and made part of the system for future use. In summary, the programming environment alone favorably matches to the experimental problem-solving strategy proposed by the Systems Theory instrumented problem solving methodology.

Fig. 1.3 shows some of the currently existing facilities of the CAST.FOURIER system. Some images and the results of different "general" transforms of these images are displayed in several modes and windows. The text window represents the CL listener which, at present, is the primary interface to the user or programmer.

It must be noted that, at the current state of development, CAST.FOURIER must be considered rather a collection of algorithms (a program package) than a mature and coherent "method bank". The textual man/machine interface is still somehow inconvenient and must be replaced in the future by a state-of-the-art graphical user interface as represented by the browsers in CAST.FSM, for instance. The system does not keep track of the system transformations and does not record the state of the problem-solving process as do the so-called realization trees in CAST.FSM. Hence, it is the user's responsibility to bind and remember intermediate results so that they can be referenced in further processing, e.g., by typing

```
(setq spectral-image (fft image)).
```

However, all the facilities to provide these convenient features are available in the CL and OOPS environment. It is our plan of research to realize these important components of a CAST system in our tool since its acceptance and success may crucially depend on these features. In addition, the potential of Lisp offers the opportunity to build a sort of "intelligent assistant" on top of the CAST system to make the knowledge gathered about (generalized) spectral techniques available for the average user.

In the meantime (1991) VKS development has been stopped in favor of the DIP environment KBVision and CAST.FOURIER has to be integrated with this DIP tool.

1.1.2 Research Goals and Methods

CAST.FOURIER extension and explorative use

Although many of the goals of our proposed research project have become already obvious from the previous discussion, they are briefly summarized here. They can be roughly divided into short-term goals and long-term goals.

The primary short-term goal is to thoroughly investigate the GDFT described above. First, a deeper understanding of the "general" transforms must be gained and reasonable interpretations be found. (The concept of "generalized frequency" that has been proposed [AR75] is very abstract and difficult to comprehend.) Secondly, experiments should be performed to explore potential applications of the GDFT in DIP. Comparison with established methods will reveal whether the above hypothesis holds that one or more "general" transforms can outperform the standard DFT (or WHT). If so, the final attempt will be to establish guidelines which direct the user in selecting a transform well suited for a given problem (application).

The long-term goal of the project is to make CAST.FOURIER a versatile software environment which enables the user/researcher to conveniently apply/investigate the most important (generalized) spectral techniques in DIP problems. The user should be in a position to learn about available techniques, to perform experiments, to explore new application areas, to develop new methods based upon available spectral transforms, and to solve certain DIP problems.

To this end, other generalizations of the DFT will have to be implemented and investigations similar to the ones described for the GDFT will have to be performed. For test, comparison and evaluation purposes, standard frequency-domain as well as corresponding spatial-domain methods will have to be implemented as well. Furthermore, in order to increase the performance, a software component which provides the user with some kind of assistance and guidance will have to be added. Clearly, it is desirable that a CAST system can offer advice and is able to support the user in selecting suitable methods. This demand is especially stringent in a domain like the role of generalized spectral techniques in DIP where the user is not expected to be an expert.

The 2D Generalized Fast Fourier Transform (GFFT), a fast algorithm for the 2D GDFT, and corresponding methods for scaling and displaying (generalized) spectral images have already been implemented by us. As principal further investigations for CAST. FOURIER, three directions should be followed:

1. The GDFT investigated in a 1D context. This work is to provide insight into and yield a comprehensible interpretation of "generalized spectra".

2. Case studies should be performed which test the usefulness of the GDFT in specific DIP problems. Image compression seems here to be the most promising application area. Results obtained by means of the GDFT - more precisely, selected "general" transforms - are compared with and evaluated against established spectral techniques and/or corresponding spatial-domain methods. Since it is computationally infeasible to explore the entire class of transforms which the GDFT represents, the experiments are being performed with specific "general" transforms. Selecting representative transforms poses a major problem because, as stated above, our knowledge about the properties of "general" transforms is limited. Again, studying the principles of the GDFT is required.

3. In order to extend the experiments we will use also other spectral transforms which are implemented in CAST.FOURIER. These include the generalized transform as defined by [AR75], the DCT, DST, HT, DHT, as well as the Karhunen- Loeve Transform.

In the course of these developments and experiments, the system is permanently improved. So, first applications and enhancements of the tool go closely together.

Development of new algorithms with CAST.FOURIER

By the availability of an operational prototype of CAST.FOURIER the following development of specific algorithms should be considered as part of our research.

Wiener optimal filters for images over groups To investigate the selection of the proper group structure in the spatial domain to obtain the "best" WIENER filtering for noise reduction or prediction in the case of a given digital image disturbed by additive noise.

The theory of optimal Wiener filtering in the general sense has been developed by Karpovsky and Trachtenberg [Kar77].

Earlier fundamental contributions for the case of the Walsh-Hadamard transform have been made by Pichler [Pic70] and Pearl [Pea75]. Computational aspects can be found in [Pra72]. The efficient proper selection of a generalized Wiener filter which is "best" in the sense, that it fulfills a certain performance criteria [Tra85, KE86] is a desirable goal. CAST.FOURIER which allows the user to perform efficient experiments with images and image-transformations can be an important tool for reaching this goal. Important applications in the area of digital image processing (e.g. [Tra86]) seem to be feasible.

Development of special 2-D filtering methods for images

(a) "ROTH Multiplicity Filtering" To investigate the approximation properties of different kind of ROTH multiplicity filters in digital image processing with the emphasis on approximation of polynomial images.

The concept of multiplicity (in german "Vielfalt") plays a very important role in generalized Fourier analysis. As Weisz [Wei67] pointed out first, the generalized Fourier expansion of a polynomial of degree n has the property that all Fourier coefficients a(k) which have a "Rademacher-multiplicity" k¿n are zero (k is defined to be equal to the number of Rademacher-functions which are necessary to generate the character function ch(k) associated with a(k)). Roth [Rot73] investigated in his research linear dyadic filters with respect to the multiplicity concept. In the master thesis, performed at our institute [Goe90] the properties of 2-D multiplicity filters which approximate piecewise polynomial images have been investigated using CAST.FOURIER. Furthermore an extension of the multiplicity theorem of Liedl [Lie64] for the dyadic case has been derived. It should be mentioned here, that the hypothesis for the existence of such a theorem, was achieved by results derived by CAST.FOURIER experiments. Our research should continue with explorations of similar kind with the goal to get a full understanding of ROTH multiplicity filtering in the general sense of the Fourier-analysis on finite groups. It is hoped that results will help to find efficient methods for image-coding and image restauration for polynomial-like images.

(b) "GIBBS Derivatives in Digital Image Processing" To investigate the effects of Gibbs-differentiation on digital images and to explore its relation to other image-trans-formations.

The generalized derivative of Gibbs [BW72] has recently found new attention in applied mathematics [BE89]. With help of CAST.FOURIER it was possible to explore some of the properties of Gibbs derivation for digital images [Pic89a, Goe90]. Further experimental research with CAST.FOURIER should explore the relation of Gibbs derivation and Roth multiplicity. Furthermore it would be desirable to investigate the general role of Gibbs differential operators and Gibbs differential equation systems in digital image processing.

Prigogine transformation of baker-sequences of images The baker-transformation is a mixing measure-preserving bijective transformation for 2-D images defined on pixel-level. In our recent research [Pic92] we showed, that Walsh-Fourier analysis is the proper approach to study computational features of the baker transformations. We have implemented it as a modul of our digital image processing tool

CAST.FOURIER and of the tool KBVision [PS93]. This implementation allows us to visualize results and to apply - in the case of KBVision - common digital image processing and vision methods and also more specific methods, as they are available in CAST.FOURIER.

Prigogine et al. point out in their work that a baker-dynamics is a specific example of a deterministic conservative dynamical system, which allows the construction of a family of associated stochastic dissipative dynamical systems (very specific Kolmogoroff systems).

Furthermore, Prigogine et al. give the concept of the Λ-transform, which realizes this construction by state-assignment. In the report [Pic92] it is been shown, that Prigogines Λ-transform can be considered equivalent to the concept of a dyadic convolution operator, as it is known in Walsh-Fourier analysis. In earlier papers such operators have been interpreted as linear I/O systems which are invariant with respect to dyadic shifts. Our interpretation allows to characterize the requirements on Λ in terms of the transfer function of a dyadic- invariant system.

The irreversibility of the dynamical system assigned by Λ is based on the facts that for every state (as described by a distribution function on the unit-square) the states of the associated state-trajectory have an exponential increasing demand on bandwith (as defined for signals by Walsh-Fourieranalysis) and inversion needs an exponential grow of the transfer-function. We already have implemented the Λ-transform and the associated baker- Prigogine dynamics and we have integrated them as moduls into CAST.FOURIER and KBVision. Further research should focus on possible generalizations of the baker-transformation and the related dynamical systems. Special attention will be given to appropriate visualisation methods. On basis of this results of mathematical systems theory and by our CAST.FOURIER implementations of this results we want to investigate different possible applications in DIP, such as applications in image scrambling and image coding. However, also applications which are in line with Prigogines motivations to study phenomena of living systems behavior by basing such studies on deterministic machines on a molecular level seem to be promising. Certainly the Baker-Prigogine dynamics will serve here only as (mathematical working) example and other similar constructions of dynamical systems will have to be investigated. Different applications of such rather new methods in DIP in biological and medical fields could be investigated. CAST.FOURIER, if properly extended to such specific spectral representations will certainly serve as an important tool for such research.

A co-operation in research with an austrian medical center is planned.

Variable receptive field transform: applications in image-scrambling By the work of R. Moreno-Diaz and his group (based on earlier work on neural nets, PhD thesis at MIT with Prof. W. McCulloch) there exists a class of digital image transformations (linear and non-linear) which have been so far mainly be investigated for applications in robotics and in quality control of industrial processes [CMBB90, RMD90]. The more recently developed nonlinear receptive field transform (RFT) - as reported in the PhD thesis of O. Bolivar-Toledo - requires the use of a large number of parameters, which can be used

(a) to tune on image to a semi-intelligent image (as e.g. wanted in TV scrambling) and

(b) to serve as a cryptographic key.

Since the mathematics of these transform is based on nonorthogonality this applied research would allow - together with other considerations and planned cooperations within the overall DIP research project - to compare the methods derived by CAST.FOURIER with the RFT methods. By the existing contacts with the group in Las Palmas it would be naturally to base the research on a cooperation with this group.

1.2 Statistical Methods for Pattern Recognition

Task leader:

Univ. Prof. Mag. Dr. Georg Pflug,
Department of Statistics, Operations Research and Computer Science,
University of Vienna.

1.2.1 Problem Statement

Pattern recognition deals with extracting information from images. Often this information is corrupted by random noise. Therefore pattern recognition may be viewed as a statistical estimation problem, where the signal has to be estimated and separated from noise. Since an image is mathematically idealized by a function from $[0,1] \times [0,1]$ to a continuous scale (e.g. continuous gray scale), this estimation problem is of infinite dimension (with respect to the number of parameters) and even if the discretization in resolution is considered, its dimension is beyond feasibility. Therefore the application of statistical methods is classically done in two steps.

In the first step a low dimensional vector of features is selected from the image and in the second step this vector is handled by methods of multivariate statistics, such as discriminant analysis, principal component analysis, cluster analysis, nonparametric density estimation et cetera.

The recent progress in theoretical statistics and in the theory of stochastic processes makes it however possible to handle infinite dimensional models and it is one aim of the research project to incorporate results from the modern theory of stochastic processes (which deals with either infinite dimensional or at least very high dimensional objects) into applicable methods for pattern recognition.

Statistical pattern recognition is based on statistical models of observed images. The part of theory handling with statistical modelling of image processes is called *pattern synthesis* ([Gre76b]). Statistical models are needed for:

(I) Classification of patterns by reversing their generation process.

(II) The noise process which corrupts the image.

(III) Design of optimal sampling sets.

(IV) Standard methods for classification and cluster analysis

Models for (I) are based on a stochastic theory of shapes ('deformable templates'). A random shape is an equivalence class of random sets in $\mathbb{R}^2$, the equivalence is with respect to shape invariant transformations (translation, rotation, scalar multiplication).

Shape models may be found in the books by Grenander ([Gre76b, Gre77, GK91]), Harding and Kendall ([HK74]), Matheron ([Mat75]), Serra ([Ser82]). Some of the possible approaches to pattern recognition via shapes are listed below.

1. The theory of random sets. It deals with random distributions on the appropriately metricized space of compact sets in $\mathbb{R}^2$ ([SW81]). Applying special adapted random processes to those sets (i.e. to the shape) the shape generating mechanisms can be modelled ([GK86, Ken77, Kno88, Osb86, Eve85, HP80]). This generating mechanism helps to develop algorithms for pattern recognition by reversing the generating process, i.e., we are looking for optimal parameters of the underlying stochastic process which are then tested against the expected ones, given by the generating process.

2. Mathematical morphology. A theory of operators on sets, (such as dilation, translation, modulation, erosion, openings, closings etc.) applied to random sets ([Mat75, Ser82]).

3. The theory of *connector graphs* ([AGP91, CGK91, GM92, GK91]). Connector graphs are graphs whose vertices correspond to points in $\mathbb{R}^2$ and whose edges model relations between these vertices. A random structure is a probability distribution on a transformation group, which acts on this graph.

4. Time-series, change-in-angle characterization of shapes. Instead of modelling the set as an object in $\mathbb{R}^2$, this method only considers the change-in-angle function of the closed contour of a set as a periodic function, which may be analyzed by time-series methods ([ZR72, PF77, KS89, KC81]).

   Among these time-series methods are methods based on series expansions which are both, very easy to interpret and highly adaptable to the problem. Little attention was so far paid to wavelet expansions of the contour function. Donoho ([Don92b]) has shown that wavelets have the specific feature to adapt to local scales, in the sense that the degree of smoothing by a wavelet approximation depends on the local behaviour of the data. This property will be investigated for the purpose of pattern recognition.

5. Random transformations of the plane. There is increasing interest in this transformations, which transform an ideal template into real observable images. This approach was considered by Gabe ([Gab88]), Amit ([AGP91]), and Piccioni ([BP92]) and will be developed in the project. These stochastic deformations of the plane can be used to deal with biologically variable shapes, such as the shapes of leaves or handwritten signatures ([GK86, PCUG93, FPC+93, Ken77, AGP91]).

   Another application of these transformations is pattern matching, e.g., when two different images with different distortions of the same object have to be matched.

Models for (II) are random processes with index set $\mathbb{R}^2$, which are homogeneous in space, i.e., either two dimensional stationary or Markovian with respect to a clique structure in the plane (Gibbs processes) ([GG83, GG84, Der79, WD86, KS80, WDM87, ZVC89]).

Statistical models for (III) have to be investigated; with imaging of natural scenes you never know where to put your sampling points. Recent work is done by Hans G. Feichtinger and his research team in the theory of non-uniform sampling but no theoretical models on the choice of the sampling points have been introduced so far. Optimal sampling is of value for image analysis and image synthesis with respect to optimal reconstruction and data compression.

The methods in (IV) which are used for pattern classification (Bayes discriminant analysis, et cetera) and clustering are well known [Fuk90]. These problems are all used by the image processing and pattern recognition community. We introduce new methods using wavelet analysis (cf. cites below) and methods of operations research. Although there is a lot of commercial and non-commercial software available dealing with different algorithms, there is so far no method database where one can get help on statistical problems arising in this research area.

1.2.2 Research Goals and Methods

The goal of this research topic is the theoretical and experimental comparison of different (new) methods in the theory of stochastic processes and statistics. The recent progress in hard- and software has enabled us to implement very complex algorithms which are used to test the theoretical results in experiment and in concrete applications, such as offline signature verification (sponsored research by the Austrian Nationalbank, [PCUG93]). Furthermore computational experiments help us to find new results, even before any stable theory exists.

Therefore it is planned to setup a statistical laboratory for image processing and pattern recognition (StatLab) consisting of an UNIX workstation and a PC, in order to have the two main platforms for software development available (in order to use international standard freeware for academic research) and to be able to use the existing soft- and hardware and to easily share it with the other research groups of the program (program packages for statistics and mathematics, available at the Department of Statistics, and via InterNet at the other sites of the research program, furthermore the existing input and output devices).

In this StatLab a library of statistical methods in digital image processing and pattern recognition should be collected and developed, with the main emphasis to our area of interest (cf. research topics below). This *Statistical Image Processing Library*(STIPL) will be extended to a method database, which can be shared with the other groups of the research program via the Austrian Academic Net at any time (using the workstation as server). Thus a knowledge database of statistical methods for pattern recognition and image processing with concentration to our main research topics (see below) should be available to all researchers in Austria by the end of the research program.

The application of statistics to the recognition and classification of biologically variable forms, such as classification of trees by the shape of their leaves, or, recognition of persons by the shape of their hands or by their fingerprints, will be one of the main research areas of our team ([GK86, Amm91, Amm89a, Bro85, Ede68, Spr73, Boo78, PCUG93, FPC+93]). The proposed algorithms decrease the possibility of missclassifications introduced by using standard methods only.

The main research topics are

- T1 Application of the theory of deformable templates
- T2 Contour line recognition by means of orthogonal and non-orthogonal series expansions (wavelet approximations)
- T3 Parameter free classification and discrimination using wavelets
- T4 Experimental design methods for optimal sampling
- T5 STIPL – a statistical image processing library and knowledge base

These topics will be dealt with by developing the proper theory, implementing and testing of the algorithms in experimental setting (StatLab) and, at last, by field experiments, i.e., by application of the developed theory to useful problems in industry, such as signature verification and recognition and classification of templates by their shape (with applications to machine vision).

Application of the theory of deformable templates

A deformable template is a set S in $\mathbb{R}^2$ together with a stochastic process mapping $T : \mathbb{R}^2 \to \mathbb{R}^2$ ([AGP91, GM92, GK91, BP92]). The 'deformed' set $S^T = \{T(x)|x \in S\}$ is a random set. Suppose we start with k sets, say $S_1, S_2, \ldots S_k$ and observe a set $\widetilde{S}$. $\widetilde{S}$ may be the deformation of any of the sets S_1 to S_k, but which 'original' set is the most likely ?

This question leads to the so-called maximum-likelihood estimate under partial information (MLPI). Since we are not able to observe T (we do not know which point was deformed to which in the underlying set) but only the set $\widetilde{S}$ as a whole. Theoretical questions like consistency and asymptotic behaviour, and practical questions like efficient matching algorithms will be investigated.

Contour line recognition by means of orthogonal and non-orthogonal series expansions (wavelet approximations)

The contour of a Jordan (closed contour) set can be transformed to a periodic function by using the change-in-angle function, normalized to contour length 1. This periodic function has been analyzed by time series methods in the past ([ZR72, PF77, KE89, KC81]). We intend to use wavelet approximations for this purpose as it is known that they give a good compromise between over- and undersmoothing ([Don92b, DJ92, Don93, Don92c]).

Parameter free classification and discrimination using wavelets

Wavelet approximations may not only be used for noise reduction (contour smoothing) but also for information extraction. We will experiment with different wavelet bases to find interpretable coefficients and classification methods based on these (cf. [PC93]).

Experimental design methods for optimal sampling

The optimal placing of sampling points for best approximation of a function is an important question in signal processing. Statistical methods of experimental design ([Fed77]) will be discussed with the colleagues from the mathematical department (Hans Georg Feichtinger et al.) who study this question in connection with non-orthogonal series expansions.

STIPL – A statistical image processing library and knowledge base

This database will collect all algorithms with extensive documentation developed during the research program (T1–T4). Furthermore standard methods concerning statistics and image processing (noise reduction, filtering, classification, clustering, set theory, et cetera) will be looked for (in cooperation with the other research teams), adapted to the problems and infrastructure, modularized and stored.

1.3 Numerical Harmonic Analysis & Image Processing

Task leader:

a.o.Prof. Dr. H.G. Feichtinger,
Department of Mathematics,
University of Vienna.

1.3.1 Problem Statement

Irregular Sampling Problem

The **irregular sampling problem** is concerned with the partial or complete reconstruction of a smooth signal from non-uniformly spread sampling values. A typical example is the restauration of an image with missing or wrong pixel values spread out throughout the image, or complete reconstruction of an audio signal from an oversampled signal with missing values.

The group has developed a variety of new algorithms for the complete reconstruction of band-limited functions, given the exact values and locations of the measurements. A signal is called **band-limited** if its Fourier Transform (in the natural setting) has a certain support,i.e. certain high "frequencies" are known to be excluded in the signal. The reconstruction of such a function from the values over a sampling grid is a simple task, and the corresponding formula is well known as Shannon's Theorem, and has important applications in signal analysis (only to mention CD players).

For "small" problems the reconstruction can be seen as a problem in linear algebra. However for typical applications in image analysis, with maybe 30.000 unknown spectral coefficients and 40.000 sampling points it is not anymore possible to solve the system of linear equations directly. Not even a supercomputer will be able to store the corresponding auxiliary matrices to obtain the solution through standard iterative methods within reasonable time. Therefore the full mathematical structure of the problem has to be exploited, and only a combination of recent mathematical methods together with experience from numerical experiments can give "superfast" algorithms.

Non-orthogonal Expansions

The main problem in **non-orthogonal expansion** is to expand a given signal into a series $f = \sum_{n=1}^{\infty} c_n f_n$ involving a given collection of possible building blocks $(f_n)_{n=1}^{\infty}$. This collection has to be "rich" enough in order to allow to write arbitrary elements in a given signal space by the choice of an appropriate sequence of coefficients $(c_n)_{n=1}^{\infty}$. In the classical case of non-harmonic Fourier series the building blocks are sine- and cosine function over some interval, but with given frequencies, which do not have to fit to the interval length (cf. [You80]).

Also different collections will be useful for different applications. It has turned out in the last few years that fast algorithms can be expected if such collections are built in a "coherent" way, i.e. if the building blocks are derived from a "mother wavelet" by simple transformations, such as shifts, or frequency shifts (also called modulations), or dilation operators. More recently further transformations have been found interesting by various authors (cf. Baraniuk/Jones). In general one can say that one would like to use smooth building blocks, because then partial sums are nice, smooth functions (in contrast to partial sums of the Haar expansions, which consists of step functions only), as well as a guarantee that a smooth function f requires only a small number of big coefficients (in order to have high data compression rate).

The task of finding the (unique) appropriate coefficients for a given sequence of atoms $(f_n)_{n=1}^{\infty}$ is obvious if the building blocks form a complete orthonormal system with respect to the standard inner product structure of signals of finite energy, because then it is enough to calculate the inner products. Unfortunately classical orthonormal expansions $f = \sum_{n=1}^{\infty} \langle f, f_n \rangle f_n$, e.g. using Hermite functions, are not local and minor local changes of a function may result in significant change of all the coefficients.

This is one of the reasons why non-orthogonal expansions which on the other hand are local and use well localized building blocks have been considered very extensively in the last few years. Among them approximation of images by sums of radial basis functions, (cf. Powell and Buhmann) Gabor expansions (cf. a list of 150 papers on that direction) or non-orthogonal (affine) wavelets.

In the case of non-orthogonality one faces the problem of finding the appropriate coefficient sequence $(c_n)_{n=1}^{\infty}$. They should depend linearly on the signal or satisfy other requirements, which may make them useful for adaptive data compression. If furthermore the system of building blocks is highly redundant the representation coefficients are not uniquely determined.

Depending on the situation various criteria for obtaining "good" sequences of coefficients are in use, e.g. minimal norm, or maximal concentration. Although the theory of **frames** has found a lot of attention (cf. [HW89, Hei90]) it does not appear as the most efficient approach for the determination of the coefficients in terms of speed of approximation and actual calculations of the coefficients.

Among the systems of most interest in the last few years are **Gabor expansions**, where the family of building blocks is obtained from one single function, typically a symmetric bump function, through time-frequency shifts. Another promising field in which we have carried out extensive theoretical work are **wavelet expansions**, where the family of atoms is obtained from a single "mother wavelet" through dilations and translations.

We have collected a list of approximately 160 papers on Gabor expansions and their applications in the last few years, many of them still unpublished. So despite the vast literature in the field there are still many important open problems.

1.3.2 Research Goals and Methods

We can describe the Research Goals and Methods in general terms as the search for efficient algorithms for the purpose of image processing, image transmission and storage, and pattern recognition. These algorithms will be based on ideas using wavelet theory, non-orthogonal expansions and new methods for irregular sampling. Depending on the demand and scientific importance of the corresponding problems we are going to deal with a variety of applied problems, which require additional mathematical modelling, systematic theoretical investigations and efficient implementations of the algorithms under consideration. The number of potential applications is very large, and therefore - depending on the development in the field, suitable topics will have to be chosen for the short-term planning of research.

The main research topics are

1. Irregular Sampling
   - Band-Limited 1-Dimensional Signals
   - Band-Limited and Non-band-limited (smooth) and noisy Images
   - Short-Time Fourier Transforms
   - Wavelet Transforms
2. Nonorthogonal Expansions
   - Shifted Bump Functions
   - Gabor Expansions
   - Wavelet Expansions
3. Applications of irregular sampling and non-orthogonal expansions

- image coding (transform coding)
- data compression (e.g. large wavelet coefficients)
- pattern recognition
- subimage-matching
- PET-scanner image reconstruction
- system identification from output signals (sampling of STFT)
- applications in radio-astronomy
- cooperation with Garching, Hubble space telescope site
- geophysical applications
- progressive image transmission (wavelet coefficients for large scales first)
- sharp image from several distorted ones (deconvolution mixed with data fusion)

In the course of the previous projects and through the theoretical contributions ([Gro90, Gro91b, Gro92, Gro]) of our coauthor K.Gröchenig from the Math.Dept. of the Univ. Storrs/CT in USA significant improvements for the 1D theory as well as very efficient implementations could be obtained.

The corresponding publications are just coming up and have turned out to be more efficient than those of the leaders in the field (cf. the separate list of publications of the working group). The current state of affairs in the field are described by the recognized authors of an edited book by R.Marks II, which came out in 1993, about "Advanced Topics in Shannon Sampling Theory" [MI93].

The fact, that the work of our group has been ignored by the authors of these articles is based on the fact that our results have not been yet published when those manuscripts where written. On the other hand, we can point out that our group has meanwhile obtained substantial improvements over the results on irregular sampling as described by F.Marvasti or H.Stark.

Band-limited 1-D Signals

The joint survey "Theory and Practice of Irregular Sampling" (cf. [FG93c]) done with K.Gröchenig describes theoretical background and practical performance of the 1D methods tested extensively so far. In particular if features the so-called ADPW (Adaptive Weights Method) as suggested on the basis of theoretical investigations by Feichtinger/Gröchenig (cf. [FG92b]).

The ADPW-method is shown to be fast, efficient and stable. [FG93c] also describes the significant advantage of this method compared to others published methods in the field. For example, methods, which make use of Voronoi-step-functions or piecewise linear interpolation as proposed in [SA87], are computationally rather expansive. On the other hand, the simple frame iteration, also called Wiley-Marvasti method, has the serious draw-back of being very sensitive to irregularities, such as clusters of sampling points.

The literature (including our own publications) describes almost exclusively methods with a geometric rate of convergence, and give often poor results in terms of speed. Also, auxiliary matrices tend to grow beyond manageable sizes even for images of standard size, such as 512×512. Therefore one has to look for efficient methods from numerical linear algebra (such as the Conjugate Gradient Method, with appropriate preconditioning) and try to perform iterations without actually storing the relevant matrices in full size. For a number of structured problems (e.g. sampling sets which are a finite union of regular lattices) such "superfast" algorithms can be designed, but numerical experiments and further theoretical investigations have to show which combination of "tricks" works best for a given class of problems.

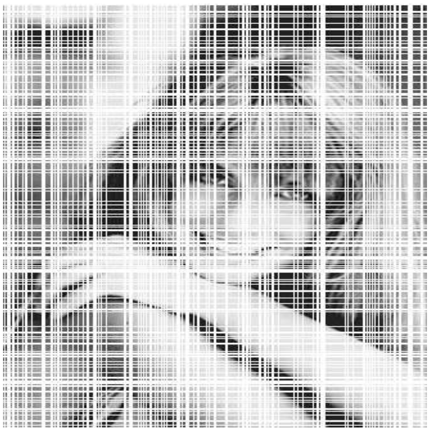
This design problem may also take into account the means available to solve these equations, such as the special features of a digital signal processor or the structure of the computational machine (either purely sequential, or parallel to a given degree) on which the calculations have to be carried out, and even programming language involved.

Compiled C-code for the 1D-case has been produced which allows to demonstrate the efficiency of this algorithm also to people with commercial interest in such methods.

Reconstruction of band-limited and non-band-limited images

The aim of this task is to develop theory and implement 2D versions of ideas explained above, which fully exploits the structure of the mathematical problem at hand. The findings so far also demonstrate in a very clear way that it makes sense to study such a "simple" applied problem in all its mathematical aspects to find highly efficient algorithms for its solution.

(a) Sampled Image



70850 sampling points

(b) Reconstructed Image

error = 1.588e-05

Figure 1.4:

As an example Fig. 1.4a shows an image where a number of vertical and horizontal lines have been removed. Fig. 1.4b shows a quite good reconstruction out of Fig. 1.4a.

From our present knowledge it is clear that the resulting methods will be more than competitive with other standard methods in the field. It appears realistic to expect that the fully developed algorithm will become a new standard at least for a certain type of applications. Also the most successful areas of applications have to be identified.

There are, however many further applications to the problem of image restauration or scattered data interpolation which should be exploited as well, and maybe even put into hardware, in case appropriate partners can be found.

Also, there are many more different types of "semiregular" sampling sets in the 2D case compared with the 1D setting which have to be studied and for which perhaps special methods may be developed. Let us just mention the conversion problem from polar coordinates to a rectangular grid which is a critical step in the direct Fourier method of computer tomography and probably the weak point in this method so far.

Short-Time-Fourier-Transform and Gabor Expansions

It is also clear from our experience and from theory (cf. [FG93c]) that efficient methods for the solution of the 2D irregular sampling method will yield a very good basis for the reconstruction of a short time Fourier transform (STFT) from irregular samples.

Such a problem has useful applications in the identification of a linear, translation invariant system from the responses to a family of frequency shifted input signals with known envelope. Such a test for a system appears much more realistic than the traditional ones based on the concept of impulse response, or a transfer function. If methods in this direction work as expected it will be also possible to develop testing rules for the application of this method of systems identification.

For example, new and more efficient methods to determine the "dual Gabor window" have to be discussed in detail. It is used as "window" for the STFT through which one can determine suitable coefficients for a Gabor expansions with given building block (cf. [Chr93]).

First experiments aiming at the use of the conjugate gradient method to solve this task have been quite promising and should be continued, and also transferred to the 2D case. The new approach will allow to calculate such dual Gabor windows for more realistic signal sizes. Until now rather big auxiliary matrices have to be inverted, and therefore the use of Gabor expansions has been limited due to computational costs.

Wavelet Transforms and Wavelet Expansions

The group has already preformed systematic research on the irregular sampling problem for the wavelet transform. This problem is closely related to the problem of constructing one sharp image from several smeared ones, taken at different resolutions. Numerical experiments are expected to confirm the efficiency of the algorithms described so far only theoretically, and to improve them on the basis of experimental findings.

Another direction that has gained recently much interest is adaptive signal processing in the following sense. Any given set of building blocks is usually designed to be large enough to allow representation of arbitrary signals of finite energy by means of a suitable choice of square summable coefficients. But it is often *not* well adapted for certain types of signals. For example Gabor representation may not be good for the representation of chirp signals. One possible solution is the use of a larger "dictionary" of building blocks (cf. [MZ92, Mal89a, MZ90]).

Therefore – in addition to the problem of finding appropriate coefficients – we have the extra problem of a clever selection of those building blocks that should be used for a signal representation.

The algorithms presented so far appear to be - despite their theoretical interest - rather slow and far below the potential of this new approach. Nevertheless there is justified hope that within the next few years new and more efficient methods will be found. The odds that our group can make relevant contributions appear to be good on the basis of the research done so far.

One also should mention the wavelet packages developed mainly at Yale University, for which a well developed theory is already available (cf. [Wic90]). They use big families of orthonormal systems of wavelet type, and also allow adaptive expansions of signals.

Image coding, data compression and progressive image transmission

In such applications the purpose is efficient coding of images or general signals of several dimensions, including image sequences. The interpretation of wavelet or Gabor expansions coefficients in terms of time/scale suggests that "relevance" of given contributions can be handled in a natural way. E.g. details of an image might be required only in a certain part of the image (for medical imaging), or high frequencies in a signal only in some areas of high oscillation, whereas smooth parts should require considerable less storage.

Therefore using the largest coefficients in such an orthogonal or non-orthogonal expansions is reasonable and is under discussion be different research groups all over the world (e.g. Jawerth-deVore in South Carolina, Wickerhauser-Coifman at the Yale Univ., or St. Mallat at Univ. New York, with his interpretation of such maxima as multi-scale edge information).

Besides the choice of a suitable family of coefficients which are to be stored or transmitted, there is also the problem of reconstructing efficiently and fast from the stored information.

For progressive image transmission, such as image-telephone, it is important, that high resolution can be transformed on top of quick low resolution images. Local signal expansions are used such that additional information only has to be transmitted in small quantities if only certain parts of the image are changing (e.g. small object moving through a larger image).

In this area both theoretical and practical work has to be done. Especially the problem of reconstruction from a given set of coefficients which will be chosen in an adaptive way in the next generation of such compression tools is closely related to irregular sampling and therefore the use of efficient methods from irregular sampling for such applications can be expected to give new and practical reconstruction algorithms.

Pattern recognition

The choice of certain families of Gabor building blocks or wavelet systems can be used to map the signal space into a suitable coefficient space. The choice of a finite subfamily allows to access the pattern recognition and classification task on the level of those coefficients.

There are realistic expectations that the choice of the basic wavelet will have some influence on the viability of this approach. Therefore the restriction to orthogonal wavelets appears as not well motivated for this particular task. Actually, minimal redundancy rather corresponds to reduced stability of related algorithms.

It appears, that somewhat redundant systems, such as Wavelet or Gabor frames will be more suitable for this kind of application. Theoretical work on the theory of frames in connection with non-orthogonal expansions is carried out by Ole Christensen in his thesis [Chr93], written as a member of our group.

Subimage-matching

Among others there is a series of papers by Ben-Arie and Rao on the problem of identifying subimages within a larger image by the attempt to obtain a non-orthogonal expansion of the large image in terms of shifted copies of the small one. Significant coefficients in such an expansion indicate very clearly and often better than correlation considerations the presence of the given subimage within the large ones.

The efficiency of such a method depends of course on the easiness of obtained such non-orthogonal expansions for arbitrary subimages. Carrying out this task on the basis of a marked subarea of a given real world image and testing the performance of this approach will show its efficiency.

PET-scanner image reconstruction

We are in contact with a medical group in Philadelphia around Prof. G.Herman. Many reconstruction methods for tomographic devices are based on a decomposition of a 3D-volume in terms of volume elements, called **voxels**.

In mathematical terms this corresponds to the approximation of the given (unknown) variable density of a 3D body under inspection by a step function, constant over cubic areas. In the case of **Photo Emission Tomography** (=PET) the variable density has some smoothness. Therefore step functions are not the best form of approximation through space shifted copies of a given function (the cubic block). In fact, smooth building blocks can be expected to perform better.

Investigations should be directed towards the question of finding optimal pairs of suitable radial basis functions and regular shift-lattices (e.g. hexagonal ones) which result in optimal approximation with the minimal number of terms. Such results will help to reduce the complexity of PET reconstructions and/or will probably improve the quality of this method.

System identification from output signals by sampling of STFT

Usually **Translation Invariant Linear Systems** (=TLIS) are identified by the impulse response function. This requires to check the output for a kind of "shock"-input, which might destroy the system, or come into conflict with the frequent "approximate linearity" assumption for real world time-invariant systems.

An alternative is the description through a transfer function, which requires testing a large number of frequencies, which is rather time-consuming. Checking the output of the system for input signals which are just frequency-shifted copies of some fixed bump-function would allow at least partial reconstruction of the signal, even if the output is only sampled in an irregular way.

Mathematically the problem is equivalent to reconstruction of a signal from sampled short-time Fourier transforms. This can be done for one and two-dimensional systems (in practice). However, the complexity of the task requires very efficient methods in order to avoid endless calculations. Even for the 1D-setting performance is a question of interest.

Applications in radio-astronomy

During the workshop on wavelets and irregular sampling organized by H.Feichtinger in early March 1993 at the Math.Dept. Univ.Vienna it became clear that significant interest from radio-astronomers (e.g. Prof. Usowicz, Turon/Poland) in the 2D irregular sampling problem.

The fact that targets observed by the huge telescopes are rather small imply that the data obtained from interferometric experiments can indeed by well interpreted as the 2D-sampling of a band-limited image, and that the reconstruction of the "true" image (e.g. some configuration of stars) based on Fourier methods is equivalent to the main task of this sub-project.

Astronomical observation are also often bound to accept long integration times, because of weak signals and high SNR (signal to noise ratio). Therefore, one often obtains rather average values than sampling values of the function in question. The newly developed method of reconstructing a function from averages might be a very helpful general approach to such problems.

Sharp image from several distorted images

In this field we are in cooperation with Garching, Hubble space telescope site. A recent visit of Dr.Adorf from the European site of the Hubble space telescope has indicated the need of irregular sampling methods for certain mathematical problems, related to the deficiencies of the Hubble space telescope.

The deformed images due to the deformation if the main-mirror of the telescope might be partially corrected through the use of suitable methods related to irregular sampling.

Also in this area the question of obtaining sharp images from several smoothed versions is of importance. It arises in the following sense: The limited resolution of CCD-cameras allows only to obtain a given resolution by taking one measurement, even if one has high-performance instruments. However, it is possible to take a number of subsequent images over a period of times. Each of those images has the same resolution (to low for certain purposes). The movement of the earth and the telescope with it within the galaxy corresponds to different shifted versions of the same, smooth image. Since the telescope (or the earth) are not freely movable according to he wishes of scientist making measurements one has to start from data which correspond to irregularities if one tries to match those data directly with a fine grid. Irregular sampling methods will help to overcome the mathematical difficulties related to problems of this kind.

In this case we tend to mix mathematical methods for deconvolution of such images with the concept of data fusion – also investigated within this research program.

Geophysical applications

Typical geophysical measurements (e.g. magnetic or gravitational field measurements) are taking place with high precision but at positions which do not cover a rectangular lattice. This is either due to the difficulties of taking measurements at exactly the grid points, or due to the movement of the measuring instrument taking up the data (e.g. helicopters, aircrafts or ships)

Nevertheless, in order to obtain good interpolation to a rectangular grid efficient methods are required. Since typically gravitational potentials or magnetic fields do not show too sharp local changes the assumption of approximate band-limitedness for the signal (the basis for our reconstruction methods) is fulfilled in a satisfactory way.

Further aspects to be investigated in the subproject

This section mentions some general aspects to be investigated in the subproject. They report some main ideas of your further work.

irregular sampling for higher dimensions

So far we have dealt mainly with (very) efficient new methods for 1D irregular sampling, most recently we are able to say more about the 2D setting (i.e. reconstruction of band-limited images from irregular sampling). The full potential of those new methods, in particular for sampling patterns showing some kind of regularity, appears to be high and has to be investigated next.

For several applications it will also be necessary to go to higher dimensions. 3D and 4D applications appear if one studies sequences of images or sequences of 3D data. Of course, this will require sufficient computing power in order to do realistic examples.

use of real image data

For the evaluation of algorithms until now mainly synthetic data have been used, which allow a verification of speed and quality of the complete reconstruction.

But since during the project the application of our methods to real data, such as noisy images, have to be carried out, a detailed study of stability of our algorithms with respect to perturbations has to be undertaken.

The indications from theoretical research (cf. [FG93a]) in this direction are encouraging. This circle of ideas also requires to deal with questions of statistical stability, and therefore cooperation with the statistical group will be important.

multigrid and domain decomposition approach

For large images even some of the promising methods under consideration right now will not be applicable anymore (independent of questions of speed). Therefore a splitting or multigrid approach will be necessary. Contacts with Prof. Oswald in Jena (an expert in approximation theory and finite element spaces, who also works on the connections to wavelet theory) will be helpful in this respect (cf. [Osw91, Osw]).

Furthermore, domain decomposition methods are expected to be relevant for images with changing smoothness and variable sampling density.

development of parallel algorithms

As far as we can see the algorithms under development are very well suited for parallel processing. Therefore, during the time of the 5-year project we will have to observe both the relevant applied and mathematical literature as well as the developments on the hard and software sector.

Contacts to mathematicians and computer scientists working within the Austrian Center for Parallel Computing (supported by the FWF) will be helpful to obtain efficient algorithms on that basis, but definitely the mathematical structure of the problems under investigations has to be studied very carefully and will provide better results than direct parallel implementations of the present algorithms.

But parallel algorithms are not only based on a multiple datastream – often used in images analysis. We also tend to develop parallel algorithms based on the parallel mathematical structure of the problem. This idea also leads to multigrid and domain decomposition approaches – as described above.

Bibliography of relevant references

It is hard to give "the" list of references to the field, because it is growing on a daily basis. For example, Mr. Stefan Pittner from the Technical University of Vienna, is collecting just the mathematical wavelet literature since about 1986 (cf. [Pit93]). His next edition will contain about 700 entries, about 300 of them within the last 2-3 years.

We hope that it will be made available to the interested public via anonymous ftp at our server tyche.mat.univie.ac.at(131.130.22.31) with his kind permission.

A bibliography on Gabor representations put together very recently by H.G.Feichtinger contains at the moment more than 150 entries, about one third of the papers still unpublished, but available to our working group through personal contacts to the authors. It is also available via anonymous ftp from tyche.mat.univie.ac.atas LaTeX-file (pub/gaball.tex).

In the field of irregular sampling the bibliography of a recent book ([MI93], published in 1993, R.Marks, Ed., "Advanced topics in Shannon Sampling and Interpolation Theory"), which contains 1003 !! bibliographical references, indicates the wide-spread interest in this problem.

The list of publications on wavelet theory, Gabor representations, and irregular sampling within the Numerical Harmonic Analysis Group at the Dept. of the Univ. of Vienna will be given separately. The actual version is kept available via ftp on tyche.mat.univie.ac.atas pub/grouppub.tex.

We also like to mention the so-called "WAVELET Digest", published through Electronic Mail about every third week. It announces new papers, books, special conferences and workshops in this field each time it appears. We are trying to keep track of those developments in our seminar and in private discussion within our working group. The enormous flow of information in this field cannot be handled by a single person, but it is important to have a small group of people working in this field in order to provide students in Austria (and Europe) preparing for a master or PhD thesis, as well as scientist from other areas, with basic information in order to allow them to do efficient research in this field.

1.4 Parallel Processing Strategies for Large Image Data Sets

Task leader:

o.Prof. Dipl.-Ing. Dr. techn. Franz Leberl,
Institute for Computer Graphics,
Technical University, Graz.

1.4.1 Problem Statement

Introduction

Computer vision is often regarded as one of the most complex and computationally intensive problems (Choudhary et al., 1991). Image processing and pattern recognition solutions have traditionally been successful only if special computers were available to accelerate the massive number of pixel operations needed in most vision problems (Kumar, 1991). There exist, however, at least two classes of image processing domains:

real time vision, often for a narrow purpose in support of industrial automation, and of other real time applications;

systematic processing of large image data sets, not in real time.

We believe that these domains are fundamentally different in the manner in which computers and algorithms are being used. As we observe the rapid increase of electronic imaging technology and the resulting acceptance of machine vision in industrial automation, we also observe an enormous proliferation of computing architectures, operating systems and languages to accomplish the fundamental goal of such vision systems: process an image in real time, i.e. in about 16 milliseconds. This is the time it typically takes to create the image with a square-array CCD-camera consisting of 0.25 Megapixels.

On the other hand there exist systematically collected image data sets, for example in remote sensing, planetary exploration or medicine, when a repository of thousands of images accumulates about a medical specialty and phenomenon, about a planet, or about the Earth. Data quantities in remote sensing are simply a result of data collection abilities. Processing and analysis tools lag far behind. In orbit, a satellite may generate images at a rate of 18 Megabytes or more per second. Image data repositories would be filled very rapidly in the event where a satellite continued to image day and night, for months on end. Fortunately, data rates are not being sustained continuously at 18 MB/s. Instead, data sets are being collected for a specific purpose and a certain time is subsequently available to process and analyse the images.

We believe that these issues are important, that we have special capabilities to address them, and that our work will propel us as an Austrian team into the center of the European research scene. Therefore it is in this domain that we propose to contribute to the Austrian Forschungsschwerpunkt FSP. We plan to work on image analysis algorithms, data structures and processing strategies. While historically such large data quantities were addressed with especially built computers, we believe that this problem domain lends itself to a solution using so-called "open computing systems" for parallel processing on various computing tools, to include

massively parallel computers such as the Intel-Paragon and similar systems;

multi-processor RISC workstations such as the Silicon Graphics Power Series and similar systems;

clusters of multiple low cost graphics work stations such as the SGI IRIS, its competitors and its successors.

In honing our skills in this domain we expect to contribute to the solution of a significant image processing problem which appears to be often neglected. As experts in dealing with massive amounts of image data we expect to inspire applications in international environmental programs, in regional and national remote sensing, in planetary research and in the systematic analysis of vast repositories of medical images.

Motivation

(a) Computing Resources through Parallel Processing

The Principal Investigator for this task has participated in the US space research program since about 1974 (Leberl, 1975, 1976; Leberl and Goetz, 1976). Since about 1979, opportunities developed for close participation in the most ambitious planetary imaging program ever undertaken, the mission Magellan to planet Venus (Leberl, 1980; Leberl et al., 1991, 1992a,b; 1993a). It resulted in an image data set encompassing about 300 Gigabytes of data. Simultaneously, work on images from Earth observing satellites, for example to routinely track sea ice motion (Leberl et al., 1979, 1983; Leberl, 1989 and 1993b), created great concern for the need to deal with large image data sets in a sensible manner, and that is by avoiding especially built hardware.

While such data have been processed on special hardware at the rate of collection, this hardware seems by now to be obsolete and replaceable by off-the-shelf high performance computers. We believe the time has come to move into so-called "open system parallel computation" (Webb, 1991) and indeed to employ generally available computers for the analysis of large digital image repositories. This is one of several drivers for the proposed research.

It is through the skills of Austria's leading experts in parallel computing, Prof. Zima and Ms. B. Chapman, that the work on the proposed problem domain promises to be efficient and at a high level.

(b) From Data Base to Image Information System

There exists a tradition to place images upon collection and initial inspection into an image repository where access to individual scenes is by the order in which such images were collected. This is a rather unintelligent manner to store large images, since it is highly unlikely that any future use of, and access to, the images will be by time and order of collection. Instead it will in all likelihood be by geographic region, by neighbourhood, by type of phenomenon under investigation.

A need develops to create an image information system from large image repositories. The result is to consist of a data structure that will support the ease of access to large data sets, and the queries as a function of a user's likely needs. Without such data organisation, the size of a large image data set may overwhelm its user. National and regional image data bases will have to convert into image information systems to become manageable and useful to the intended user. For this reason we propose to investigate good structures, query and interaction systems which support the use of large images as members of very large data bases.

(c) Algorithms

In industrial machine vision there exists a need for a great multitude of individual, but often limited computing tasks which need to operate on small images for the robust automation of simple manual tasks. In large image data sets, in contrast, there often is a need for only very few algorithms, but they must be applied to very large data quantities and perform work that a human may not even be able to perform, or can perform only after months or years of training. A common example is the stereo-process for shape reconstruction from overlapping images (e.g. Haralick and Shapiro, 1992), or the surface classification on the basis of multi-spectral or other characteristics.

We believe that there exists a need to analyse images to reconstruct from them not only the shape of an object, but also its surface reflection properties. Commonly, reconstruction of shape is considered a separate task, and a difficult one for which no generally accepted successful methods exist.

Instead, solutions are for restricted special cases (Gruen, 1989; Norvelle, 1992). Manifestations of reflective properties of an object often are considered a perturbation to the surface shape reconstruction, particularly in echo-range images such as radar (Leberl, 1989). We propose to prove the opposite: we believe that the reconstruction of shape should not be done without the simultaneous reconstruction of surface reflective properties. As a result we expect that we not only extract from images considerably more information about an object than is customary, but that we accomplish improvements in the shape reconstruction as well.

Specifically we want to start by investigating echo-range images, again from the Venus-satellite and from other sensing sources. We want to develop concepts that combine stereo-analysis with shape-from-shading in a manner which will permit one to employ images even if they are too dissimilar under traditional stereo-rules. This will not only satisfy the need for improved shape measurement accuracies and robustness so that full automation can be accomplished, but will also satisfy the need to describe an object's surface type and material.

State of the Art

Our proposed task builds on various but disparate foundations in signal processing, surface reconstruction, image matching, visualization; in parallel computing and automated parallelisation; in data structures and spatial information systems. However, large image data sets have not been a distinct focus of research. Of course, medical image processing finds itself faced with sizeable image quantities in Picture Archiving and Communications Systems (PACS, Leotta and Kim, 1993). However, these images are typically somewhat smaller than images in remote sensing. Yet, throughout a year, a typical hospital may still accumulate a total of 1 to 2 Terabytes of images (Osteaux, 1992).

(a) About Large Images

a1. System for Large Images.

Satellite remote sensing images have been traditionally processed in centralized locations by space organizations. Systems to cope with large data quantities of 100 MBytes per image have been the result of massive computer system development, such as those at NASA's Goddard Space Flight Center, or at the European Space Agency. The Earth Observation Systems era is dawning by 1997 and massive amounts of image data are expected to become available for routine processing. At issue is therefore the concept of an EOS-workstation (Kober et al., 1988). Recently, Archiving and Operations Systems have been set up for large image quantities in de-centralized locations, for example at the University of Alaska, without special hardware, but off-the-shelf computer facilities and coping with 40.000 to 100.000 images. First prototypes of smaller systems for processing and coping with routine large image data flows address a specific applications domain, and are built with commercial off-the-shelf computers and accelerator boards (Kwok et al., 1990; Leberl et al., 1993b). A contemporary system of relevance is of course NASA's ground system for the Magellan mission itself: it is built with special hardware (Leung, Jin et al. (1992). Such a system consists of the routine processing of all incoming image data, as well as the applications-specific processing of selections of the entire data set. Visualizations of processing results again have been based on special hardware (Kirk et al., 1992), but is also being considered for transfer to off-the-shelf parallel computing systems (Li and Curkendall, 1992).

a2. Interaction with Large Images.

Interaction with large images and groups of such images is a problem of its own. Interaction is often a neglected issue in image processing research, presumably since the optimism is high that all tasks can be fully automated. Yet there seems to always exist a need for interaction with the images for quality control or machine support. Concepts for such interactive tools for large images and image groups to browse, view, zoom, roam, comparing of image pairs etc.

need to be integrated into any system for large image data sets. Therefore the development at Pixar of a so-called electronic light table (ELT) received so much attention (Pixar, 1989). An entire family of interactive work stations for large images was developed at General Dynamics (see Leberl et al. (eds), 1992), or at Matra in France (Lamoure, 1992). These tools, even today, remain supported by special hardware. A recent and promising development was inspired by NASA's Magellan mission and resulted in a sophisticated pyramidal, tiled structure of the data in support of rapid interactive data access (Leberl et al., 1992). It differs from previous work by strict reliance on open computing systems and on the integration of a measuring function into the viewing of multiple large images. This project has clearly shown the viability of refraining from special hardware, yet accomplishing real or near real-time functionality for interaction.

(b) Information Extraction Algorithms

b1. General.

Software for image processing and pattern recognition has been developed since the early 1960's, where space applications were always a prime driver for innovation (Rindfleisch 1966; Rindfleisch et al., 1971). The current paradigm for image processing is based on a "tool box" of algorithms. A task-specific solution is configured by combining various elements from the tool box into a working system. The creation of a complex structure from individual elements has at times been denoted as "plumbing" in analogy to piping systems for a specific data flow. An example is the applications-specific solution to determine the motion of sea ice, and its type, from large satellite radar images (Leberl et al., 1993b). We will discuss in the following the status of only a limited algorithm domain, since we propose to focus our efforts on a specific set of algorithms in the context of our large image data base paradigm.

b2. Radar Image Matching Algorithms.

Surface reconstruction from overlapping images requires that geometric differences between these images be detected. A great amount of work has been performed in this area (Holden et al, 1992; Mueller et al., 1988). Any attempt at merging, comparing, mosaicking, sequence analysis etc. will always lead to a matching requirement. The vast image matching literature has recently been reviewed by Gruen (1989) and in Foerstner's chapter 16 of the monograph by Haralick and Shapiro (1992). Foerstner is the source for the most widely used sub-pixel matching algorithm, least square matching (Foerstner, 1982). Matching may be performed in the spatial or spectral domain (Anuta, 1970). However, almost all algorithms neglect the strong non-geometric, i.e. radiometric dissimilarities between candidate images for an image match. This is caused by the focus of most matching research on optical imagery, thereby neglecting issues of matching active sensor images. In light of massive satellite data repositories currently being collected with such sensors (European ERS-1, Japanese ERS-1, Radarsat, Shuttle SIR-C, Russian Almaz), this neglect needs to be corrected. A review of radar image matching work was presented by Leberl (1989) and Leberl et al. (in print), referring to work by Ramapryan et al. (1986), Guindon et al. (1986), McConnel et al. (1989, 1991) and others. One will conclude from these reviews that this subject has not been studied extensively.

A particular concern in need of a solution is the ability to match overlapping radar images in the presence of strong illumination differences. There exist practically no previous studies on this subject, with the exception of McConnel's work (1991). Fullerton et al. (1986) and Kirk (private communication)[1] reported also on some of a very small set of efforts, however either without any, or with questionable success.

b3. Surface Reconstruction from Radar Images.

[1] R. Kirk, US Geological Survey in Flagstaff, Arizona, has orally reported about unsuccessful attempts to use overlapping radar images taken from opposite sides of a surface, to use in a multi-image shape-from-shading approach.

Image matching and detection of geometric differences is but one of several ways in which to obtain surface shape information. Others include shape-from-shading, use of shadows, altimetry, interferometry. In the current context we find stereopsis, i.e. the use of geometric differences in tow images, and shape-from-shading as very promising concepts, particularly if there exist no alternative data sources for shape information, as is the case in the Venus environment. Work on radar stereo began in 1963 with LaPrade (1963). Several reviews were authored by Leberl (1979, 1989), and limitations are dominated by illumination differences which always accompany the required geometric differences; these limitations are reasonably well-understood.

A separate method is shape from shading (SfS). The technique has been discussed by Horn (1989) and others for optical imagery. Wildey (1986) began a series of efforts on radar images, followed by Frankot and Chellappa (1987), Kirk (1987) and Thomas et al. (1991). The radar relevant work was reviewed by Leberl (1990). The limitations are obviously in the need to know the surface reflection properties to infer from image brightness the surface slope, and then to be able to integrate slope measurements into a surface shape. SfS could use multiple images to overcome the limitations from unknown reflective properties, but suffers the inverse problem of stereopsis: it requires that overlapping images already match, so that the radiometric differences can be used to infer surface slope. As a result it was Thomas et al. (1991) who first proposed and demonstrated the combined use of stereopsis and SfS. However, this work was preliminary only. We hope to achieve progress through the work proposed here by combining stereopsis also with SfS, but in an approach that models not only surface shape, but also surface reflective properties.

b4. Other Algorithms.

Surface shape reconstruction and image matching clearly are the most important tasks in the initial work with large image data sets. It is through these that images can be related to one another, and that hope exists to convert an image data set into a useful image information system. However, there are numerous other image processing algorithms in need of study and improvement. These include fast methods of low and intermediate level processing, for example image filtering, resampling, edge, line and region extraction, pixel and region classification, visualizations etc. While these individual topics are important, they are not subject to the same degree of urgency in finding innovative solutions as we had in the previous topic. Instead, the issue is one of adapting well-known approaches of low and intermediate level vision (Gonzalez and Woods, 1991; Haralick and Shapiro, 1992) to the specific characteristics of range images. No work has been reported that would address the application of low and intermediate level image processing algorithms to groups of overlapping, matched radar images. Therefore we believe that we will be able to chart new territory in our proposed work by employing the matched radar images in the form of an "image cube" for systematic feature detection and classification.

(c) Parallel Architectures for Vision in Large Images

c1. Machine Vision.

In machine vision there has been an avalanche of work on parallel computing architectures solving rather specific vision algorithms (Duff, 1983; Pitas, 1992; Cantoni and Levialdi, 1988; Uhr, 1987; Maresca and Lavin, 1988 and others). There certainly exist at least 25 international groups with excellence in this domain of research (Kumar, 1991). Numerous issues are being addressed with taxonomies, network topology, data path widths, data path types (Shu and Nash, 1990), operating systems, languages etc. This work is justified and meaningful in the context of real time analysis in industrial settings; such a system may still be based on special boards but increasingly employs off-the-shelf processing elements (Webb, 1991). In fact routine industrial automation would not be affordable, were there a special board to be developed for each solution . As a result, much non-routine work in this domain is purely speculative

(Krikelis, 1991). Regarding typical machine vision applications, an extensive bench mark has been organized in the USA under DARPA-sponsorship as part of the US Image Understanding program to assess and compare various parallel architectures (Weems et al, 1991). Not surprisingly, architectures especially built for Image Understanding, the IUA (Image Understanding Architecture, see Weems et al., 1992; Shu and Nash, 1990) and ASP (Lea, 1988) performed in the best category. Yet we believe that this research in Image Understanding is not relevant when we move from the real-time machine vision issue to the issue of large image data sets.

c2. Fast Processing of Large Image Data Sets.

All space agencies have had an ongoing need to cope with large data flows from imaging satellite sensors. This is going to increase with the advent of the "Mission to Earth" and the Earch Observations System EOS. Traditionally, these large data streams were subject to massive processing in special computers or attachments to off-the-shelf computers. The needs of space agencies differ from those sketched in this proposal; we are concerned with the creation and use of large image data bases *after* the data collection has been completed. Real time is not a concern in this context, whereas in a ground receiving station, it is. Discipline-oriented applications processing of many large images can be performed on fast multi-processor systems or on massively parallel systems (Mueller et al., 1988). Holden et al (1992) have demonstrated the use of a cluster of workstations to deal with a complex, time consuming algorithm for matching of large optical images.

c3. Open Parallel Computing.

An open computing system is not designed to specifically deal with vision algorithms, although "open parallel computing" is a concept occasionally addressed in the vision research community (Wallace et al., 1988; Webb, 1991; Bruner and Reeves, 1982; Reeves, 1989). While massively parallel computers or workstation clusters may not help significantly with industrial vision problems, the compute power and storage facilities they offer are likely to help process large images by means of the adoption of a programming methodology based on the segmentation of input data and a data-parallel approach to computing (Karp,1987; Chapman,1991). High-speed access to the large distributed files used by the application program is essential to parts of this task.

Massively parallel systems offer the promise of virtually unlimited scalability in the number of computational nodes, and hence their computing power, as well as in the size of the associated memory subsystems. The use of off-the-shelf processors enables their construction at reasonable cost. Current systems include Intel's Paragon and iPSC/860, Thinking Machine's CM5, the Meiko Computing Surface CS-2, and various transputer-based machines. A major characteristic of these parallel systems is that each processing node has its own local memory; hence, they are known as distributed memory parallel computers (Zima et al.,1990). The time required to access a non-local datum, i.e. data not stored in a processor's local memory but in that of another processor, may be an order of magnitude higher than the time taken to access locally stored data. This has important consequences for the strategies adopted by an applications programmer. In particular, the management of data, with the twin goals of both spreading the computational workload and minimizing the delays caused when a processor has to wait for non-local data, becomes of paramount importance (Chapman et al., 1992b) when attempting to achieve good performance.

There are two specific issues which are of overriding importance to the tasks approached in this project, as far as the adaptation of the algorithms and their coding for parallelization is concerned. One of them is the need for a method to express algorithms by determining a portable representation of the code which can be efficiently utilized on a range of target architectures. The state of the art in this area is described separately in c4 below.

The other major issue is the necessity to handle large amounts of data in secondary storage efficiently. The initial processing of large amounts of data as required within this project, as well as the efficient subsequent retrieval of data sets, will place very heavy demands upon the management of secondary storage facilities. Efficiency in this area will be critical for the overall success of the parallelization of these methods.

Massively parallel systems, such as those named above, address the problem of providing large amounts of storage by combining them with powerful concurrent I/O systems [Asbury et al., 1989; Pratt et al.,1989]. The hardware and software currently available from various vendors differs, however. The Concurrent File System (CFS) developed by Intel for the iPSC/2 and iPSC/860 machines (Pierce, 1989), for example, is based on a so-called striping scheme, in which a single file is spread across multiple discs of the I/O system to improve access speed and decrease congestion in communication links. This striping is done at the logical block level. Several methods of accessing striped files are provided, but they must be controlled by the user. To obtain efficiency, it is critical that the data structures used in the application code are mapped to the disks of a distributed filing system in a manner which enables efficient reading and writing of data by the computational nodes in parallel. The actual mapping and mode of access will therefore depend on the way in which file data is accessed in the nodes.

Language constructs for supporting parallel I/O at a higher level than that available currently have been put forward, but have not been generally adopted (Brezany et al., 1992; Pase, 1991). This project will require a comparison of the capabilities of I/O systems and access to distributed files, and will include an investigation into the compiler's ability to generate code to use these efficiently.

c4. Programming Parallel Processing Computers.

A major difficulty with distributed memory parallel computing systems as provided commercially is that, so far, they generally lack programming tools for software development at a suitably high level. The user is forced to deal with all aspects of the distribution of data and work to the processors, and must control the program's execution at a very low level. This results in a programming style which can be likened to assembly programming on a sequential machine. It is tedious, time-consuming and error prone. This has led to particularly slow software development cycles and, in consequence, high costs for software. It has also made the resulting software not very portable.

The fact that the time taken to adapt existing sequential code and to develop new applications is prohibitive in comparison to conventional machines, including vector supercomputers (Chapman 1992b). This has been a major factor in the relatively slow acceptance of massively parallel machines in the market place.

Thus much research and development activity is now concentrated on providing suitable programming tools for these architectures. A major portion of this work is aimed at the provision of appropriate high-level language constructs to enable users to design programs in much the same way as they are accustomed to on a sequential machine. Several proposals have been put forth recently for a set of language extensions to achieve this (Chen et al.,1992, Loveman, 1992; Fox et al., 1991, Steele, 1992; Zima et al., 1992, Benkner, 1992), in particular (but not only) for Fortran, and current compiler research is aimed at implementing them.

This work was focussed in the High Performance Fortran Forum, a group of academics and representatives from industry which met regularly in the USA during 1991 and resulted in a preliminary set of language extensions, High Performance Fortran (HPF) which are expected to be supported on a number of massively parallel machines (HPFF, 1993). The Vienna Fortran programming language was a major input to this effort, and many features of HPF have been taken directly from it, although in some aspects Vienna Fortran goes far beyond the current HPF extensions.

Research in compiler technology has so far resulted in the development of a number of prototype systems, such as Kali (Koelbel et al., 1991), SUPERB (Gerndt 1989, Zima et al., 1988), and the MIMDizer (PSR 1991). In contrast to the current programming paradigm, these systems enable the user to write code using global data references, as on a shared memory machine, but require the specification of the distribution of the program's data. This data distribution is then used to guide the process of restructuring the code into an SPMD (Single Program Multiple Data) program for execution on the target distributed memory multiprocessor. The compiler analyzes the source code, translating global data references into local and non-local data references based on the distributions specified by the user. The non-local references are satisfied by inserting appropriate message-passing statements in the generated code. Finally, the communication is optimized where possible, in particular by combining messages and by sending data at the earliest possible point in time.

The Vienna Fortran Compilation System (VFCS) (Chapman et al., 1992) is the state of the art in compiler tools for distributed memory systems, and is a successor to the SUPERB system. It accepts, in the current version, Fortran 77 codes and a specification of the distribution of a program's data in the Vienna Fortran syntax and generates message-passing code for the SUPRENUM, Intel iPSC/860 and Paragon machines. Code may also be generated which makes use of the Parmacs macros, a portable message-passing layer which has been implemented on several different architectures, including workstation clusters, and thus provides a basis for use of the system with a number of machines. Further targets are planned.

In this project we intend to develop a portable version of the resulting code in Vienna Fortran. This will then be compiled for several different target architectures within the VFCS. Use of these existing language extensions and tools will not only permit rapid testing and experimentation with different implementations, it will also be a test of the capabilities of the language and compiler. In particular, we expect that it will enable further evaluation and consideration of the problems posed by large data sets, and drive compiler work to improve the utilization of concurrent filing systems where appropriate.

(d) Image Information Systems

Large images are typically stored as raw data without regard for their likely uses. The data structure is dictated by the manner in which the data were collected, not the manner in which they might be used. An image information system makes the transition from the collection of images to a well-thought-out interaction with images and information extracted from them. The major paradigm is the Picture Archiving and Communication System (PACS, Osteaux, 1992). Usually, however, medical PACS deal with images of a size of 2 to 6 MBytes, and this is small when compared to remote sensing images. However, over a period of a year, the medical PACS easily may accumulate 1 to 2 Terabytes of images.

The proposal to replace a collection of large images by an image information system is new; we are currently unaware of any literature, except that which we referred to earlier under the heading "Interaction with large images".

1.4.2 Research Goals and Methods

We believe that the management and analysis of large images and image data sets deserves considerable research and development. We also believe that parallel computing is on the verge of becoming a mainstream resource, at the expense of continued development of special task-specific hardware. It is therefore our general goal to develop concepts and algorithms to deal with large image data sets, and to investigate the utility of parallel processing architectures for this problem domain. We describe in the following our specific objectives. We then follow with a set of individual steps by which we expect to attain them. This will not be a complete list of methods and steps, but will provide a representative glance at them.

Defining the Objectives

Our general goal consists of numerous smaller elements as follows:

(a) Definition of efficient algorithms for low and intermediate level preprocessing, and of matching of overlapping large active-sensor images in an image data base.

(b) Creation of efficient algorithms to cope with dissimilarities in overlapping images for the purpose of precision matching them.

(c) Development of new algorithms to reconstruct an imaged surface shape from multiple large dissimilar images, and thereby also reconstruct the surface reflective properties.

(d) Concept definitions and demonstrations for interactive information systems of large images.

(e) Development of an understanding for the usefulness of "open parallel computing architectures" in image processing and pattern recognition.

(f) Understanding the relative merits and drawbacks of various parallel computing approaches, e.g. of massively parallel computers, computer clusters or multi-processor computers.

(g) Understanding of concepts, components and applicability of automated conversion of sequential image processing software into parallel form.

(h) Persuasive demonstration of the concepts through a software package for image processing on an open parallel computing system.

(i) Creation of a processing strategy, and demonstration of its validity, to process a very large image data set for surface reconstruction, reconstruction of reflective surface properties, and for conversion from large images into an interactive image information system.

By meeting these objectives we will evolve into a leading force in the understanding of, and development of solutions to, the issues of large images and image data bases.

Attaining the Goals by a Team of Two Austrian Groups

To accomplish our goals, we will perform our research task in a project team with members from two participating institutions, namely the leading group in parallel computing from the University of Vienna and the image processing group at Graz University of Technology. One partner (Graz) will be responsible for the overall task and will contribute the understanding for the problem domain of large images and image data bases, for image processing and pattern recognition, and it will contribute data and the existing software to deal with such data. The other group (Vienna) will contribute knowledge in the use of parallel computers and their various architectures, and will employ its tools to automatically or interactively convert existing software into a parallel environment.

Attaining the Goals by International Cooperation

We are embedded into an international collaboration of groups with an interest in coping with the massive data quantities produced by imaging sensors on board of Earth observing and planetary spacecraft. In fact, the Graz group will become a node in the computer network called the US Planetary Data System (PDS) to provide comfortable and network access to all planetary data currently on file in NASA's archives. Close and project-specific cooperation will especially exist with the signal processing group (Dr. S. Hensley) and the parallel computing team at NASA's Jet Propulsion Laboratory (Dr. P. Li), the INTEL-PARAGON group at the California Institute of Technology (Dr. H. Lorenz-Wirzba), the image data users at Washington University (Prof. R. Arvidson, St. Louis, Missouri), at Brown University (Prof. J. Head, Providence, Rhode Island) and Princeton University (Prof. J. Suppe, Princeton, New Jersey). Once the work gets going, we expect to broaden these collaborations to partners in Europe through existing contacts in academia and at the European Space Agency (ESA); the need exists at that agency to also deal with massive image data sets.

Attaining the Goals by Building on Existing Algorithms and Software

We do not start from scratch, but can build from a large base of existing software for low, intermediate and high level image analysis and image signal processing which was built with the participation of this task's principal investigator while under contract to NASA. This software has been developed either in an applications-neutral manner for use with all kinds of images, or for planetary data. Some of it executes on special hardware that needs to be replaced by an open parallel computer architecture, some of it serves at this time to operate on small data samples in an interactive image analysis environment. Algorithms have been collected into these software systems from various sources, most notably from the remote sensing and planetary image analysis communities. As a result there are three software systems, each consisting of code in the range of about 100 000 lines or more:

signal processing for radar Image Formation Processing;

interactive planetary image analysis tool kit;

digital stereo light table for real-time interaction with large images.

Since these packages are to a large extent experimental, they contain multiple algorithms for some of the tasks. Many of these algorithms are preliminary only and must be replaced by improvements which we hope to create in the proposed task. We address two examples of algorithms in the following two subsections to provide a flavor of a fairly mature system (Subsection 5.2.5) and of a rather preliminary set of algorithms (Subsection 5.2.6).

Software Example "Image Formation Processor"

(a) Purpose of the Software

The image formation processor (IFP) is a software system to convert raw synthetic aperture radar (SAR) video signals to a geometrically and radiometrically corrected image product. The resulting output image is representative of the backscattered power per unit surface area of the imaged surface, with each measurement placed in the image in the proper location for the given object coordinate system. The IFP compensates for phase variations in the electro-magnetic pulses of the raw SAR signal, which is dispersed in both the along- and cross-track dimensions of the moving sensor-platform (the satellite).

The cross-track dispersion is designed into the radar system to achieve a high signal to noise ratio without a large peak power. This is implemented by coding the transmitted pulse (e.g. a binary phase code for Magellan's radar sensor) and performing a matched filtering process to collapse the dispersed pulse into an impulse representative of the point target. This matched filtering is performed in the image formation processor.

The along-track dispersion results from the target being in the near field of the synthetic array. To form the image of a point target, the quadratic phase variation of the near-field target must be compensated. This is performed in the IFP by applying a range dependent filter function that compensates for the phase variation so that all returns from the target can be added coherently.

Typically, performing these operations which effectively "focus" the SAR image, requires on the order of 400-500 floating point operations per input data sample. For Magellan, where there are on the order of 1.0 Giga-samples per orbit and over 5 000 orbits of data, the net result is a massive processing task.

(b) Algorithm

This algorithm is commonly known as SPECAN (for spectral analysis) and consists of 3 main parts:

* Range Pulse Compression: This module operates on each echo line to collapse the time dispersed pulse to an impulse. This is a matched filter operation that can be performed by a convolution in the time (or space) domain, or as a multiplication in the frequency domain (which is most computationally efficient). For the frequency domain approach, the processing consists of two FFTs and a complex multiply. Typically, the set of reference functions are pre-generated for this purpose.
* Azimuth Compression (synthetic aperture formation): This is a similar operation to the range compression where a reference function is multiplied time the data (complex multiply in the along track direction). The reference function is derived from the relative sensor-target position using the nominal pointing geometry and the platform tracking data.
* Radiometric and Geometric Corrections: The resulting imagery from the range and azimuth compression's is distorted and must be resampled to a map projection. This resampling should use the terrain information to minimize distortion resulting from surface relief. It is essentially a two dimensional complex interpolation. Additionally, a radiometric correction is needed to compensate for range dependent distortions from the antenna pattern and the resolution cell variation.

The software is a product of many person-years of NASA research in advanced radar signal processing and executes on hardware especially built for this purpose. Major elements have recently been ported into a work station computer to "publish" the algorithm for use by a small set of pre-qualified scientists. However, the software is not capable of processing anything more than some small test data. Reasonable data throughput will require that some form of parallel architecture and computing is being used.

Algorithm Example "Shape-from-Shading" and Stereopsis

(a) Purpose of the Algorithm and Current Status

The current status of surface reconstruction is based on a sequence of two-image stereopsis and two-image shape-from-shading. Stereo image matching is based on one of three algorithms without, thus far, an authoritative assessment of a "best" one. The algorithms are all hierarchical, i.e. operating on a resolution pyramid, and matching by looking for an extremum of the cross-correlation function in the vicinity of a match candidate. They differ in the choice of windows for the matches, in the strategy for predicting likely match points, in the use of interest operators, and in the rejection methods to weed bad matches.

Dissimilarities between overlapping images are being ignored; their damaging effect on match quality must be accepted. To minimize these effects, dissimilarities are kept to a minimum by use of image overlaps that are as similar as possible. Image pairs with larger dissimilarities are not being matched.

Once the stereo process is completed, the two contributing images are modified to accomplish geometric superposition to within the quality of the matches. In addition, the surface shape is being computed and an elevation measure is assigned to each image pixel, resulting in four values at each pixel location, namely surface reflection for each of the two images, surface slope and surface elevation.

These four values are now being entered into a shape-from-shading process. In essence the known estimate of surface shape is being used to "simulate" two images as if they had been illuminated identically to the two actual image configurations. Two reflective values are being computed at each pixel and compared to the actually observed ones. Differences are being used to change the surface shape so that the surface and two images become consistent with one another under the known illumination. This approach is not being performed in the spatial but spectral domain. Heavy use is therefore being made of Fast Fourier Transforms (FFTs) of the entire data set for conversion between spatial and spectral representations.

(b) Approaches for Improved Algorithms

In contrast to IFP-software (Subsection 5.2.5) we are faced here not only with the need to develop processing strategies on open parallel computers, but we must first find methods and algorithms to accomplish the following:

* be able to employ three or more rather than just two images;
* cope with strong dissimilarities among images of the same surface;
* model as unknowns not only the surface shape, but also the surface reflective properties.

We expect that a least squares model should be defined for determining not only the optimum shift between "n" overlapping input images, but also the optimum parameters of a polynomial reflectance curve. This would result in a simultaneous least squares method for computing the stereo solution and the surface backscatter properties at each pixel.

Computing Architecture: Parallelizing

(a) Manual versus Automated Parallelizing

When code is parallelized by hand, the programmer must distribute the program's work and data to the processors which will execute it. One of the common approaches to do so makes use of the regularity of most numerical computations. This is the so-called SIMD or SPMD (Single Instruction or Program Multiple Data) or data parallel model of computation. With this method, the data arrays in the original program are each partitioned and mapped to the processors. This is known as *distributing* the arrays. The specification of the mapping of the elements of the arrays to the set of processors is called the *data distribution* of that program. A processor is then though of as *owning* the data assigned to it; these data elements are stored in its local memory. Now the work is distributed according to the data distribution: computations which define the data elements owned by a processor are performed by it - this is sometimes known as the *owner computes* paradigm. The processors then execute essentially the same code in parallel, each on the data stored locally.

It is, however, unlikely that the code on one processor will run entirely without requiring data which is stored on another processor. Accesses to non-local data must be explicitly handled by the programmer, who has to insert communication constructs to send and receive data at the appropriate positions in the code. This is called *message passing.* The details of message passing can become surprisingly complex: buffers must be set up, and the programmer must take care to send data as early as possible, and in economical sizes. Several issues arise which do not have their counterpart in sequential programming. New types of errors, such as deadlock and livelock, must be avoided. The programmer must decide when it is advantageous to replicate computations across processors, rather than send data. Moreover, for code which is explicitly parallel, debugging is a serious problem.

A major characteristic of this style of programming is that the performance of the resulting code depends to a very large extent on the data distribution selected by the programmer. The data distribution determines not only where computation will take place. It is also the main factor in deciding what communication is necessary. The total cost incurred when non-local data is accessed involves not only the actual time taken to send and receive data, but also the time delay when a processor must wait for non-local data, or for other processors to reach a certain position in the code. Note that the performance of a program can no longer be estimated solely by the amount of computation it comprises: extra computation is not necessarily costly, and the communication delay inherent in a particular data distribution could be prohibitive.

The message-passing programming style requires that the communication statements be explicitly hardcoded into the program. But these statements are based upon the chosen data distribution, and as a result, the data distribution is also implicitly hardcoded. It will generally require a great deal of reprogramming if the user wants to try out different data distributions. The result of this low level style of programming is that the user spends a great deal of time organizing the storage and communication of data. In consequence, the time taken to produce a program is considerably longer than for comparable codes on shared-memory machines. Moreover, once written, the code is hard to modify or improve to run in some other way, even on the same machine. For example, if instead of dividing into square subblocks, the user wanted to experiment with blocking in only one dimension, e.g., blocks of rows or columns, most of the code dealing with specification and communication would have to be modified. We therefore propose to employ an approach with which code can be parallelized efficiently.

(b) Parallel Computing Strategy for Image Formation Processor IFP

The SAR image formation algorithm (Subsection 5.2.5) is most naturally suited to a pipeline architecture, as the processing is a sequential set of operations. However, this processing task is easily parallelized in that the data set can be broken into blocks, where each block of data can be processed by a different processor.

The most obvious approach is to create a two dimensional topology of processors and to divide the data set across these processors with the appropriate overlap between the data blocks. Each processor would then perform all three major operations on its block of data (i.e., range compression, azimuth compression and radiometric and geometric compression) to produce a subset of the desired output image.

(c) Parallel Computing for Shape from Shading (SfS)

The natural way in which SfS is being applied to large images is image block by image block of only 128 x 128 pixels, to avoid an accumulation of random errors. Individual blocks are being computed with a small overlap and then assembled from the individual blocks or tiles into a seamless array of surface shape values.

It is thus obvious that one manner in which SfS would be parallelized is by data segmentation. At the same time each block, or tile, requires iterative refinements of surface shape computations through repeated transfoms from spatial to spectral domains, and vice-versa. These operations can also benefit from smart parallelizing of the individual transform steps.

Other Steps to Attain the Goals

We present in Fig. 1.5 the concept of an image information system as a collection of images in a common object coordinate system. To enable ease of query and retrieval, the images are “cut” into small segments or “tiles” This is a common method in systems for the fast interaction with large data sets, namely to store the data in logical entities that are likely to be queried as an entity. All image segments of a certain tile are stored with that tile so that retrieval *by area* is fast and efficient. This breaks the original organization by data collection time (satellite orbit or sensor path). However, use of images is not by item of data collection, but by object area of interest or by phenomenon to be studied. Also, relationships among images can only be established and used in a common coordinate system. This must be the system of the object, not any particular sensor related system.

Among the two collaborating groups of the University of Vienna and the Graz Technical University, we have access to numerous parallel processing systems, and we can configure several multi-computer systems from groups of work stations. This variety of computing systems will permit us to authoritatively develop a broad view and understanding of image processing on open system parallel machines.

We expect great benefits from the collaboration within the larger project, of which this current task is but an element. Data compression for the image information system, for example through wavelets, image transforms to support the information extraction, for example with Walsh, Haar and others, statistical methods to describe texture and their use to segment overlapping images prior to stereopsis and SfS, are all expected to benefit the increase of our understanding of the current problem domain.

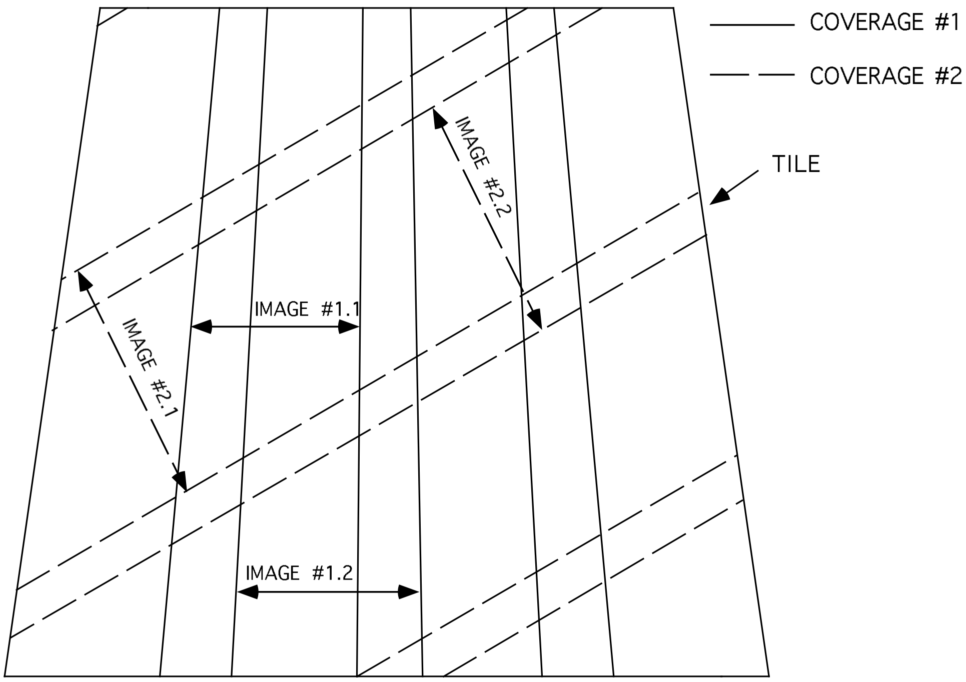


Figure 1.5: Concept of an Image Information System's Tiling Structure. All Data for a Tile (Geographic Area) Are Collected in this Tile.

Bibliography

[AGP91] Y. Amit, U. Grenander, and M. Piccioni. Structural image restoration through deformable templates. *Amer. Stat. Ass.*, 86:376–387, 1991.

[Amm88] Maan Ammar. Off-line preprocessing and verification of signatures. *Int. Journal of Patern Recognition and Artificial Intelligence*, 2 (4):589–602, 1988.

[Amm89a] Maan Ammar. Applications of signature analysis by computer and the consequence of its possible missuse. *Proc. 5^{th} Int. Conf. on Image Analysis and Processing (5ICIAP)*, pages 535–542, 1989.

[Amm89b] Maan Ammar. *Feature extraction and selection for simulated signature verification.* World Scientific Publ. Co., 1989.

[Amm91] Maan Ammar. *Progress in verification of skillfully simulated handwritten signatures*, pages 337–352. World Scientific Publishing Co., 1991.

[Anu70] P. Anuta. Spatial Registration of Multispectral and Multitemporal Digital Imagery Using FFT Techniques. *IEEE Trans. Geoscience Electronics*, GE8, 1970.

[AR75] N. Ahmed and K.R. Rao. *Orthogonal Transforms for Digital Signal Processing.* Springer, Berlin, 1975.

[AS89] R.K. Asbury and D.S. Scott. Fortran I/O on the iPSC/2: Is There Read After Write? In *Proc. of the DMCC 4*, pages 129–132, 1989.

[AYF86] Maan Ammar, Yuuji Yoshida, and Teruo Fukumura. A new effective approach for automatic off-line verification of signatures by using pressure features. *Proc. 8^{th} Int. Conf. on Image Analysis and Processing (8ICIAP)*, pages 566–569, 1986.

[BAR92a] Jezekiel Ben-Arie and K. Raghunath Rao. Novel approach to template recognition by image expansion with non-orthogonal wavelets. *SPIE Visual Communications and Image Processing*, 1818:991–1003, 1992.

[BAR92b] Jezekiel Ben-Arie and K. Raghunath Rao. Template Recognition by Image Expansion with Non-Orthogonal Wavelets. *IEEE Int. Conf. on SMC*, 1992.

[Bas93] M. J. Bastiaans. Gabor's Signal Expansion and Its Relation to Sampling of the Sliding-Window Spectrum. In R. J. Marks II, editor, *Advanced Topics in Shamnnon Sampling and Interpolation Theory*, pages 1–36. Springer Verlag, 1993.

[Bat81] E. Batschelet. *Circular Statistics in Biology.* Academic Press, London, 1981.

[BCF76] G. Bongiovanni, P. Corsini, and G. Forsini. One-dimensional and Two-dimensional Generalized Discrete Fourier Transforms. *IEEE Trans. Acoustics, Speech, and Signal Processing*, 24:97–99, feb 1976.

[BCZ92] S. Benkner, B. Chapman, and H. Zima. Vienna Fortran 90. In *Proc. of the Scalable High Performance Computing Conference 1992*, pages 51–59, Los Alamitos, April 1992.

[BD89] M. G. Beaty and M. M. Dodson. Derivative sampling for multiband signals. *Numerical Functional. Analysis and Optimization*, 9 & 10:179–195, 1989.

[BE89] P.L. Butzer and W. Engels. Background to an Extension of Gibbs Differentation in Walsh Analysis. In P.L. Butzer and R.S. Stankovic, editors, *Proc. of the 1st International Workshop on Gibbs Derivatives*, pages 19–57, Kupari-Dubrovnik, Yugoslavia, 1989.

[Bea84] K.G. Beauchamp. *Applications of Walsh and Related Functions.* Academic Press, London, 1984.

[Bes75] J. Besag. Statistical analysis of lattice data. *The Statistican*, 24:179–195, 1975.

[Bet84] T. Beth. *Verfahren der schnellen Fourier-Transformation.* Teubner, Stuttgart, 1984.

[Bey90] G. Beylkin. Fast wavelet transforms. *SIAM Conf. Chicago 1990*, 1990.

[BGMH92] P. Brezany, H.M. Gerndt, P. Mehrotra, and H.Zima. Concurrent File Operations in a High Performance FORTRAN. In *Proc. of Supercomputing 92*, pages 230–237, Los Alamitos, November 1992.

[BH] J. Benedetto and W. Heller. Irregular sampling and the theory of frames, I. Dedicated to the Memory of Prof. G. Köthe. to appear.

[BH88] P. L. Butzer and L. Hinsen. Reconstruction of bounded signals from pseudo-periodic irregularly spaced samples. *Signal Processing 17*, pages 1–17, 1988.

[BH89] P. L. Butzer and L. Hinsen. Two-dimensional nonuniform sampling expansions - An iterative approach. *I, II. Appl. Anal. 32*, pages 53–68 and 69–85, 1989.

[Bia59] H. Bialy. Iterative Behandlung Linearer Funktionalgleichungen. *Arch. Rat. Mech. Anal.*, pages 166–176, 1959. nicht aufnehmen.

[Boo78] F. L. Bookstein. *The measurement of biological shape and shape change*, volume 24 of *Lecture Notes in Biomathematics.* Springer, 1978.

[BP92] Marco Bonetti and Mauro Piccioni. Deformable Templates: A Statistical Setting. Technical Report 9, Dipartimento di Matematica Pura ed Applicata, Universita di L'Aquila, Italy, 1992.

[BR82] J. Bruner and A.P. Reeves. An Image Processing System with Computer Network Distribution Capabilities. In *Proc. of the International Conference on Pattern Recognition*, pages 447–450. IEEE, 1982.

[Bra84] R.N. Bracewell. The Fast Hartley Transform. *Proc. IEEE*, 72(8):1010–1018, 1984.

[Bra86] R. Bracewell. *The Fourier transform and its application.* Mc-Graw Hill, second edition, 1986. Chapter 10: Sampling series).

[Bro85] E. Brocklehrust. Computer methods for signature verification. *Journal of Forensic Science Society*, 25:445–457, 1985.

[BS] P. L. Butzer and R. L. Stens. Linear Prediction by Sampling from the Past. In R. J. Marks II, editor, *Advanced Topics in Shannon Sampling and Interpolation Theory*, pages 157–184. Springer Verlag.

[BSS88] P. L. Butzer, W. Splettstößer, and R. L. Stens. The sampling theorem and linear prediction in signal analysis. *Jahresbericht des Deutschen Mathematischen-Vereins 90*, pages 1–70, 1988.

[Bur88] W. Burger. Programmierung von Computer Vision Systemen durch graphische Instruktion. In F. Pichler and A. Pinz, editors, *Statistik und Mustererkennung.* Oldenbourg, Wien, 1988.

[Bur89] W. Burger. On Interactive Programming of Computer Vision Systems. In R. Albrecht and A. Pinz, editors, *Wissensbasierte Mustererkennung.* Oldenbourg, Wien, 1989.

[But83] P. L. Butzer. A survey of the Whittaker-Shannon sampling theorem and some of its extensions. *J. Math. Res. Expositions 3*, pages 185–212, 1983.

[BW72] P.L. Butzer and H.J. Wagner. On a Gibbs Type Derivative in Walsh-Fourier Analysis with Applications. In R.E. Horton, editor, *Proc. of the 1972 National Electronics Conference*, pages 393–398, Oak Brook, 1972.

[CBB+92] B. Chapman, S. Benkner, R. Blasko, P. Brezany, M. Egg, T. Fahringer, H.M. Gerndt, J. Hulman, B. Knaus, P. Kutschera, H. Moritsch, A. Schwald, V. Sipkova, and H. Zima. *Vienna Fortran Compilation System Version 1.0 – User's Guide.* Vienna, 1992.

[CC87] F. S. Cohen and D. B. Cooper. Simple parallel hierarchical and relaxation algorithms for segmenting noncausal Markovian random fields. *IEEE Transactions PAMI*, 9:195–219, 1987.

[CC90] D. Cochran and J. J. Clark. On the sampling and reconstruction of time-warped band-limited signals. *ICASSP 1990: D*.2*, pages 1539–1541, 1990.

[CECS80] D. Cooper, H. Elliot, F. Cohen, and P. F. Symosek. Stochastic boundary estimation and object recognition. *Comput. Graphics Image Processing*, 12:326–355, 1980.

[CFM+92] C. Cenker, H. G. Feichtinger, M. Mayer, H. Steier, and Strohmer T. New variants of the POCS method using affine subspaces of finite codimension, with applications to irregular sampling. *Conf. SPIE 92 Boston*, page , 1992.

[CGK91] Y. Chow, U. Grenander, and D. M. Keenan. *Hands: A Pattern Theoretical Study of Biological Shape.* Springer Verlag, New York, 1991.

[Chr93] O. Christensen. *Frames in Hilbert Spaces.* PhD thesis, Universiy of Vienna, 1993.

[Chu92a] C. K. Chui. *Wavelets - A Tutorial in Theory and Applications.* Academic Press, 1992. nicht aufnehmen.

[Chu92b] Charles K. Chui, editor. *Wavelets: A Tutorial in Theory and Applications.* Academic Press, London, 1992.

[CL88] V. Cantoni and S. Levialdi. Multi-Processor Computing for Images. In *Proc. of the IEEE*, volume 76(8), pages 959–969, 1988.

[CL92] M. Chen and J. Li. Optimizing Fortran 90 programs for data motion on massively parallel systems – Technical Report YALE/DCS/TR-882. Technical report, January 1992.

[CMBB90] S. Candela, J.A. Munoz-Blanco, and O. Bolivar. Systemic Approach to Variable Receptive Field Transforms. In R. Trappl, editor, *Proc. of the Tenth European Meeting on Cybernetics and Systems Research*, Vienna, 1990. World Scientific.

[CMQW] R. Coifman, Y. Meyer, S. Quake, and V. Wickerhausen. Signal Processing and Compression with Wave Packets. .

[CMW92] R. Coifman, Y. Meyer, and V. Wickerhausen. Wavelet Analysis and Signal Processing. In Mary Beth Ruskai, editor, *Wavelets and Their Applications.* Jones and Bartlett, 1992.

[CMZ92a] B. Chapman, P. Mehrotra, and H. Zima. *Vienna Fortran - A Fortran language extension for distributed memory systems.* Elsevier Press, 1992.

[CMZ92b] B. M. Chapman, P. Mehrotra, and H. Zima. Programming in Vienna Fortran. *Scientific Programming*, 1(1), 1992.

[Coh91] L. D. Cohen. On active contour models and balloons. *Comp. Vis. Graph. and Image Proc.*, 53:211–218, 1991.

[Coh92] A. Cohen. Wavelets and Digital Signal Processing. In Mary Beth Ruskai, editor, *Wavelets and Their Applications*, pages 105–121. Jones and Bartlett, 1992.

[CPL85] J. J. Clark, M. R. Palmer, and P. D. Lawrence. A transformation method for the reconstruction of functions from nonuniformly spaced samples. *IEEE Trans. ASSP*, 33(4):1151–1165, 1985.

[Cre92] R. Creutzburg, 1992. Referenzliste über Wavelets, 700 References.

[CW89] A. Calway and R. Wilson. A unified approach to feature extraction based on an invertible linear transform. *Proc. 3rd Int. Conf. Imag. Proc. and its Appl.*, pages 651–655, 1989.

[CW93] S. Casey and D. Walnut. Systems of Convolution Equations, Deconvolution, Shanon Sampling and the Wavelet and Gabor Transforms. *preprint*, 1993.

[Dau] I. Daubechies. The wavelet transform, time-frequency localization and signal analysis. *Trans. IEEE Inf. Theory*, ? to appear.

[Dau88a] I. Daubechies. Orthonormal bases of compactly supported wavelets. *Comm. Pure Appl. Math*, 41:909–996, 1988.

[Dau88b] I. Daubechies. Time-frequency localization operators: a geometric phase space approach. *IEEE Trans. Inf. Theory*, 34:605–612, 1988.

[Dau92a] I. Daubechies. Wavelets and Applications. to appear in CBMS-Lecture Notes - nicht aufnehmen, 1992.

[Dau92b] Ingrid Daubechies. *Ten Lectures on Wavelets.* CBMS-NFS Regional Conference Series in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1992.

[dBDR92] C. de Boor, R. DeVore, and A. Ron. On the reconstruction of multivariate (pre)wavelets. Technical Summary Report, 1992.

[dBDR93] C. de Boor, R. DeVore, and A. Ron. The structure of finitely generated shift-invariant spaces in $L_2(\mathbb{R}^d)$. *preprint*, 1993.

[Der79] E. J. Derp. Image data compression using autoregressive time series models. *Pattern Recognition*, 11:313–323, 1979.

[DGM86] I. Daubechies, A. Grossmann, and Y. Meyer. Painless non-orthogonal expansions. *J. Math. Phys.*, 27:1271–1283, 1986.

[DH91] P. Deuflhard and A. Hohmann. *Numerische Mathematik.* de Gruyter, 1991.

[DJ92] David L. Donoho and Iain M. Johnstone. Minimax estimation via wavelet shrinkage. Technical report, Department of Statistics, Stanford University, 1992.

[DJJ89] I. Daubechies, St. Jaffard, and J. Journe. A simple Wilson orthonormal basis with exponential decay. *preprint*, 1989.

[DJP] R. DeVore, B. Jawerth, and V. Popov. Compression of wavelet decompo *preprint*, .

[Don92a] David L. Donoho. De-Noising by Soft-thresholding. Technical report, Department of Statistics, Stanford University, 1992.

[Don92b] David L. Donoho. Ideal spatial adaption by wavelet shrinkage. Technical report, Department of Statistics, Stanford University, June 1992.

[Don92c] David L. Donoho. Interpolating Wavelet Transforms. Technical report, Department of Statistics, Stanford University, October 1992.

[Don93] David L. Donoho. Unconditional Bases are Optimal Bases for Data Compression and for Statistical Estimation. Technical report, Department of Statistics, Stanford University, 1993.

[DS52] R. Duffin and A. Schaeffer. A class of nonharmonic Fourier series. *Trans. Amer. Math. Soc.*, 72:341–366, 1952.

[DS89a] D. Donoho and P. Stark. Recovery of a sparse signal when the low frequency information is missing. Technical Report 179, Dep. Stat., UCB, June 1989.

[DS89b] D. Donoho and P. Stark. Uncertainty principles and signal recovery. *SIAM J. Appl. Math*, 48(3):906–931, 1989.

[Duf83] M.J.B. Duff. *Computing Structures for Image Processing.* Academic Press, 1983.

[Eat89] M. L. Eaton. Group Invariance Applications in Statistics. *CBMS Reg. Conf. Series in Prob. and Stat.*, 1989.

[Ede68] M. Eden. *Handwriting generation and recognition*, pages 138–154. MIT Press, Cambridge, 1968.

[Eve85] M. Evett. A study of the form and extent of natural variation in genuine writings with age. *Journal of Forensic Science Society*, 25:207–215, 1985.

[FC87] R. Frankot and R. Chellappa. Application of Shape from Shading Technique to Synthetic Aperture Radar. In *Proc. IGARSS Symposium*, Ann Arbor, 1987.

[FCH91] H. G. Feichtinger, C. Cenker, and M. Herrmann. Iterative algorithms in irregular sampling: A first comparison of methods. In *Conf. ICCCP'91, March 1991, Phoenix, USA*, pages 483–489, 1991.

[FCS91] H. G. Feichtinger, C. Cenker, and H. Steier. Fast iterative and non-iterative reconstruction methods in irregular sampling. *Conf. ICASSP'91, May, Toronto*, 1991.

[Fed77] V. Fedorov. *Experimental design.* Wiley, 1977.

[Fei] H. G. Feichtinger. Reconstruction of band-limited signals from irregular samples, a short summary. *2. International Workshop on Digital Image Processing and Computer Graphics with Applictions.* ÖCG, 1991.

[Fei89] H. G. Feichtinger. Coherent frames and irregular sampling. *Proc. Conf. Recent Advances in Fourier Anal. and Its Appl.*, NATO ASI Series C, Vol. 315:427–440, 1989. NATO conference, Pisa.

[Fei91] H. G. Feichtinger. Pseudo-inverse matrix methods for signal reconstruction from partial data. In *SPIE-Conf., Visual Comm. and Image Proc., Boston*, pages 766–772, 1991. Int.Soc.Opt.Eng.

[Fel82] H. Fellner. Realisierung der allgemeinen Faltung durch die schnelle verallgemeinerte Fouriertransformation. Master's thesis, Johannes Kepler Universitaet, Linz, 1982.

[FG] H.G. Feichtinger and K.H. Groechenig. Non-Orthogonal wavelet and Gabor expansions, and group representations. In I.Daubechies et al. G.Beylkin, R.Coifman, editor, *Wavelets and their Applications*. Jones and Bartlett, .

[FG89] H. G. Feichtinger and K. Groechenig. Multidimensional irregular sampling of band-limited functions in L^p-spaces. *Conf. Oberwolfach Feb. 1989*, pages 135–142, 1989. ISNM 90 (1989), Birkhäuser.

[FG92a] H. G. Feichtinger and K. Groechenig. Irregular sampling theorems and series expansions of band-limited functions. *J. Math. Anal. Appl*, 167:530–556, 1992.

[FG92b] H. G. Feichtinger and K. Groechenig. Iterative reconstruction of multivariate band-limited functions from irregular sampling values. *SIAM J. Math. Anal. 231*, pages 244–261, 1992.

[FG92c] H.G. Feichtinger and K.H. Groechenig. Gabor wavelets and the Heisenberg group: Gabor Expansions and Short Time Fourier transform from the group theoretical point of view. In C.K. Chui, editor, *Wavelets – A Tutorial in Theory and Applications*, pages 359–397. Academic Press, 1992.

[FG93a] H. G. Feichtinger and K. Groechenig. Error analysis in regular and irregular sampling theory. *Applicable Analysis*, 1993. to appear.

[FG93b] H. G. Feichtinger and T. Groechenig, K.and Strohmer. Superfast Reconstruction Algorithms in Irregular Sampling Theory. 1993. to appear.

[FG93c] H.G. Feichtinger and K.H. Groechenig. Theory and Practice of Irregular Sampling. In Benedetto J. and Frazier M., editors, *Wavelets:Mathematics and Applications*. CRC Press, 1993. to appear.

[FHK+91] G. Fox, S. Hiranandani, K. Kennedy, C. Koelbel, U. Kremer, C. Tseng, and M. Wu. *Fortran D language specification*, March 1991.

[FJ85] M. Frazier and B. Jawerth. Decomposition of Besov spaces. *Indiana Univ. Math. Journal*, 34:777–799, 1985.

[FJ90] M. Frazier and B. Jawerth. A discrete transform and decomposition of distribution spaces. *Journal Functional Analysis*, 93:34–170, 1990.

[FLM86] K. Fulberton, F. Leberl, and R. Marque. Opposite Side SAR Stereo. 52:1487–1498, 1986.

[Foe82] W. Foejekt1rstner. On the Geometric Precision of Digital Correlation. *International Archives of Photogrammetry*, 24(III):176–189, 1982.

[For93] High Performance Fortran Forum. High Performance Fortran Language Specification – Version 1.0. Technical report, Houston, May 1993.

[FPC+93] H. G. Feichtinger, G. Pflug, C. Cenker, D. Schaffer, and A. Gritsevski. Mathematical methods and tools for digital image processing: Stochastic shape theory. *FWF Project 8784*, 1992/93.

[FS92] H. G. Feichtinger and T. Strohmer. IRSATOL – Irregular Sampling of Band-limited Signals TOOLBOX. In K. Dette, D. Haupt, and C. Polze, editors, *Conf. Computers for Teaching, Berlin*, pages 277–284, 1992.

[FS93] H. G. Feichtinger and T. Strohmer. Fast Iterative Reconstruction of Band-Limited Images from Irregular Sampling Values. 1993. to appear.

[Fuk90] Keinosuke Fukunaga. *Introduction to Statistical Pattern Recognition.* Academic Press, 1990.

[Gab46] D. Gabor. Theory of Communication. *Journal Inst. Elect. Eng*, 93:429–457, 1946.

[Gab88] A. Gabe. Stationäre Gaußprozeße in der Ebene und ihre Anwendung in der Statistischen Bildanalyse. Master's thesis, Universität Gießen, Germany, 1988.

[Ger89] H. M. Gerndt. *Automatic Parallelization for Distributed-Memory Multiprocessing Systems.* PhD thesis, University of Bonn, December 1989.

[Get92] Neil Getz. A Fast Discrete Periodic Wavelet Transform. Technical report, Electronics Research Laboratory, College of Engineering, University of California,

[GG83] D. Geman and S. Geman. Parameter estimation for some Markov random fields. *Complex Systems*, 11, 1983.

[GG84] D. Geman and S. Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. and Machine Intelligence*, 6:721–741, 1984.

[Gid85] B. Gidas. Nonstationary Markov chains and convergence of the annealing algorithm. *J. Statist. Phys.*, 39:73–131, 1985.

[GJ84] A. Grossmann and Morlet. J. Decomposition of functions into square integrable wavelets of constant shape. *SIAM J. Math.*, 15:723–736, 1984.

[GK86] U. Grenander and D. M. Keenan. On the shape of plain images. Technical report, Div. Appl. Math, Brown University, R.I., 1986.

[GK91] U. Grenander and D. M. Keenan. *HANDS: A pattern theoretic study of biological shapes*, volume 2 of *Research Notes in Neural Computing.* Springer, 1991.

[GM86] B. Guindon and H. Marmyama. Automated Matching of Real and Simulated SAR Imagery as a Tool for Ground Control Point Requisition. *Can. J. Remote Sensing*, 12(4):149–158, 1986.

[GM92] U. Grenander and K. Manbeck. A stochastic shape and color model for defect detection in potatoes. *Rep. Patt. Th.*, 156, 1992.

[GMP85] A. Grossmann, J. Morlet, and T. Paul. Transforms associated to square integrable group representations, I: general results. *J. Math. Phys.*, 26:2473–2479, 1985.

[Goe90] W. Goetz. Entwicklung und Implementierung von Verfahren zur Digitalen Bildverarbeitung mittels Walsh-Funktionen. Master's thesis, Johannes Kepler Universitaet, Linz, Oct 1990.

[Gre76a] U. Grenander. *Pattern Analysis*, volume 1 of *Lectures in Pattern Theory.* Springer Verlag, New York, 1976.

[Gre76b] U. Grenander. *Pattern Synthesis Vol I.* Applied Mathematical Sciences. Springer Verlag, New York, 1976.

[Gre77] U. Grenander. *Pattern Analysis Vol. II.* Applied Mathematical Sciences. Springer Verlag, New York, 1977.

[Gre81a] U. Grenander. *Lectures in Pattern Theory*, volume 1–3. Springer Verlag, Berlin, 1976, 1978, 1981.

[Gre81b] U. Grenander. *Abstract Inference.* Academic Press, New York, 1981.

[Gre89] U. Grenander. Advances in Pattern Theory. *Annals Statist.*, 17(1):1–30, 1989.

[Gro] K. Groechenig. Sharp results on random sampling of band-limited functions. *preprint*, .

[Groa] K. Groechenig. A discrete theory of irregular sampling. *Lin. Alg. and Appl.* to appear.

[Grob] A. Grossmann. Wavelet Transforms and Edge Detection. to appear in Stochastic Processes in Physics and Engineering.

[Gro90] K. Groechenig. A new approach to irregular sampling of band-limited functions. In J.S. and J.L. Byrnes, editors, , volume 315 of *NATO ASI Series 'Recent Advances in Fourier Analysis and Its Applications', Series C*. Kluwer Acad. Publ., 1990.

[Gro91a] K. Groechenig. Describing functions: atomic decompositions versus frames. *Monatshefte für Mathematik*, 112:1–41, 1991.

[Gro91b] K. Groeochenig. Efficient algorithms in irregular sampling of band-limmited functions. In *Proceedings of the* 10^{th} *annual ICCCP, March '91, Scottsdale, Arizona, USA*, pages 490–495. IEEE Computer Society Press, 1991.

[Gro92] K. Groechenig. Reconstruction algorithms in irregular sampling. *Comput. Math*, 59:181–194, 1992.

[Groar] K. Groechenig. Acceleration of the frame algorithm. *IEEE Trans. SSP*, to appear.

[Gru89] A. Gruen. Digital Photogrammetric Systems: Current Status and Prospect. *Photogrammetric Eng. and Remote Sensing*, 55:581–586, 1989.

[GS85] U. Grenander and J. Sethuraman. Limit theorems in metric pattern theory. private communication, 1985.

[GW87] R.C. Gonzalez and P. Wintz. *Digital Image Processing.* Addison-Wesley, Reading, Mass., 2 edition, 1987.

[GW91] R.C. Gonzales and R.E. Woods. *Digital Image Processing.* Addison-Wesley, 1991.

[Hei90] C. Heil. *"Wavelets and frames" in Signal Analysis.* IMA, Springer Verlag, 1990.

[Hel89] H. Hellwagner. *Systolische Architekturen fuer die Verallgemeinerte Diskrete Fourier Transformation.* PhD thesis, Johannes Kepler Universitaet, Linz, 1989.

[Hes56] M.R. Hestenes. The conjugate-gradient method for solving linear systems. In J. Curtiss, editor, *Numerical Analysis.* McGraw-Hill, 1956.

[Hig76] J. R. Higgins. A sampling theorem for irregularly sampled points. *IEEE Trans. Information Theory*, pages 621–622, Sep. 1976.

[HK74] E. F. Harding and D. K. Kendall, editors. *Stochastic Geometry.* 1974.

[HK91] Y. He and A. Kundu. 2-d shape classification using hidden Markov model. *IEE Trans. PAMI*, 13 (11):1172–1184, 1991.

[Hor89] B. Horn. *Shape from Shading.* MIT-Press, Cambridge, Mass., 1989.

[Hou92] Christian Houdré. Wavelets, probability and statistics: some bridges. Technical report, Department of Statistics, Stanford University, October 1992.

[HP80] W. J. Hale and B. J. Paganini. An automatic personal verification system based on signature writing habits. *Proc. of the Carnahan Conference on Crime and Countermeasures*, pages 121–125, 1980.

[HS52] M.R. Hestenes and E.L. Stiefel. Methods of conjugate gradients for solving linear systems. *Nat.Bur.Std.J.Res.*, 49:409–436, 1952.

[HS92] R. Haralick and L. Shapiro. *Computer and Robot Vision.* Addison-Wesley, 1992.

[Huc92] T. Huckle. Some Aspects of Circulant Preconditioners. *preprint*, 1992.

[Hur89] N. E. Hurt. *Phase Retrival and Zero Crossings, Mathematical Methods in Image Reconstruction.* Kluwer Acad. Publ., 1989.

[HW89] C. Heil and D. Walnut. Continuous and discrete wavelet transforms. *SIAM Rev.*, 31(4):628–666, 1989.

[HW90] C. Heil and D. Walnut. Gabor and Wavelet Expansions. In J.S. Byrnes and J.F Byrnes, editors, *Recent Advances in Fourier Analysis and Its Applications*, pages 441–454. Kluwer Academic Publishers, 1990.

[HY81] L.A. Hagman and M. Young. *Applied Iterative Methods.* Academic Press, 1981.

[HZM92] M. Holden, M. Zemerly, and J.-P. Mueller. *Parallel Stereo and Motion Estimation.* John Wiley & Sons Ltd., 1992.

[Jer77] A. J. Jerri. The Shannon sampling theorem - its various extensions and applications, a tutorial review. *Proc. IEEE*, 65:1565–1596, 1977.

[JS93] Björn Jawerth and Wim Sweldens. An Overview of Wavelet Based Multiresolution Analyses. Technical report, Department of Mathematics, Department of Computer Science, University of South Carolina, Columbia, SC 29208, Universiteit Leuven, B 3001 Leuven, Belgium, February 1993.

[Kar77] M.G. Karpovsky. Some Optimization Problems for Convolution Systems over Finite Groups. *Information on Control*, 34:227–247, 1977.

[Kar87] A.H. Karp. Programming for Parallelism. *Computer*, 20(5):43–57, May 1987.

[Kas90] A. Kastner. Implementierung von Spektraltransformationen fuer die digitale Sprach- und Bildverarbeitung. Master's thesis, Johannes Kepler Universitaet, Linz, Jan 1990.

[Kat89] A. K. Katsaggelos. Iterative image restoration algorithms. *Optical Eng.*, 28(7):735–748, 1989.

[KC81] R. L. Kashyap and R. Chellappa. Stochastic models for closed boundary analysis: Representation and reconstruction. *IEEE Transactions on Inf. Theory*, 27:627–637, 1981.

[KCMP90] R. Kwok, J. Curlander, R. McConnell, and S. Pany. An Ice Motion Tracking System at the Alaska SAR Facility. *IEEE Journal of Ocean Engineering*, 15:44–45, 1990.

[KE86] M.C. Karpovsky and E.A.Trachtenberg. Statistical and Computational Performance of a Class of Generalized Wiener Filters. *IEEE Trans. on Info. Th.*, 32(2):303–307, Mar 1986.

[KE89] R. L. Kashyap and K. B. Eom. Texture Boundary Detection Based on the Long Correlation Model. *IEEE Transactions PAMI*, pages 58–67, 1989.

[Ken77] D. G. Kendall. The difffusion of shape. *Adv. in Appl. Prob.*, 9:428–430, 1977.

[Ken89] D. G. Kendall. A survey of the statistical theory of shape. *Statistical Science*, 4:87–120, 1989.

[Kir87] R. L. Kirk. *A Fast Finite Element Algorithm for Two-Dimensional Photoclinometry.* PhD thesis, Cal. Inst. of Technology, 1987.

[KL87] H. M. Kalayek and D. A. Landgrebe. Stochastic Model Utilizing Spectral and Spatial Characteristics. *IEEE Transaction PAMI*, 9 (3):457–461, 1987.

[KLT+88] W. Kober, F. Leberl, J. Thomas, D. Meyer, and J. Cimino. Requirements for an EOS-Oriented Workstation. In *IGARSS'88*, pages 1741–1747, Edinburgh, 1988.

[KM91] C. Koelbel and P. Mehrotra. Compiling global name-space parallel loops for distributed execution. *IEEE Transactions on Parallel and Distributed Systems*, 2(4):440–451, October 1991.

[Kno88] A. Knoerr. Global models of natural boundaries: Theory and applications. Technical report, Div. Appl. Math, Brown University, R.I., 1988.

[Kri91] A. Krikelis. *Real Time Image Understanding with the Associative String Processor.* Academic Press, 1991.

[KS80] R. Kinderman and J. L. Snell. *Markov Random Fields and their Applications.* Amer. Math. Soc., Providence, RI, 1980.

[KS89] B. Kartikeyan and A. Sarkar. Shape description by Time Series. *IEEE Transaction PAMI*, 11 (9):977–984, 1989.

[KSL92] R.L. Kirk, L.A. Soderblom, and E.M. Lee. Enhanced Visualization for Interpretation of Magellan Radar Data: Supplement to the Magellan Special Issue. *J. Geophysical Research*, 97(E10):16,371–16,380, 1992.

[Kum91] V.K.P. Kumar. *Parallel Architectures and Algorithms for Image Understanding.* Academic Press, San Diego, 1991.

[Kun77] H. Kunz. *Approximation linearer Transformationen durch eine Klasse schneller, verallgemeinerter Fouriertransformationen.* PhD thesis, ETH (5832), Zuerich, 1977.

[KWT87] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: active contour models. *Int. Jour. Comp. Vis.*, 1:321–331, 1987.

[KY73] J. Kittler and P. Young. An new applicaton to feature selection based on the Karhunen Loeve expansion. *Pattern Recognition*, 5:335–352, 1973.

[Lam92] J. Lamoure. Avantages et Contrainites de la Stereophotogrammetrie Numerique. Matra MS2i, 1992.

[Lan80] D. A. Landgrebe. The development of spectral/spatial classifier for earth observational data. *Pattern Recognition*, 12:16–175, 1980.

[LaP63] G. LaPrade. An Analytical and Experimental Study of Stereo for Radar. *Photogrammetric Engineering*, 1963.

[LBE+79] F. Leberl, M.L. Bryan, Ch. Elachi, T. Farr, and W. Campbell. Mapping of Sea Ice and Measurement of Its Drift Using Aircraft Synthetic Aperture Radar Images. *J. of Geophysical Research*, 84/094:1827–1835, 1979.

[LC92] P.P. Li and D. Curkendall. Parallel Three-Dimensional Perspective Rendering. In *Proc. European Workshop on Parallel Computing*, Barcelona, 1992.

[LD89] S. Lakshmanan and H. Derin. Simultaneous Parameter Estimation and Segmentation of Gibbs Random Fields Using Simulated Annealing. *IEEE Transactions PAMI*, 11 (8):799–813, 1989.

[Lea88] R. Lea. The ASP: A Cost-Effective Parallel Microcomputer. In *IEEE-Mic.*, pages 10–29, 1988.

[Leb75] F. Leberl. Lunar Radargrammetry with ALSE-VHF Imagery. In *Proc. American Society of Photogrammetry*, pages 266–285, Phoenix, 1975.

[Leb76] F. Leberl. Mapping of Lunar Surface from Side-Looking Orbital Radar Images. *The Moon*, 15:329–342, 1976.

[Leb79] F. Leberl. Accuracy Analysis of Stereo Side-Looking Radar. *Photogrammetric Engineering and Remote Sensing*, 45:1083–1096, 1979.

[Leb80] F. Leberl. Die Erforschung der Oberfläche des Planeten Venus. *Mitteilungen de Geodätischen Institute der Techn. Univ. Graz*, 35, 1980.

[Leb89] F. Leberl. *Radargrammetric Image Processing.* Artech House, Norwood, MA, 1989.

[Leb93] F. Leberl. How To Extract Venus Topographic Information From Magellan Radar Images. In *SPIE Proceedings*, volume 1943, 1993.

[LED92] F. Leberl, H. Ebner, and I. Dowman, eds. Softcopy Photogrammetric Workstations. *Photogrammetric Engineering and Remote Sensing*, Special Issue, January 1992.

[Len90] R. Lenz. *Group Theoretical Methods in Image Processing*, volume 413 of *Lect. Notes in Comp. Sc.* Springer Verlag, Berlin, 1990.

[LG76] F. Leberl and A.F.H. Goetz. Übersicht über den gegenwärtigen Stand der digitalen Bildverarbeitung. *Österr. Zeitschrift für Vermessungswesen und Photogrammetrie*, 63(3):101–126, 1976.

[Lie64] R. Liedl. Ueber eine spezielle Klasse von stark multiplikativ orthogonalen Funktionensystemen. *Monatshefte fuer Mathematik*, 68(2):130–137, 1964.

[LJ92] K. Leung and J. Jin, et al. SAR Data Processing for the Magellan Prime Mission. In *Proc. IGARSS'92*, pages 1301–1304, Houston, Texas, 1992.

[LK93] D.F. Leotta and Y. Kim. Requirements for Picture Archiving and Communications. *IEEE Engineering in Medicine and Biology*, 12(1):62–69, 1993.

[LKB93] F. Leberl, R. Kwok, and T. Baltzer. Sea Ice Mapping With ERS-1 Images. In *Proc. of the 25th International Symposium on Remote Sensing and Global Environmental Change*, Ann Arbor, MI, 1993. ERIM.

[LMT+92] F. Leberl, K. Maurice, J. Thomas, C. Leff, and S. Wall. Images and Topographic Relief at the North Pole of Venus. *J. of Geophysical Research*, pages 13667–13674, 1992.

[LMTK91] F. Leberl, K. Maurice, J. Thomas, and W. Kober. Radargrammetric Measurements from the Initial Magellan Coverage of Planet Venus. *Photogrammetric Engineering and Remote Sensing*, 57(12):1561–1570, 1991.

[LMTM93] F. Leberl, K. Maurice, J. Thomas, and M. Millot. Automated Radar Image Matching Experiment. *J. of Photogrammetry and Remote Sensing*, 1993.

[LMW+91] F. Leberl, M. Millot, R. Wilson, M. Karspeck, B. Mercer, and S. Thornton. Radargrammetric Image Processing with a Softcopy Stereo Workstation. In *Proceedings, 8th Thematic Conference on Geologic Remote Sensing*, pages 639–647, Denver, CO, 1991.

[Lov92] D. Loveman. High Performance Fortran: Proposal. In *High Performance Fortran Forum*, Houston, TX, January 1992.

[LREC83] F. Leberl, J. Raggam, Ch. Elachi, and W. Campbell. Sea Ice Motion Measurements from SEASAT-SAR Images. *J. of Geophysical Research*, 88(C3):1915–1928, 1983.

[LTM92] F. Leberl, J.K. Thomas, and K.E. Maurice. Initial Results From the Magellan Stereo-Experiment. *J. of Geophysical Research*, pages 12675–13687, 1992.

[Mal89a] St. Mallat. Complete signal representation with multiscale edges. *NYU, Computer Science Tech. Report*, 483, Dec. 1989.

[Mal89b] St. Mallat. Multifrequency Channel Decompositions of Images and Wavelet Models. *IEEE Trans. ASSP*, 37:699–717, Dec. 1989.

[Mal89c] St. Mallat. Multiresolution approximation and wavelet orthonormal bases of L^2. *Trans. Amer. Math. Soc.*, 315:69–87, 1989.

[Mal89d] St. Mallat. A theory of multiresolution signal decomposition: the wavelet transform. *IEEE Trans. Pattern Anal. and Mach. Int.*, 11:674–693, 1989.

[Mal91] St. Mallat. Zero Crossings of a Wavelet Transform. *IEEE Trans. Information Theory*, 37:674–693, 1991.

[Mar89] F. A. Marvasti. An iterative method to compensate for the interpolation distortion. *IEEE Trans. ASSP 37/10*, pages 1617–1621, 1989.

[Mar93] F. A. Marvasti. Nonuniform Sampling. In R. J. Marks II, editor, *Advanced Topics in Shannon Sampling and Interpolation Theory*, pages 121–156. Springer Verlag, 1993.

[Mat75] G. Matheron. *Random sets and integral geometry.* Wiley and Sons, New York, 1975.

[May] M. Mayer. *The POCS method and some of its variations.* PhD thesis, Universiy of Vienna. in preparation.

[MC92] F. A. Marvasti and P.M. Clarkson et al. Reconstruction of Speech Signals with Lost Samples. *IEEE Trans. Sig. Proc.*, 40:2897–2903, 1992.

[MCOR88] J.-P. Mueller, K. Collings, G. Otto, and J. Roberts. Stereo Matching Using Transputer Arrays. *International Archives of Photogrammetry and Remote Sensing*, 27(B3):559–586, 1988.

[Mey86] Y. Meyer. "Ondelettes et fonctions splines", Sem. Equ. aux Derivees Partielles. *Ecole Polytechn., Paris, France*, Dec. 1986.

[Mey90a] Y. Meyer. *"Ondelettes et Opérateurs I - Ondelettes.* Hermann, 1990.

[Mey90b] Y. Meyer. *"Ondelettes et Opérateurs II - Operateurs de Calderon-Zygmund.* Hermann, 1990.

[Mey86] Y. Meyer. Principe d'incertitude, bases hilbertiennes et algebras d'operateurs. *Sem. Bourbaki*, 662, 1985/86.

[MI93] R. J. Marks II, editor. *Advanced Topics in Shannon Sampling and Interpolation Theory.* Springer Verlag, 1993.

[MKC+91] R. McConnell, R. Kwok, J. Curlander, S. Pany, and W. Kober. PSI-S Correlation and Dynamic Time Warping: Two Methods for Tracking Ice Floes in SAR Images. *IEEE Trans. Geoscience Remote Sensing*, 29:1004–1014, 1991.

[MKL+89] R. McConnell, W. Kober, F. Leberl, R. Kwok, and J. Curlander. Automated Tracking of Arctic Ice Floes in Multitemporal SAR Imagery. In *Proc. IGARSS'89*, volume 2, pages 1112–1114, 1989.

[MLL88] M. Maresca, M. Lavin, and H. Li. Parallel Architectures for Vision. In *Proc. of the IEEE*, volume 76(8), pages 970–981, 1988.

[MM92] Farzin Mokhtarian and Alan K. Mackworth. A Theory of Multiscale, Curvature Based Shape Representation for Planar Curves. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 14 (8), August 1992.

[MMK+91] M. Millot, K. Maurice, M. Karspeck, W. Kober, and F. Leberl. Generation, Visualization and Analysis of Image Cubes. In *Proceedings, IS&T's 44th Annual Conference*, St. Paul, MN, 1991.

[MR91] P. Mehrotra and J. Van Rosendale. *Programming distributed memory architectures using Kali*, pages 364–384. Advances in Languages and Compilers for Parallel Processing, Pitman/MIT-Press. Elsevier Press, 1991.

[MRH88] C. G. Morris, M. A. Richards, and Hayes M. H. Fast reconstruction of linearly distorted signals. *IEEE Trans ASSP*, 36(7):1017–1025, 1988.

[MRTT53] N. Metropolis, M. Rosenbluth, A. Teller, and E. Teller. Equations of state calculations by fast computing machines. *J. Chem. Phys.*, 21:1087–1091, 1953.

[MZ90] St. Mallat and S. Zhong. Compact image representation from multiscale edges. In *Proceedings of 3. Int. Conf. on Computer Vision*, Dec. 1990.

[MZ92] St. Mallat and S. Zhong. Characerization of Signals from Multiscale Edges. *IEEE Trans. Pattern Anal. and Mach. Int.*, 14:710–732, 1992.

[Neb92] J.A. Nebb. *Machine Independent Parallel Image Processing: Apply and Adapt.* John Wiley & Sons, 1992.

[Nic71] P.J. Nicholson. Algebraic Theory of Finite Fourier Transform. *Journal of Computer and System Sciences*, 5:524–547, 1971.

[Nor92] R. Norvelle. Stereo Correlation: Window Shaping and DEM Corrections. *Photogrammetric Eng. and Remote Sensing*, 58(1):111–115, 1992.

[Nus82] H.J. Nussbaumer. *Fast Fourier Transform and Convolution Algorithms.* Springer, Berlin, 2 edition, 1982.

[Osb86] B. Osborn. *Parameter estimation in pattern theory.* PhD thesis, Div. Appl. Math, Brown University, 1986.

[Ost92] M. Osteaux, editor. *A Second Generation PACS Concept.* Springer-Verlag, 1992.

[Osw] P Oswald. Approximations spaces and mutilevel methods. *preprint*, .

[Osw91] P Oswald. Two Remarks on mutilevel preconditioners. *Forschungsergebnisse*, 91/1(7), 1991.

[Pap66] A. Papoulis. Error analysis in sampling theory. *Proc. IEEE*, 54(7):947–955, 1966.

[Pap77] A. Papoulis. *Signal analysis.* McGraw-Hill, New York, 1977.

[Par70] W. J. Park. A Multiparameter Gaussian Process. *Annals of Mathematical Statistics*, 41:1582–1595, 1970.

[Pas92] D. Pase. *MPP Fortran programming model.* Houston, TX,, January 1992.

[Pav89] M. Pavel. *Foundations of Pattern Recognition.* Dekker, New York, 1989.

[PC93] Georg Ch. Pflug and Christian Cenker. Wavelet regression pattern analysis. *to appear*, 1993.

[PCUG93] Georg Pflug, Christian Cenker, Albert Unterberger, and Andrei Gritsevskii. Erkennung von Unterschriften mit Hilfe statistische Methoden der Mustererkennung. *Project 3974 of the 'Jubiläumsfonds' of the Austrian National Bank*, 1993.

[Pea75] J. Pearl. Optimal dyadic models of time-invariant systems. *IEEE Trans. Comp.*, 24:598–602, 1975.

[PF77] E. Persoon and K. S. Fu. Shape discrimination using Fourier descriptors. *IEEE Transactions Syst. Man. Cybern.*, 7:170–179, 1977.

[PFDSJ89] T.W. Pratt, J.C. French, P.M. Dickens, and Jr. S.A. Janet. A Comparison of the Architecture and Performance of Two Parallel File Systems. In *Proceedings of the DMCC 4*, pages 161–166, 1989.

[Pfl89] Georg Ch. Pflug. Random transformations of the plane and their applications to pattern synthesis and pattern analysis. *Schriftenreihe der OCG, Band 49*, 1989.

[Pfl92] Georg Ch. Pflug. A Statistical Method for the Recognition of Polygonial Planar Shapes. *to appear*, April 1992.

[Pic70] F. Pichler. Walsh-Fourier Synthese optimaler Filter. *Archiv der Elektr. Uebertragung*, 24:350–360, 1970.

[Pic80] F. Pichler. Fast Linear Methods for Image Filtering. In D.G. Lainiotis and N.S. Tzannes, editors, *Applications of Information and Control Systems*, pages 3–11. Reidel, Den Haag, 1980.

[Pic81] F. Pichler. Schnelle Faltung von Bildern. In F. Leberl and H. Ranzinger, editors, *Mustererkennung und Bildverarbeitung in Oesterreich*, pages 13–30. OCG, Wien, 1981.

[Pic89a] F. Pichler. Experiments with 1-D and 2-D Signals using Gibbs Derivatives. In P.L. Butzer and R.S. Stankovic, editors, *Proc. of First International Workshop on Gibbs Derivatives*, pages 181–196, Kupari-Dubrovnik, Yugoslavia, Sep 1989.

[Pic89b] F. Pichler. Some Historical Remarks on The Theory of Walsh Functions and Their Applications in Information Engineering. In P.L. Butzer and R.S. Stankovic, editors, *Proc. of First International Workshop on Gibbs Derivatives*, pages xxv–xxx, Kupari-Dubrovnik, Yugoslavia, Sep 1989.

[Pic92] F. Pichler. Realisierung der Transformation von Prigogine mittels dyadischer Faltungsoperatoren. In R. Trappl and W. Horn, editors, *Research Report of the Austrian Society for Cybernetic Studies*. ISBN 3-85206 127 X, Dec 1992.

[Pie89] P. Pierce. A Concurrent File System for a Highly Parallel Mass Storage Subsystem. In *Proceedings of the DMCC 4*, pages 155–160, 1989.

[Pit92] I. Pitos. *Parallel Algorithms and Architectures for Digital Image Processing*. John Wiley & Sons Ltd., 1992.

[Pit93] S. Pittner. Wavelets - The Bibliography, 1993. 3rd Ed., May.

[Pix89] Pixar. *ELT Technical Summary - Product Technical Description*. S. Rosa, Ca., 1989.

[Plu81] C. Plumeri. *Probability measures on regular structures*. PhD thesis, Brown University, 1981.

[Pos85] A. Possolo. Estimation of binary Markov random fields. Unpublished, 1985.

[Pra72] W.K. Pratt. Generalized Wiener filtering computation technique. *IEEE Trans. Comp.*, 21:636–641, 1972.

[PRS82] E. I. Plotkin, L. M. Roytman, and M. N. S. Swamy. Nonuniform sampling of band-limited modulated signals. *Signal Proc.*, 4:295–303, 1982.

[PRS84] E. I. Plotkin, L. M. Roytman, and M. N. S. Swamy. Reconstruction of nonuniformly sampled band-limited signals and jitter error reduction. *Signal Proc.*, 7:151–160, 1984.

[PS90] F. Pichler and H. Schwaertzel. *CAST - Computerunterstuetzte Systemtheorie.* Springer, Berlin, 1990.

[PS93] F. Pichler and J. Scharinger. Visualisation of Baker-Prigogine Dynamical Systems. In *Proc. of EUROCAST 93*, Las Palmas, feb 1993. to be published.

[RDF+71] T.C. Rindfleisch, J. Dunne, H. Frieden, W. Stromberg, and W. Ruiz. Digital Processing of the Mariner 6 and 7 Pictures. *J. Geophysical Research*, 76:394, 1971.

[Ree89] A.P. Reeves. VISIX 2.0 User's Manual. Technical report, 1989.

[Res91] Pacific Sierra Research. *MIMDizer User's Guide, Version 7.02.* Placerville, CA.,, 1991.

[RF89] O. Rioul and P. Flandrin. Wavelets and time-scale energy distributions. *Rapport ICPI TS-8910, Lab. Traitm. Signal, ICPI Lyon*, 1989.

[Rin66] T.C. Rindfleisch. Photometric Method for lunar Topography. *Photogrammetric Eng.*, 32:262, 1966.

[RMD90] R. Moreno-Diaz jr. R. Moreno-Diaz. Systems Approach to Fan-In-Out in the Retina. In R. Trappl, editor, *Proc. of the Tenth European Meeting on Cybernetics and Systems Research*, Vienna, 1990. World Scientific.

[Rot73] D. Roth. Digitale Vielfaltfilter,. *Archiv fuer Elektronik und Uebertragungstechnik*, 27:72–78, 1973.

[RSHj86] H.K. Ramapriyan, J.P. Strong, Y. Hung, and C.W. Murray jun. Automated Matching of Pairs of SIR-B Images for Elevation Matching. *IEEE Trans. on Geoscience and Remote Sensing*, GE-24(4):462–472, 1986.

[Rus92] Mary Beth Ruskai, editor. *Wavelets and Their Applications.* Jones and Bartlett, Boston, 1992.

[SA87] K. D. Sauer and J. P. Allebach. Iterative reconstruction of band-limited images from non-uniformly spaced samples. *IEEE Trans. CAS-34/12*, 1987.

[Sau93] K. Sauer. A Local Update Strategy for Iterative Reconstrucion from Projections. *IEEE Trans. Sig. Proc.*, 41:534–548, 1993.

[Sch93] W. Schmidt. Reconstruction of images from structured sampling sets. Master's thesis, Universiy of Vienna, 1993.

[Ser82] J. Serra. *Image Analysis and Mathematical Morphology.* Academic Press, New York, 1982.

[SN72] J. Slansky and P. J. Nahin. A parallel mechanism for describing silhouettes. *IEEE Trans. Comp.*, 11:1233–1239, 1972.

[SN90] D. Shu and J.G. Nash. Multiple-level Heterogeneous Architecture for Image Understanding. In *Proc. 10th International Conference on Pattern Recognition*, pages 629–634, Atlantic City, 1990. IEEE.

[Sou88] M. Soumekh. Band-Limited Interpolation from Unevenly Spaced Sampled Data. *IEEE Trans. Sig. Proc.*, 36:110–122, 1988.

[Spr73] L. V. Sprouse. Automatic verification system for human signature images. *Technical Report*, pages 24–73, 1973.

[SS88a] H. W. Schüssler and P. Steffen. Some Advanced topics in filter design. In J. S. Lim and A. V. Oppenheim, editors, *Advanced Topics in Signal Processing*, chapter 8. Prentice Hall Signal Proc. Series, 1988.

[SS88b] M. I. Sezan and H. Stark et al. Recent analysis in signal processing in medical tomography. In C. H. Chen and Marcel Dekker, editors, *Signal Processing Handbook*, chapter 23, pages 729–795. , 1988.

[Sta87] H. Stark, editor. *Image Recovery: Theory and application.* Acad. Press, 1987.

[Sta93] H. Stark. Polar, Spiral and Generalized Sampling and Interpolation. In R. J. Marks II, editor, *Advanced Topics in Shannon Sampling and Interpolation Theory*, pages 185–218. Springer Verlag, 1993.

[Ste92] G. Steele. Proposals for amending High Performance Fortran. In *High Performance Fortran Forum*, Houston, TX, January 1992.

[Str80] G. Strang. *Linear Algebra and Its Applications.* Academic Press, New York, second edition, 1980.

[Str91] T. Strohmer. Irregular Sampling, Frames and Pseudo-inverse. Master's thesis, Universiy of Vienna, 1991.

[SW81] G. Salinetti and R. Wets. On the convegence of closed-valued measurable multifunctions. *Transactions AMS*, 266 (1):275–289, 1981.

[Tch87] Ph. Tchamitchian, editor. *Wavelets, Time-Frequency Methods and Phase Space.* Springer-Verlag, Berlin Heidelberg, December 1987.

[TKL91] J. Thomas, W. Kober, and F. Leberl. Multiple Image SAR Shape-from-Shading. *Photogrammetric Engineering and Remote Sensing*, 57(1):51–59, 1991.

[Tra85] E.A. Trachtenberg. Construction of Group Transform Subject to Several Performance Criteria. *IEEE Trans on Acc., Speech, and Signal Processing*, 33:1521–1531, Dec 1985.

[Tra86] E.A. Trachtenberg. On Predetection Filtering in the Pilot Blackout Problem using Group Filter. In *Proc. of Second International Workshop on Spectral Techniques*, Ecole Polytechnique de Montreal, 1986.

[Uhr87] L. Uhr. *Parallel Computer Vision.* Academic Press, 1987.

[Wal a] D. Walnut. Continuity properties of the Gabor frame operator. *preprint*, .

[Wal b] D. F. Walnut. Lattice size estimates for Gabor decompositions. *preprint*, .

[Wal92] Gilbert G. Walter. Estimation with Wavelets. Technical Report 4, Department of Mathematical Sciences, University of Wisconsin – Milwaukee, WI 53201, summer 1992.

[WCP92] R. Wilson, A. D. Calway, and E. R. S. Pearson. A Generalized Wavelet Transform for Fourier Analysis: The Multiresolution Fourier Transform and its Application to Image and Audio Signal Analysis. *IEEE Trans. Information Theory, Special Issue on Wavelet Transforms and Multiresolution Signal Analysis*, 38 (2), October 1992.

[WD86] Wegman and DePriest, editors. *Application of the Gibbs distribution to Image Segmentation.* Marcel Dekker, New York, 1986.

[WDM87] J. W. Woods, S. Dravida, and R. Mediavilla. Image Estimation Using Doubly Stochastic Gaussian Random Fields Models. *IEEE Transactions PAMI*, 9 (2):245–253, 1987.

[Weg92] Edward J. Wegman. Wavelets and Nonparametric Function Estimation: A Function Analytic Approach. Technical Report 79, Center of Computational Statistics, George Mason University, Fairfax, VA 22030, June 1992.

[Wei67] P. Weisz. Zusammenhang von Walsh-Fourierreihen mit Polynomen. *Monatshefte fuer Mathematik*, 71(2):165–179, 1967.

[Wic90] M. V. Wickerhauser. Acoustic signal compression with wave packets. *SIAM conference 1990 Chicago*, 1990.

[Wil78] R. G. Wiley. Recovery of bandlimited signals from unequally spaced samples. *IEEE Trans. Com.*, 26(1):135–137, 1978.

[Wil86] R. Wildey. Radarclinometry for the Venus Radar Mapper. *Photogrammetric Engineering and Remote Sensing*, 52:41–50, 1986.

[Win92] D. J. Wingham. The Reconstruction of a Band-Limited Function and its Fourier Transform From a finite Number of Samples at Arbitrary Locations by Singular Value Decomposition. *IEEE Trans. Com.*, 40:559–570, 1992.

[W.K78] W.K.Pratt. *Digital Image Processing.* Wiley, New York, 1978.

[WMW93] A. Weeks, H. Myler, and H. Weenas. Computer-generated noise images for the evaluation of image processing algorithms. *SPIE-Optical Engineering*, 32/5:982–992, 1993.

[WRH91] C. Weems, E. Riseman, and A. Hanson. *Preliminary Results from the DARPA Image Understanding Benchmark Exercise.* Academic Press, 1991.

[You78] D. C. Youla. Generalized image restoration by the method of alternating projections. *IEEE Trans. Circuits Syst. CAS-25*, pages 694–702, 1978.

[You80] R Young. *An Introduction to Nonharmonic Fourier Series.* Academic Press, New York, 1980.

[YS90] S. Yeh and H. Stark. Iterative and one-step reconstruction from nonuniform samples by convex projections. *J. Opt. Soc. Amer. A*, 7(3):491–499, 1990.

[YT67] K. Yao and J. O. Thomas. On some stability and interpolatory properties of nonuniform sampling expansions. *IEEE Trans. Circuit Theory*, 14:404–408, 1967.

[Zak92] A. Zakhor. Two-Dimensional Polynomial Interpolation From Nonuniform Samples. *IEEE Trans. Sig. Proc.*, 40:169–180, 1992.

[ZBC$^+$92] H. Zima, P. Brezany, B. Chapman, P. Mehrotra, and A. Schwald. Vienna Fortran – a language specification. Technical report, Hampton, VA, 1992.

[ZBG88] H. Zima, H. Bast, and M. Gerndt. Superb: A tool for semi-automatic MIMD/SIMD parallelization. *Parallel Computing*, 6(1):1–18, 1988.

[ZC90] H. Zima and B. Chapman. *Supercompilers for Parallel and Vector Computers.* ACM Press Frontier Series, Addison Wesley, New York, 1990.

[ZHD90] W.R. Zettler, J. Huffman, and Linden D. Application of Compactly Supported Wavelets to Image Compression. *SPIE/SPSE Syposium on Electronic Imaging, Science and Technology*, 1990.

[ZR72] C. T. Zahn and R. Z. Roskies. Fourier Descriptions for Plane Closed Curves. *IEEE Transactions on Computers*, 21:269–281, 1972.

[ZVC89] Y. T. Zhou, V. Venkateswar, and R. Chellapa. Edge Detection and Linear Feature Extraction Using a 2D Random Field Model. *IEEE Transactions PAMI*, 11 (11):84–94, 1989.

Chapter 2

Robust and Adaptive Methods for Image Understanding

Project coordinator:

o. Univ.-Prof. D.I. Dr. Walter G. Kropatsch,
Department of Pattern Recognition and Image Processing (PRIP),
Technical University, Vienna.

2.0 Project Overview

2.0.1 General Problem Statement

The goal of Image Understanding (IU) is to find semantic interpretations of images, in particular to localize and name objects contained in a scene and to assess their mutual relationships in order to interact with the environment. The central problem is to recognize known objects reliably, independent of variations in position, orientation, and size, even when those objects are partially occluded. Although some information is lost by the image formation process, biological vision systems demonstrate that this task can be accomplished reliably and efficiently even under difficult viewing conditions.

In the past 25 years, many specialized techniques have been developed for various aspects of IU without reaching a comparable or even sufficient degree of robustness. The overall complexity of the problem makes it difficult to devise suitable architectures for vision systems. On the other hand, specialized systems components (i.e., algorithms, control structures, knowledge representations) cannot be developed without a global strategy in mind. Clearly, the complexity of the problem domain calls for a systems-oriented approach that can only be accomplished through intensive cooperation between the individual research groups involved.

Almost any current IU approach can be described as a sequential 3-stage process (comp. Fig. 2.1).

In the first stage, the raw image data, given as an array of intensity values (DIGITAL IMAGE), are processed in a context-free, bottom-up fashion to produce a set of data primitives, often called "features". In the second phase, features are assembled to more complex structures (e.g. REGIONS) that are useful components for scene descriptions and object models. In the third phase, these complex features are matched against existing object models in a data base (e.g. OBJECTS in the WORLD vs. OBJECTS in the SCENE). Research during the decade of the 1980's demonstrated that most vision problems are either ill-posed or inherently complex (NP-complete). From these considerations the paradigms of active vision, qualitative vision, animate vision, purposive vision emerged.

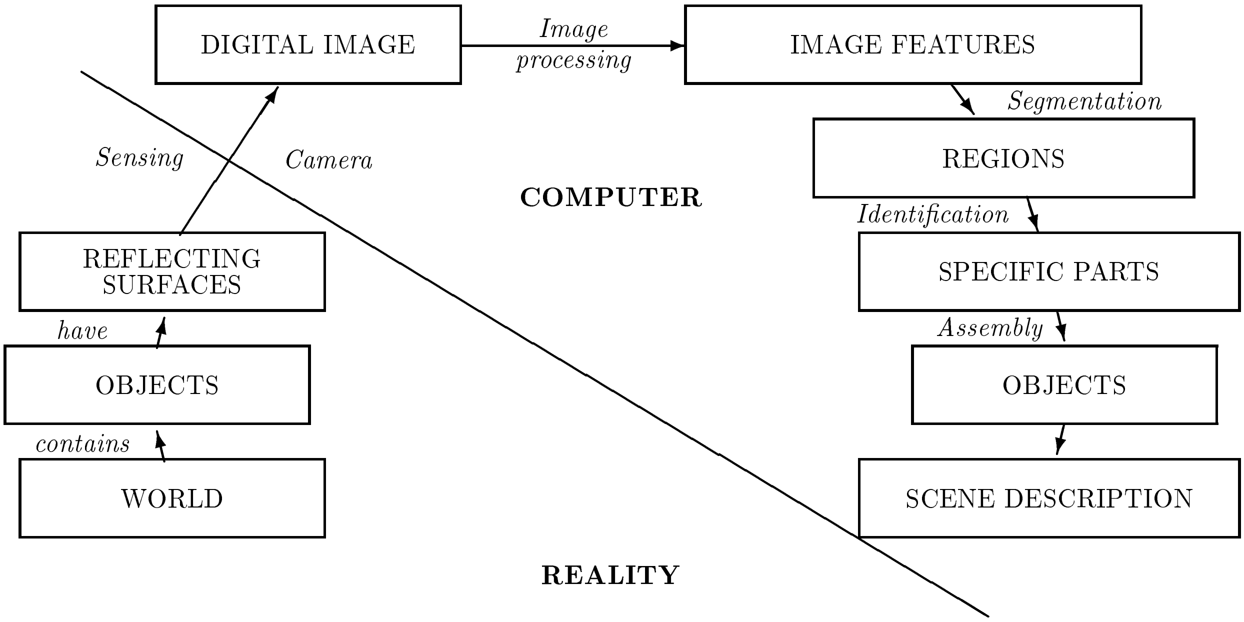


Figure 2.1: Processing steps in digital image analysis

Most current model-based IU approaches are restricted to well-defined man-made objects that are usually located in controlled, unstructured industrial environments. In contrast, this project shall put stronger emphasis upon *variable-form objects* as they populate natural environments.

We have identified three dominating problem areas that appear to be essential for successful realization of Image Understanding. *First*, there is a clear lack of robust methods for low- and intermediate-level image analysis. *Second*, current object representations and matching strategies are mostly confined to a single type of feature. *Third*, the construction of object models is a task too complex to be done manually and therefore must be supported with suitable tools (i.e adaptive methods).

2.0.2 Specific Research Goals

In summary, this research project shall focus upon the following issues:

- Robust Hierarchical Methods for Image Analysis
- "Polymorphic" Structural Feature Detection and Matching
- Machine Learning for Image Understanding
- Image Understanding Applications

Task	Section	Title	Institution
1	2.1	*Robust Hierarchical Methods*	TU Wien
2	2.2	*Polymorphic Structural Feature Detection and Matching*	JKU Linz
3	2.3	*Machine Learning for Image Understanding*	JKU Linz

The cooperative nature of this project will require an intense exchange of ideas, concepts, and algorithms. It is therefore an important part of this project to provide the fundamental organizational and technical prerequisites to allow for this multi-directional flow of information, including conventions for the exchange of data and software. This goal will be supported by two facts: First, both institutes use KBVision as their main platform for Image Analysis. Second, the cooperation project (c.f. Project 1) will provide the necessary means for efficient communication.

2.1 Robust Hierarchical Methods

Task leader:

o. Univ. Prof. D.I. Dr. Walter G. Kropatsch,
Department of Pattern Recognition and Image Processing (PRIP),
Technical University, Vienna.

2.1.1 Problem Description

Objects from the real world are mapped into an image and can be recognized there by their properties and the spatial relations of their parts. In many cases the complex mapping from three dimensions to the two dimensions of an image disturbs or even destroys this information. However, images still contain enough information that a human observer can easily recognize its content. Often, the results from image analysis methods depends essentially on the circumstances under which the image was taken (e.g. sun angle, viewing direction, etc.). Besides the incompleteness and the ambiguity of the solutions (image recognition is an ill-defined problem) digital imagery has the additional problems caused by the huge amount of data that have to be processed in relatively short time.

In general, it cannot be assumed that features extracted from the image are accurate. There are many sources that may corrupt the results of image analysis:

- Noise, distortion, partial occlusion, shadows, and specular reflections are among the effects that cause features to be distorted during image formation.
- Several errors can be introduced during processing of the data (e.g. by numerical errors, by quantization, by sampling, etc.)
- The image representation often depends strongly on the arbitrarily chosen or randomly given mapping parameters. A little shift or rotation may result in dramatic changes in the representation.
- The model for interpretation makes wrong assumptions which lead to wrong results (e.g. many restoration methods assume normally distributed noise, which is not always the right model).
- Sometimes the representation is incapable of capturing the necessary information. It is then lost when the data are transformed into such a representation.

Such influences during the formation and the processing of the image data cause differences in the results as expected under 'ideal' conditions. Features may be distorted in shape or size, may be incomplete or missing and also may cause false features to be detected. While perfect segmentation has been assumed in many implementations, any recognition system working under real condition must handle such deficiencies. A method is called *robust*, if the result of applying the method is not sensitive to little perturbations in the input.

Most existing approaches to model-based object recognition rely upon *contour-based features* as the basic object descriptors. In particular, they use straight line segments and geometric relationships between those segments, such as the work done by [Low87, BF84, GB87]. Straight line primitives result in relatively simple object models and straightforward matching algorithms. For example in [KT89] an algorithm was presented which can determine the straightness of digital lines in $O(n)$ steps. Other researchers have employed additional contour features, such as circular features and smooth curves, as in the work done by [AF86, RB84, Gri90]. One reason for the poor reliability of contour-based features is the differential nature of the underlying edge enhancement process.

Feature groupings are assemblies of primitive features to more complex salient structures. Examples are groupings of line segments to form smooth curves, corners, parallels, etc. Relevant groupings are made of image features that are *significantly related* to each others. These relations are significant if they are likely to be caused by an object in the scene rather than by accidental external conditions like for example the viewpoint of the observer. Only geometric relations that are invariant against changes of the observer's viewpoint can have any significance for the presence of an object in a (3D) scene. This limits drastically the number of relations to consider. Only five perceptually significant classes (non-accidental properties, cf. [Bie87]): straightness, connectivity, parallelism, curvilinearity, and lines meeting in a common point, have to be taken into account. Grouping criteria for line segments, such as proximity, continuation, parallelism, etc., have been applied successfully [Low87]. However, efficient implementations of line groupings (i.e., algorithms, transformations, and data structures) are still open problems that need to be addressed. Similarly, grouping processes for other types of features and groupings of different features must be investigated (see also task II.2).

2.1.2 Hierarchical Data and Processing Structures

Any intermediate of final results extracted from the raw image data are required to be robust in quality in the sense that they are reasonably insensitive against diverse variations in the data and the processes. Also, since features are usually computed in a data-driven, bottom-up fashion, they must also be computed efficiently. To solve both the qualitative as well as the quantitative problems, hierarchical data and processor structures (multiresolution representations, image pyramids) have shown to be useful [Ros84, Bur88a].

An important parameter of a digital image is its *resolution*. The spatial resolution is described in terms of the smallest dimension of the object that can just be discriminated. In an image with high resolution many details can be observed, at low resolution only large objects are recognized. A *resolution cell* is the smallest most elementary areal constituent in a digital image. The position of a resolution cell within the 2D plane is determined by the coordinates (x, y) of its center. The number of cells to cover a given image depends on the *size* of the cells. Table 2.1.2 compares the qualitative consequences of different cell sizes.

Table 2.1: Image Qualities at Different Resolutions

	small cells	large cells
resolution	high	low
data amount	huge	smaller
computing times	(very) long	(relatively) short
details	rich and many	very few if at all
overview	bad	good
precision	high	low

An image pyramid combines several of the advantages of high and low resolutions. Tanimoto [Tan86] defines a *pyramid* as *a collection of images of a single scene at different resolutions.* The images of this collection can be ordered according to their cell sizes and numbered as *levels* of the resulting ordered set of images.

Two terms describe the structure of a (regular) pyramid: the *reduction factor* and the *reduction window.* The reduction factor determines the rate by which the number of cells decrease from level to level. The reduction window associates to every (higher level) cell a set of cells in the level directly below. In general the reduction window covers the area of its associated cell. In square or rectangular images reduction windows are mostly rectangularly shaped and can be described by ($number\_of\_columns \times number\_of\_rows$).

In the *classical pyramid* every 2×2 block of cells is merged recursively into one cell of the lower resolution. We formally describe this structure by $2 \times 2/4$ which specifies the 2×2 reduction window and the reduction factor of 4. This type of pyramid has been extensively studied [TK80, Ros84]. Starting with a 512×512 image in the base, the constant reduction factor (4) determines the sizes of the levels above: 256×256, 128×128, 64×64, ... In general, a pyramid with $m \times m$ cells in the base creates $\log(m)$ levels between the base and the apex.

Classical pyramids use constant reduction factors of four between the levels of the pyramid. That means that every lower resolution has only 1/4 of the pixels. But also other multiresolution structures with lower reduction factors are possible [Kro85b, Kro86c, Kro88c, Kro88b]. They have the advantage that the resolution is reduced gradually. Objects of sizes that fall between two successive levels of the classical pyramid (e.g. too large for the higher resolution but too small for the lower resolution) may find the appropriate level in a $n \times n/2$ pyramid.

The robustness of pyramidal algorithms to perturbations of the *structure* of the pyramid is studied by Meer et al. [MJBR88]. In these experiments the reduction function was a weighted average of the son's values. The perturbations are of three different natures:

- Gaussian noise added to the ('generating kernels') weights;
- random local offsets of child-parent links;
- stochastic pyramid structures.

Meer concludes that:

- Multiresolution algorithms appear to be very robust and only weakly dependent on the structure on which they are implemented.
- The parent should control the computation.
- At higher levels of image pyramids, the amount of noise is significantly decreased. Objects may become easier to detect. They allow fast data gathering across the input image.
- The resolution reduction in pyramids can be regarded as the discrete case of scale-space smoothing.

Perturbations of the originally regular structure results in an irregular structure which gave rise to the concept of the *stochastic pyramid* [Mee89] and the *adaptive pyramid* [JM92]. In a regular pyramid, the reduced resolution representations can be distorted when the input is shifted [Tan76]. Bister et al. [BCR90] show that most classical pyramids, which are regular and consequently rigid structures, have problems in segmenting an image, e.g., when the input image is shifted, rotated, or differently scaled. This shift-dependence problem will appear whatever the neighborhoods which are used to built the pyramid (that is non-overlapping or overlapping neighborhoods).

In the stochastic pyramid the structure is generated by a random process with some additional local constraints. Therefore the resulting pyramid is nearly independent of the location and orientation of the objects within the field of view. In Meer's stochastic pyramid [Mee89], the decimation is based on a randomized parallel algorithm. The stochastic decision process starts by assigning every cell a uniformly distributed random number. Then the *local maxima* are selected as surviving cells and gaps closed as before.

The concept of *adaptive pyramid* [JM92, MMR90] differs by the selection criterion. There the decimation process first tries to select significant cells depending on the cell's content and uses a random selection only where the data do not allow a decision.

Another hierarchical decimation method is presented in [DP88]. There, the levels form triangular networks, and the decimation is based on data-dependent criteria like in the adaptive pyramid. The neighborhood relations of the decimated levels are determined by a new version of the constrained Delaunay triangulation.

The efficiency of irregular pyramids has been shown very recently: the labeling of connected components of an image needs only $\mathcal{O}(\log n)$ parallel steps [MMR89]. This efficiency is made possible by decimating the components separately.

Dual representations enhance the description capabilities by combining numerical data and symbolic data in the concept of dual pyramids [Kro86b, Kro88a]. The usage of symbolic representations within a pyramidal framework is relatively new. The symbols used are still very simple image primitives like curve segments or peaks and ridges [CP84] on the gray level surface. But since image interpretation has the objective to describe the image in terms of a high level vocabulary, symbolic pyramid representations are a step in the direction of high level conceptual hierarchies as proposed by R. Bajcsy [BR80]. The combination of such rich representations with the robustness of irregular pyramids let us expect to come closer to our goal of having methods that produce reliable image analysis results in reasonable time.

A different approach for robust structures comes from the area of *artificial neural networks*. Though these structures have many desirable properties like robustness, massive parallelism and the adaptivity [RM86] for complex problems, such as vision, the current approach of using fully connected (three-layer) neural networks has severe deficiencies. As Le Cun [LMB+90] has stated "Expecting good performance without any a priori knowledge, relying exclusively on learning is wishful thinking". One way to incorporate a priori knowledge is to specify a proper topology of the network. Neural networks share many similarities with image pyramids. Leonard Uhr [Uhr92] points out in a theoretical analysis that hierarchical structured networks (i.e. networks which employ only local connectivity) offer efficient solutions for many problems. We were able to show [BK93b] that we can build equivalent neural networks for all kinds of pyramids (that is numeric and symbolic pyramids, regular and irregular pyramids). With this framework we can enhance pyramidal algorithms by neural network functions, and neural networks by properties of pyramids. The advantage of using neural networks in a pyramidal framework is that we have a fine grained description, which gives as more degrees of freedom. One of the main motivations for identifying the similarities between neural networks and image pyramids is to employ learning algorithms on pyramids. It should be noted that Task II.3 is also concerned with learning. However Task II.3 will study "symbolic learning" while we propose to use connectionist methods. These two approaches are complementary however a comparison of the results will be done.

2.1.3 Proposed Research Topics

With respect to feature extraction, this project shall investigate in following problem areas:

1. Robust techniques for extracting and grouping contour-based features (e.g., pyramid techniques).
2. Alternative features that are inherently more reliable.
3. Efficient techniques for feature grouping (perceptual organization).

In the above areas, the project intends to produce new insights in following actually open problems:

Properties of irregular pyramids

The local constraints that are used for the generation of irregular pyramids produce certain unwanted side effects:

- Distances between cells at higher levels need not be greater than in the levels below (as proven in [Kro90]).
- Any cell in the pyramid should have a limited number of neighbors at the same level. This is an important property needed for example for efficient addressing and indexing. It is not clear so far how such a property can be imposed on irregular pyramids.

- The requirements of the processes is that they are local and independent of one another, which gives the possibility of parallelizing the algorithms. For the concept of dual pyramids we require the planarity of the graphs. Cases where we do not have the planarity (e.g. 8-connectedness) require a different concept.
- Extend the results of 2-D decimation to higher dimensions, e.g. 3-D and 4-D irregular pyramids.

Dynamic adaptation of the pyramidal structure

Decimation pyramids adapt their structure to the data of a static image. To apply the same procedure to dynamic images (with e.g. moving objects) a new structure would have to be created an top of every image of a sequence although there will be only small changes between temporally adjacent images. We therefore search for strategies to modify a given irregular pyramid to accommodate small changes in the base image while maintaining decimation properties.

Irregular curve pyramids

In a square grid a curve crosses the boundary of a cell at two of the four sides of the square. If a cell in an irregular pyramid has a bounded number of sides, the same concept can be applied to define a symbolic curve representation.

Add new primitives to the curve pyramid

When curve segments are generalized certain properties get lost. E.g. adjacent corners could get merged, antiparallel contour segments become one curve segment although they represent a grey level cross profile (ridge) different from contours (edge). In some cases contour segments should have an orientation if grey levels on the right side are brighter than on the left side. Enhance the descriptive power of pyramidal cells. In particular, extend the existing symbolic curve representation scheme to include also line crossings and junctions.

Symbolic Pyramid Matching

Identification of parts of objects as parts of an image needs matching. Burt's pattern tree [Bur88a] is an example how matching can be done efficiently within a pyramid structure. Using the robustness measurements derived in [FK89] matching strategies should be developed that allow efficient and robust shape matching.

Symbolic refinement of symbolic pyramids

Let $C(i)$ represent level i of a symbolic pyramid. Find detail (or difference) representations $D(i)$ that allow reconstruction of $C(i)$ given $C(i+1)$ and $D(i)$ in analogy to Burt's Laplacian pyramid [BA83] or Mallat's wavelet representation [Mal89].

Gray values from contour lines

In the dual concept curve segments are detected in the grey level image by local detectors. Using simple multilevel thresholds the image can be reconstructed from the curves with a few extra seed pixels (e.g. at local extrema).

Control strategies in pyramids

Image data are input to a pyramid at the bottom level and are propagated upwards. In higher levels information is collected that often allows hypotheses about the expected image content. Such hypotheses can be verified by processes that step down to the lower levels to check certain properties. This implements some sort of focusing. 'Higher level cells tell their children what to look for and where to search it'. Control strategies must be found that combine both bottom-up data flow and top-down control flow. Speed-ups could be achieved by allowing a certain degree of asynchronous processing.

Irregular curve pyramid as a model for perceptual grouping

One method to accumulate evidence during perceptual grouping is referred to as *recursive application of structuring* by Lowe ([Low85], page 42f). This process fits very well into a (irregular) pyramidal framework. In particular the research will focus on the following topics (note also the relation to Task II.2):

- Straightness: Transfer of the RULI-chain code algorithm to curve pyramids;
- Gap bridging;
- Crossings and junctions as new primitives.

Hierarchical neural networks

The main objective of this research is to exchange knowledge between neural networks and image pyramids. In particular we would like to study:

- The concept of symbolic pyramids in a neural framework. Especially the development of new primitives by employing learning algorithms.
- Apply learning algorithms to pyramids. In particular we are interested in unsupervised learning schemes which develop optimized representations at the reduced levels. Especially the principle of "Maximum information preservation" (cf. [Lin88]) provides an reasonable approach in this direction. Some learning algorithms which will serve as a starting point are Oja's rule [Oja82], Autassociative Back-propagation [RHW86, BH89] and the algorithms developed by Linsker [Lin88, Lin86].
- Attention mechanisms for hierarchical neural networks. There exist a lot of specific mechanisms for focusing attention in neural networks (e.g. citeTsotsos91a, Milanese90a, Olshausen92a). However most of these models are concerned with mechanisms to implement the focus of attention use only simple techniques to determine the part where to focus attention. In particular we would like to answer the question, which information allows us to focus our attention at a certain region of an image. This part is also closely related to control strategies in pyramids, and to the models of perceptual grouping of this task and the of task II.2.

2.1.4 Previous Work

There exists extensive knowledge and experimental software for simulating parallel processes on pyramidal structures. Experiences with different types of reduction functions to build a gray level pyramid are described in [Her87, KP87, KH88].

Several contributions concern the symbolic representation of curves in pyramids [Kro85a, Kro86d, Kro86a, Kro87, FK92a, FK92b], necessary and sufficient conditions for the recognition of straight lines

[Toc87, KT89], and shape patterns such as extrema of curvature or corners [FK89, Fer89, FK92b]. A combination of gray level pyramid and curve pyramid has been used to preserve contours during reduction in the gray level pyramid [Kro87, Kro88a, Paa87, PK88].

There has been theoretical work on irregular pyramids [KM91, KM92, Kro92], and irregular curve pyramids [KW92a, KW92b, KRWW, KW, KW92c]. It could be proven that the dual graphs of irregular pyramids have bounded degrees. Software implementations are currently under work.

The theoretical backgrounds have been developed in close contact with one of the centers for multiresolution methods, the Center for Automation Research (CfAR) at the University of Maryland. The continued contact with the CfAR has lead to many new aspects in the theoretical foundations of irregular hierarchies [Kro90]. Another cooperation has been established with the image processing group of University Joseph Fourier in Grenoble (IMAG-TIMC). There is also strong interest in irregular pyramids in this group. IMAG and PRIP with other European partners (Ruhr University Bochum Inst. for Nero-informatik, and University Leiden) participate in an ERASMUS program, and a HCM program is planed for the near future. We hope to continue this fruitful cooperations.

A somewhat complementary approach are the works with T. Yamaguchi where the building principles of irregular neural networks with hierarchical structure such as Fukushima's Neocognitron have been implanted into the regular classical pyramid [YK89a, YK89b]. Several papers discuss the relationship between image pyramids and neural networks. In [BK93c, BK93b, Bis91] we have presented a general framework. In [BK93a] we discussed how to use Hopfield neural networks for constructing irregular pyramids. In [BK93c] we have shown how to describe the curve pyramid by a neural network. In [Bis93, Bis92] we outline some preliminary results of learning in a pyramidal framework. Some work on focusing attention in hierarchical neural networks has appeared in [BP92b, BP92a, BH93].

2.2 Polymorphic Structural Feature Detection and Matching

Task leader:

Dr. Wilhelm Burger,
Department of Systems Science,
Johannes Kepler University, Linz.

2.2.1 General Description

Current approaches to structural object recognition are typically based on a *single* type of structural primitives. This limits the applicability of these techniques to a very restricted type of scenes and provides poor performance in general. The combined use of *multiple* types of structural primitives promises increased recognition performance and reliability and allows to cover a broader range of scenes, including complex natural scenes. In this part of the project, we want to investigate the full potential of such a "polymorphic" structural approach and develop methods for overcoming the technical difficulties involved.

2.2.2 Detailed Problem Statement

Structural object recognition is based on the idea that the appearance of objects can be described by visual primitives (called "features" or "tokens") and the spatial relations between them. In the structural approach, the spatial relations between features are expressed explicitly to improve the specificity of the description. This is the reason that *structural* descriptions are preferred over *evidence-based* descriptions.

For the purpose of visual object recognition, we must distinguish between two different forms of structural descriptions:

1. the structure of the *true* object (i.e., in the three- dimensional reality), called the *object model*;
2. the structural description of the *appearance* of the object in the image, called the object's *image description*.

2-D and 3-D Structural Features

Similarly, two different types of image features can be extracted: those that are directly related to the 3-D shape of the object (or part of it) being viewed, and those features that result from the 3-D to 2-D projection. In addition to the problem of detecting both kind of features in intensity images, the latter type also causes ambiguities since part of the 3-D shape information is lost during the projection. The essence of the recognition problem is to relate the structures found in the image with the underlying object models. Other important issues involved in structural recognition are *(a)* the adequacy of the representation for the kind of objects encountered, *(b)* the selection and extraction of visual primitives, *(c)* the description of the spatial relations between primitives, *(d)* the problem of matching image structures to models, and *(e)* the inference of structural object descriptions from examples.

Feature Selection

The selection of suitable shape primitives is of central importance. For efficient recognition, tokens should be most *expressive* in the sense that a combination of only a few of them (or even a single token) can facilitate object identification [Bie87]. On the other hand, the available shape primitives should be *general* enough to model a large range of object categories. Tokens should also be detectable from images reliably in a bottom-up fashion, and be *non-incidental*, in the sense that they are unlikely to occur from random configurations in space.

Today, most structural approaches are based on a *single* type of shape primitive, usually straight line segments. Although these features are relatively easy to extract from the image and well suited for many man-made objects, they are not adequate and reliable for most complex natural scenes. In fact, the poor performance of most current image understanding systems in natural environments can be attributed mostly to their reliance upon a single feature type. The focus of this project part (2.2) is to develop methods for using *multiple* feature types for structural object recognition in a cooperative and complimentary fashion. We refer to this as a "polymorphic structural approach".

Matching Structural Descriptions

Assuming that both the contents of the image and the contents of the object model(s) are given in the form of structural descriptions, a matching procedure must solve the following three problems simultaneously:

1. *Selection Problem:* determine which image tokens belong to the same object.
2. *Identification Problem:* determine the identity of the structure, i.e., invoke the right object model.
3. *Association Problem:* assign the right model *component* to each image primitive.

Except for trivial configurations, none of these three problems can be solved without solving the other two. The *selection problem* cannot be solved without knowing the identity of the object and collecting those image tokens that are required by the model. Selecting the right model is not possible unless at least some of the image tokens have been assigned to model primitives which, in turn, requires solving the identification problem. In summary, this turns structure matching into a complex search process in a large space of alternatives.

Formalized as a graph-theoretic problem, matching structural descriptions is equivalent to finding a *morphism* between the primitives in the image and the model, such that the structure is preserved. This problem is known to be $\mathcal{NP}$-hard, i.e., generally the amount of computation grows exponentially with the number of structural elements.

Reducing Search Complexity

The search space of the matching process can be reduced mainly by three means: *(a)* using more expressive tokens [Ett88], *(b)* assembling image tokens that are likely to belong to the same object (grouping), *(c)* making a good guess for the object's identity (indexing), and *(d)* by using heuristic constraints to rule out unlikely hypotheses at an early stage [Gri90, Gri91]. In addition, *structural similarity measures* offer a way to match substructures without directly associating their individual elements.

Grouping and Indexing: When all sensory data under consideration are known to belong to a single (known) object, the expected amount of search is quadratic in the parameters of the problem and linear in the number of data-model pairings. This assumes that *(a)* the token selection problem and *(b)* the indexing problem are correctly solved. Therefore, the pre-selection of scene tokens can effectively reduce the amount of search but, unless pre-selection is perfect (all selected tokens indeed belong to a single object), search is still exponential in the worst case. If indexing fails, i.e., the wrong model is used, search is also exponential, under otherwise perfect conditions.

Current perceptual grouping methods [GB87, Low87, Sau90] are mostly based on

(i) a *single* type of primitives and

(ii) *predefined* grouping rules that are not adapted to the application domain.

Although the resulting simple representations and grouping criteria can be evaluated efficiently and the object descriptions are independent of the problem domain, the traditional grouping approach has several disadvantages:

(a) The perceptual "saliency" of groupings between different types of primitive features is not used.

(b) Groupings based on a single feature type are inherently brittle.

(c) Fixed, domain-independent grouping rules are not optimal for unknown and dynamically changing environments.

Early Termination of Search: Early termination of the search process [Gri91] is based upon local unary and binary constraints applied to pairings of scene and model tokens. These constraints are used to rule out interpretations before the full depth of the interpretation tree is reached. In effect, this combines heuristic search with constraint satisfaction or consistent labeling ("constrained tree search"). In order to achieve polynomial time recognition algorithms, the ratio of clutter tokens and real object tokens must be kept below a specified bound (which is given). This again suggests the use of a pre-selection mechanism for grouping tokens that are likely to belong to a single object.

Use of Structural Similarity: Graph morphisms are a very strong notion of *equivalence*. In the strict sense, two graphs are considered equivalent (i.e., isomorphic) when a 1:1 mapping exists between their nodes and the labels attached to the corresponding arcs are identical. This form of equivalence is not realistic in structural recognition for various reasons. *First*, it requires a finite set of possible node labels. Node labels, however, represent properties and spatial relationships of tokens, including distances, angles, etc., from a potentially infinite range of values. *Secondly*, to be of practical use, matching must be tolerant against spatial variations to a certain degree. Consequently, instead of searching for structural *equivalence*, the matching process should look for structural *similarity* that can be quantified by a suitable distance measure. A classic example for such a distance measure is the model proposed by Fischler and Elschlager [FE73], where the spatial relations between tokens are abstracted as elastic springs. The matching distance used is a function of the spring loading, among other parameters.

The use of a matching *distance* instead of a matching *predicate* turns the original graph matching problem into an optimization problem, with the goal to find the "best match" among a large number of possible ones. Practical graph matching techniques are often characterized by the use of heuristics to avoid exhaustive search. Examples are backtrack search algorithms [BB82], association graph techniques [ABB+75], relaxation labeling [BF84], and various hashing schemes [Bre89, LW88]. Considerable work has been done towards structural recognition in industrial environments, particularly under occlusion, i.e., when objects are only partially visible [AF86, BM87, BC82, KJ86].

Other Strategies for Reducing Search: Ordering of structural tokens can significantly reduce the required search effort compared to regular combinatorial search. As argued in [SMS87], where tokens are ordered according to their distances from the object's centroid, practical matching algorithms must execute in polynomial and not exponential time. Multi-stage strategies are another alternative to improve matching performance. The computational advantages of a 2-stage matching process are examined in [SR89]: First, a subgraph of the model is selected and all possible matches with the scene are evaluated. The smaller the chosen subgraph, the fewer matches have to be tried, but the number of acceptable matches and thus the work necessary at the second level will be higher. An essential statement is that the threshold for accepting a subgraph match depends upon the size of the entire model graph and not upon the subgraph only. This guarantees that expanding a rejected subgraph match will never result in an acceptable total match. It is show in [SR89] that the total matching costs are minimal for some fixed subgraph size. However, this is not necessarily the case with geometrically constrained graphs.

2.2.3 Proposed Research Topics

The main goal of this project part is to develop the *polymorphic structural approach* for increasing the robustness and efficiency of structural object recognition in complex natural scenes and with partially occluded objects. The main focus of this work is on the following subproblems:

A. Detection of Polymorphic Features;

B. Structural Object Descriptions with Polymorphic Features;

C. Grouping, Indexing, and Matching with Polymorphic Structural Features.

Adaptive and learning techniques (see also Task II.3) will be considered for some of these problems, particularly for those related to object modeling.

Detection of Polymorphic Features

We plan to investigate the following structural primitive classes with respect to their detectability, descriptiveness, robustness, invariance properties, and ambiguity:

1. *Contour-based primitives*, such straight line segments [BWR89, BHR86], parameterized curve segments [FH88, Leo93], corners [BR92, CVK91, GD91, aSN90], and multi-scale curve descriptions [LE92, Sau90]. This will include results obtained the work described in Task II.1.
2. *Region-based primitives*, such as classical region segmentations [KR84], local homogeneous blobs [Bur88b, Heu92, LE92, RS88], and oriented wavelet patterns citeBovik92,Mallat90. In particular, we will investigate the use of Gabor wavelet patterns for performing recognition under moderate variations of scale and object aspect.
3. *Qualitative 3-D surface primitives*, using invariance characteristics derived from differential geometry, based on the promising results from our current work [Sar92, Sar93].

Extraction methods are already available for some of these feature classes (contained in various software packages, such as KBVision, KHOROS, OBVIOUS, etc.), while suitable methods still need to be investigated or newly developed for others. In particular, additional work is needed for the blob detectors, the oriented wavelet patterns, and the qualitative 3-D surface primitives. Also, some effort must be spent to adapt existing feature detection methods and make them available in a common software environment.

Grouping, Indexing, and Matching with Polymorphic Structural Features

Polymorphic Grouping: Grouping is a powerful strategy for combining structural primitives into more expressive aggregates. It can dramatically improve object model indexing and reduce the search complexity. While techniques for grouping primitives of a single type are well developed (e.g., for point sets, straight line segments, and segmented regions), no general methods are available for grouping *different* structural primitives. One of the difficulties involved is that different primitives mostly share only a small set of common properties. E.g., some primitives have an *orientation* property, while others don't. Consequently, a distinct evaluation function is required for each allowed combination of feature types.

However, "polymorphic groupings" offer a much greater expressiveness than "homomorphic" groupings of the same number of primitives. Thus, to achieve a certain degree of *indexing power*, the number of primitives needed in polymorphic groupings can be kept relatively small. We can therefore be quite selective with respect to the set of permissible feature class combinations, which also reduces the number of distinct evaluation functions. Similarly, the overall complexity of indexing and matching is reduced by using smaller but more expressive assemblies of structural tokens.

In current approaches, the *grouping rules* and grouping evaluation functions are *static* and specified at the syntactic level for efficient bottom-up detection of salient groupings [DW89, GB87, Sau90]. However, the effectiveness of such general grouping criteria will strongly depend upon the relevant objects in the application domain. E.g., searching for parallel lines is effective for certain man-made objects but not in most natural scenes. Thus we want to dynamically adapt the grouping rules to the corresponding application domain, taking into account the objects encountered in this domain, their relative importance in the given task, and the (polymorphic) features used for modeling these objects.

The main subtasks are to develop

1. the database of perceptual relationships,
2. evaluation functions to measure the "saliency" of high-order polymorphic groups, and
3. efficient grouping algorithms that can handle polymorphic structures.

Indexing and Matching Indexing and matching are complementary processes in structural recognition and often incorporated in a hypothesize-and-test strategy:

Indexing is used to narrow down the set of possible object or subpart categories, by evaluating certain properties of a given image feature set. While geometrical invariance is considered important for indexing based on point sets and straight line segments [MZ92], combinations of polymorphic features carry high indexing power themselves. The connection between grouping (see above) and indexing are strong, because grouping supplies the token sets for indexing.

Matching is performed by explicitly hypothesizing correspondences between image tokens and object model entries. Indexing can effectively reduce the number of matches to be evaluated. Matching hypotheses requires a *metric* that allows to measure the distance between image features and entries in the model base. In contrast to single feature type approaches, new and specific distance measures must be developed for the polymorphic approach. The combinatorial advantages resulting from the use of polymorphic features will be analyzed and verified under realistic conditions.

Structural Object Descriptions with Polymorphic Features

Object models in this approach are viewer-centered descriptions based on the spatial arrangement of structural features. Each object model consists of a set of distinct aspects of the corresponding object. The number of different aspects needed to sufficiently describe an object depends mainly on the object's complexity, viewpoint dependency, and the number of "common views" related to that object. Individual aspects are related to each other by an adjacency relation defined on the "view sphere", such that intermediate aspects can be inferred if necessary. Each aspect is, in general, a multi-scale representation that refers to structural details at a range of different resolutions. In this context, the relationships between polymorphic features at different scale levels presents a challenging area of investigation.

To successfully cope with natural objects, the representation must accommodate a certain degree of object variability. The traditional rigid structural descriptions are not adequate for this task. We propose to use structural constraints of varying rigidness for handling object variability:

- *strong* structural constraints for describing local structural details, and
- *weak* structural constraints for describing larger structural aggregates.

In addition to providing tolerance against object variability, the use of weak structural constraints at the global object level also reduces the search complexity during recognition. Also, there appears to be evidence [Cer86] that similar forms of representations may exist in biological vision systems.

2.2.4 Application to Automated Document Processing

Theoretical concepts need to be demonstrated in real-world applications in order to be valuable, which is of particular importance in an engineering domain such as image processing and computer vision. While results of the previously described concepts will be shown on several other types of real images, we want to pursue automated document processing as a major application at the systems and implementation level. By working on basic methods and a corresponding application at the same time, we expect synergy effects and increased overall motivation. While short-time scheduling often leads to compromise solutions in industrial development environments, this project would provide the opportunity to produce engineering solutions based on solid scientific grounds.

Problem Description

Public utility companies (such as electrical power, telecommunications, water, and gas companies) maintain wide-area network facilities that require extensive documentation and planning. While a large part of this engineering effort is still done in traditional paper form, there is a great demand for computerized support. The use of computer systems for utility network documentation will provide easier maintenance, increased accessibility, higher accuracy, and reduced cost. It will further allow new types of network analysis that were impossible to perform without computer and CAD support. For example, the prediction of cross-coupling effects between high-voltage power lines and nearby wire-based communication equipment is an important engineering factor.

A central problem, however, is the acquisition of *existing* engineering drawings, plans, and maps into computerized formats. This is a very time-consuming, demanding, and thus expensive process. In many cases, sufficient manpower for performing the conversion in a reasonable time frame is not available.

The desired computerized support should include reliable means for

- precise and reliable extraction of global structural data (e.g., the network components, even when dashed lines are used),
- interpretation of network properties (e.g., thick lines have a different meaning than thin or dashed lines),
- extraction and interpretation of symbols
- extraction of written text,
- efficient means for human interaction and correction.

While several vendors of computerized geographical information systems (GIS) provide utilities for digitizing and converting existing plans, the capabilities and robustness of these systems are reportedly not satisfactory and the additional manual effort remains unacceptable. What we are faced with is a classical image interpretation problem that cannot be solved with graphics and image processing tools alone. The technology and scientific knowledge developed in other parts of this joint project provides a solid basis for solving this problem. In particular, the techniques related to structural scene analysis and object recognition developed in project parts II.2 and II.3 will find direct applications in this problem domain. In return, the work on *real* engineering problems like this will allow us to efficiently verify our theoretical results.

Proposed Research Topics

The focus of this project part is to implement a modular software system for performing intelligent technical document acquisition and processing. A commercial document processing environment will serve as the base system. The software developed in this project will consist of a set of dedicated modules that introduce the required functionality to the base system or enhance existing facilities.

The focus will be on the following problem areas:

1. Decomposition of technical schematics (blueprints) into semantic components;
2. "Intelligent vectorization" of line art (e.g., managing thick lines, gaps, dotted lines, merging of elements, color).
3. Extraction and recognition (interpretation) of text components under varying orientation and style.
4. Consideration of explicit sensor models for the document scanning process.
5. Integration with graphical databases.
6. Integration with topographical maps.
7. Combination of multi-source network schematics (e.g., power, telephone, water, natural gas, railroad, highways).
8. Visualization of specific simulation results (e.g., crosstalk effects, electrostatic / magnetic fields, noise).

Implementation Issues

System development will be performed on general-purpose UNIX workstations with intensive graphical support. Original data will be provided by public utility companies. With respect to the software implementation, we will emphasize the following issues:

1. Use of an extendible commercial base software.
2. Clean interface specifications to the new software modules.
3. Use of standard programming languages and interfacing techniques (e.g., inter-process communication).
4. Use of commercial software for the problems requiring online character recognition (OCR).

Deliverables and Technology Transfer Path

The system to be developed in this project will be directly applicable for on-site use at public electricity companies, including local and national suppliers. In addition, the system will be easily adaptable for use at agencies with similar requirement profiles (such as communications, gas, water, and railroad).

The technology developed here will be made available to non-academic personnel in the form of documentation and training seminars. On request, all relevant software developed in this project will be made available to the funding agency for subsequent use and distribution.

In addition, we will involve and educate a number of graduate students in this project who will then be capable of contributing to similar projects within the industry itself.

2.2.5 Previous Accomplishments

The Department of Systems Science at Johannes Kepler University in Linz has been active in the area of digital image processing for many years. Specific research in computer vision, particularly object recognition and related areas, began in 1987. Since then, several diploma theses and one doctoral thesis have been completed in this area.

Various new methods for extracting blob-like image structures were investigated in a thesis by M. Heumel [Heu92], using a local moment-of-inertia measure to effectively avoid full region segmentation. Two different hierarchical implementations of the Hough transform, which is very popular for extracting geometrical primitives, were developed in two other diploma theses. S. Grabner [Gra91] used partitioning of the original image domain to obtain a hierarchical description of straight line segments. F. Schneeweiß [Sch91] developed a hierarchical accumulator space version of both the classical and the generalized Hough transform.

As part of the ongoing research project, Dipl.-Ing. R. Sara is investigating the mathematical characteristics and potential use of qualitative 3-D surface properties obtained directly from intensity images. He has shown [Sar92, Sar93] that certain local surface properties are highly invariant under projection and changing lighting conditions. The resulting qualitative cues will be extremely valuable for 3-D object recognition. A practical implementation of this approach is currently under way.

The problem of structural feature detection and recognition under highly space-variant image transformations has been studied in [Bur92b, Bur92a]. It was shown that a wide class of spatial image transformations (cortical mappings) can be used as multi-scale representations of the original image and allow rapid detection of local and global structures. Currently, a salient-structure detection mechanism is being implemented that is based on local cooperative computation in the warped image domain.

In connection with an ongoing project funded by the U.S. Defense Advanced Research Agency (DARPA), Dr. Burger is investigating the use of Gabor Wavelet patterns for the purpose of object recognition in complex natural images. This research is expected to provide a valuable extension to the "polymorphic" description framework to be pursued here.

2.2.6 The Role Within This Project

Within this project, strong interrelations exists to topic "Robust Hierarchical Methods" (Task II.1). In return, the techniques developed in the context of *polymorphic structural feature detection and matching* is also important for structural problems in other projects of this research initiative, such as *image registration*, *motion analysis*, *cartography*, *stereoscopic mapping*, *remote sensing*, and *image fusion*. In the area of Gabor wavelets (Task II.2), there will be a close cooperation with teams in Project 1, particularly with Prof. Pichler (CAST.FOURIER and Applications) and Prof. Feichtinger (Numerical Harmonic Analysis and Image Processing).

2.3 Machine Learning for Image Understanding

Task leader:

Dr. Wilhelm Burger,
Department of Systems Science,
Johannes Kepler University, Linz.

2.3.1 General Description

Current object recognition and image understanding (IU) techniques lack flexibility and robustness. Research in *Machine Learning* over the past two decades has produced a set of powerful learning strategies, each having its own merits, weaknesses, and application domains. The use of specific ML tools in computer vision and image understanding is still in its early beginnings. It has already become clear, however, that ML techniques can be used successfully to improve system performance and robustness. Task-specific adaptation of IU systems is best done by learning on examples. Among several learning applications in IU, the issue of automatic or interactive acquisition of knowledge for object recognition (acquisition of object models) is one of the most urging and challenging problems.

In this proposal we focus on the problem of *machine learning* in connection with *structural object recognition* (cf. Task II.2). In particular, this includes the issues of

- adaptive perceptual grouping,
- learning of object models from examples, and
- representation, use, and learning of scene context.

2.3.2 Detailed Problem Statement

Learning and Perceptual Grouping

Model-based, structural object recognition methods require image data to be matched with models in the system database. Typically, the image data consists of unordered sets of simple geometric primitives like lines, arcs, and corners. It is well known that the computational complexity of the matching process is exponential in the number of image features for a given object model. *Grouping* has been shown to be an effective means for reducing the search complexity in structural matching [Low87]. However, most current grouping techniques use only perceptually motivated, low-order geometrical relationships, (such as collinearity, co-termination, parallelism, proximity, etc.), but no object model information, to assemble simple features of the same type. As a result, the full potential of grouping for solving the indexing problem has not been realized yet.

Groupings are commonly assembled by either bottom-up grouping criteria or pre-specified prototype patterns that are searched for in a goal-directed manner (e.g., [TNM87]). The work reported by Segen [Seg85] addresses some aspects of learning composite structural concepts from examples, however, no results have been shown on real images. The use of a small set of fixed bottom-up composite structural concepts allows efficient detection in images. Similar arguments hold for top-down search for specific arrangements when the number of possible objects is small. Unfortunately, a small but fixed set of intermediate structural concepts is generally not useful in different application domains. For using top-down, model-based composite structures, the number of models is restricted. In either case, the manual specification of suitable intermediate structures is difficult.

Learning Object Models from Examples

Traditionally, object models are developed by a human expert with considerable experience and knowledge about the application domain. Extensive experiments necessary for reliable performance make this process time-consuming and costly.

Two different directions have been taken in the past towards automatic object model generation: *(a)* approaches that predict 2-D appearance from 3-D models (e.g., CAD-based vision), and *(b)* approaches that build object models from the observed appearance of objects.

In *CAD-based vision*, one assumes that the exact 3-D shape of the objects to be recognized is known. Using CAD descriptions of an object, one tries to predict its appearance, in particular its visually significant features, and to devise a strategy for recognizing the objects [Goa83, HH88, LR87, Low87, Ike87]. Unfortunately, in many applications, CAD models are either not available (e.g., faces, fruits, trees) or would be too numerous to be practicable (e.g., cars in a traffic scene).

The idea of using the object's appearance for constructing a model originated in statistical pattern recognition, where the model is usually built in a so-called "training phase" from real images. The key idea is to use the *same* low-level processing steps for modeling and for recognition. [RB84, FHW86, MH87, YT77] have shown that it is possible to acquire structural object models from real images of (2-D) industrial parts.

Acquiring object models from real examples is a typical problem of *learning-by-example* [Bro81, CB87]. The central problem of learning-by-example is to create a valid generalization (i.e., a "concept") from the observed instances, while keeping the description *minimal* in some sense. While fully automatic model acquisition appears to be infeasible for several reasons [Bur89a], we propose that the learning system is guided by a human instructor in an interactive manner. The aim of the human instructor is to tell the system "what" to look for in the images and to supply additional semantic information, while the actual model (i.e., the visual concept) is built by the machine. The details of the model (which may be extremely complex and non- explicit) should be completely hidden from the instructor. The validity of the result (i.e., the model base) is primarily judged from its performance in processing other instances.

Use of Scene Context

The importance of *context* in object recognition and scene analysis has long been recognized as a powerful source of information [BK74, Tou78]. Unfortunately, successful examples for the actual *use* of context in vision have been rare (e.g., [SF91]) and are often limited to a particular task environment. Similarly, the importance of context in *speech recognition* is well known, although only few real applications exist (e.g. [KM90]).

In real scenes, objects usually do not appear in isolation but in combination with other objects. Recognizing an object can be viewed as making a decision, given a certain amount of evidence and remaining ambiguity. Since the real world is only observed through a projection, the available evidence is never complete. Humans are known to perform recognition even under minimal evidence as long as other parts of the scene can be identified. Thus object recognition is not only a problem of accumulating *local* evidence (i.e., structural object components) but capturing the evidence available in the entire scene.

2.3.3 Proposed Research Topics

In summary, the proposed work in the area of *machine learning for structural object recognition* will focus on the following research topics:

A. Adaptive Perceptual Grouping;

B. Learning Object Models from Examples;

C. Representation, Use, and Learning of Scene Context.

Adaptive Perceptual Grouping

We propose to develop a learning technique that enables us to automatically come up with new composite structural groups (intermediate concepts) that are useful in the application domain, without being specified in advance. The key features of our approach are:

1. The use of a two-stage grouping strategy that combines *(a)* domain-independent perceptual grouping and *(b)* model-based grouping with a database of high-order structural arrangements (intermediate visual concepts).
2. The use of machine learning to automatically infer the most useful intermediate visual concepts from real examples.
3. The incorporation of multiple types of primitive features ("polymorphic" features) to participate in feature groupings instead of a single feature type, which will lead to increased robustness (by providing redundancy) and indexing power (cf. Task II.2).

Learning of new structural concepts is based on the following criteria:

- *Perceptual saliency:* A concept must be perceptually salient, i.e., receive a high score in the first (perceptual) grouping stage.
- *Operationality:* A concept must be describable in terms of the operators that the model-based grouper is able to perform. For this purpose, knowledge about these operators is supplied in explicit form.
- *Simplicity:* Concepts that permit a simple description (i.e., one with few grouping steps / transformations) are preferred.
- *Recognition utility:* Only those concepts that are found to be useful in recognizing a particular object are eventually accepted.

Under the program described in Task II.2, we will develop methods for collecting structural primitives of different types (e.g., lines, arcs, parametric curves, blobs) into polymorphic groups, using a set of perceptually significant spatial relationships. The relationships (e.g., proximity, collinearity, symmetry, relative size) being used depend upon the type of elements contained in each particular group. The purpose of this initial bottom-up grouping process is to supply an ordered set of composite structures that have a high probability of being semantically meaningful.

A collection of intermediate structural concepts is created and maintained by a learning scheme based on Conceptual Clustering (CC) [SM86]. The major steps in this task are

- the development of a suitable representation for high-order polymorphic feature groups which can also express their variability,
- the adaptation of CC for learning parameterized geometric concepts and its implementation in software, and
- the development of efficient matching algorithms that can make use of the polymorphic nature of the feature groups.

The *advantages* of the proposed approach are that

(a) new intermediate concepts are developed directly in the context of the application domain using Conceptual Clustering;

(b) intermediate concepts are guaranteed to be significant for recognition and can still be extracted bottom-up;

(c) efficiency of indexing into the model-base using intermediate concepts is increased.

Learning Object Models from Examples

The goal of this project part is to investigate the possibilities of semi-automatic construction of object models from real images by a learning system. We propose that this system is supervised by a human instructor who aids the system in determining the relevant objects in the given images and who supplies additional semantic information (e.g., the names of objects).

In particular, this part of the project shall investigate:

1. Prerequisites and limitations of visual learning;
2. Generalization of principal views;
3. Fusion of information from multiple views;
4. Incremental model updating;
5. Modeling spatial indeterminacy and uncertainty;
6. Maintaining model compactness, consistency and robustness;
7. Construction of application world models (long-term memory);
8. Non-verbal interface mechanisms for interactive modeling.

Design of a long-term memory mechanism is an important issue in order to optimize the model base and to relate new images to examples studied in the past. For this purpose, we plan to employ the ITASCA object-oriented database system, which provides persistent storage of arbitrary objects (including images and executable code) and can be accessed from a variety of languages, including C, C++ and Common LISP.

Representation, Use, and Learning of Scene Context

The key issues related to the use of *scene context* are

1. Representation of multi-object scenes;
2. Clustering of scenes to scene classes;
3. Semantic object/object relationships in scenes;
4. Use of scene models as context to bias object recognition and grouping;
5. Learning complete scene classification from examples.

A *scene* is primarily characterized by the *objects* it contains. E.g., a house scene usually contains buildings, streets, trees, cars, etc., while an indoor picture will show completely different things. Of course, the number of possible scenes may be very large and thus there is no obvious way of categorizing or naming scenes. Instead, we use a probabilistic definition of a scene that is based on the probability that a given set of objects occur together in a scene. E.g., we would not say

"A *house-scene* contains the objects *house, tree, car, street.*"

but rather express the coherence of a scene as the probability that these four objects appear together, i.e.

$$P(House \wedge Tree \wedge Car \wedge Street)$$

Thus we do not classify scenes as such but we want to measure whether a certain collection of objects is familiar or not by the probability of their co-occurrence. This allows us to work with *partial* and new scenes. The key question is how we can use global context for recognizing individual objects. We assume that objects are recognized sequentially, one by one. I.e., at some point in time t, we have a set of recognized objects $Q(t) = \{C_1, C_2, \ldots\}$, where C_i is a known object (C stands for "concept"). We define $Q(t)$ as the *global scene context.*

In order to apply the global context for recognition, we define the *bias* established by the current context $Q(t)$ towards the candidate concept C_j as the probability that concept C_j occurs, given the context set $Q(t)$:

$$Bias(C_j, t) = P(C_j|C_1, C_2, \ldots) = P(C_j|Q(t))$$

The bias towards an object category introduced by the existence of other objects in the scene can be described as the *semantic adjacency* between different objects. Semantic networks [SF85] provide a useful representation formalism for this kind of knowledge. The *learning aspect* of the context problem is how to infer the semantic links in this network from a set of real scenes.

2.3.4 Previous Accomplishments

At Department of Systems Science at Johannes Kepler University in Linz, teaching and research in specific areas of machine learning has been pursued within the last few years, providing a high level of expertise with its staff and faculty members (Dipl.-Ing. Th. Müller, Dipl.-Ing. R. Mittelmann, Dr. H. Prähofer). This team is also experienced in Artificial Intelligence (AI) techniques and the implementation of large-scale AI-based systems, object-oriented design, and sophisticated user interfaces and has educated a significant number of excellent graduate students. Dr. Burger has investigated the problem of automatic model acquisition from examples [Bur89a, Bur89b] and is currently participating in a special DARPA project on machine learning in computer vision at the Univ. of California at Riverside.

2.3.5 The Role Within This Project

Machine learning is a key technology for significantly advancing the state of the art in computer vision and image understanding. Although many engineering solutions in this field are still accomplished in a traditional, static methodology, successful future designs will inevitably incorporate adaptive and learning strategies. Research in this field create additional links between image processing and computer vision on one side and artificial intelligence on the other side. It is therefore important to pursue specific research in machine learning as part of Project 2, not only for increasing the awareness for this technology, but also to create a resource of specialized expertise and make this expertise available to the cooperating partners.

Bibliography

[ABB$^+$75] A.P. Ambler, H.G. Barrow, C.M. Brown, R.M. Burstall, and R.J. Popplestone. A versatile system for computer-controlled assembly. *Artificial Intelligence*, 6:129–156, 1975.

[AF86] Nicholas Ayache and Olivier D. Faugeras. HYPER: A new approach for the recognition and positioning of two-dimensional objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8(No. 1):pp.44–54, January 1986.

[aSN90] R. Mehotra ans S. Nichani. Corner detection. *Pattern Recognition*, 23(11):1223–1233, 1990.

[BA83] P. J. Burt and E. H. Adelson. The Laplacian pyramid as a compact image code. *IEEE Transactions on Communications*, Vol. COM-31(No.4):pp.532–540, April 1983.

[BB82] Dana H. Ballard and Christopher M. Brown. *Computer Vision*. Englewood Cliffs, NJ: Prentice-Hall, 1982.

[BC82] R. C. Bolles and R. A. Cain. Recognizing and locating partially visible objects: The local feature focus method. *Int. J. Robotics Res.*, 1(3):57–82, 1982.

[BCR90] M. Bister, J. Cornelis, and Azriel Rosenfeld. A critical view of pyramid segmentation algorithms. *Pattern Recognition Letters*, Vol. 11(No. 9):pp. 605–617, September 1990.

[BF84] B. Bhanu and O. Faugeras. Shape matching of two-dimensional objects. *IEEE Trans. Patt. Anal. Mach. Intell.*, PAMI-6(2):137–156, March 1984.

[BH89] P. Baldi and K. Hornik. Neural networks and principle component analysis: Learning from examples without local minima. *Neural Networks*, 2:53–58, 1989.

[BH93] Horst Bischof and Karin Hraby. Focusing attention in hierarchical neural networks. submitted to Neural Networks and Artificial Intelligence at Substance Identification Technologies, 1993.

[BHR86] J. B. Burns, A. R. Hanson, and E. M. Riseman. Extracting straight lines. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8(No.4):pp.425–455, July 1986.

[Bie87] Irving Biederman. Matching image edges to object memory. In *Proceedings of the First International Conference on Computer Vision*, pages 384–392, London, England, 1987.

[Bis91] Horst Bischof. Modular, hierarchical, and geometrical neural networks. Technical Report PRIP-TR-9, PRIP, TU Wien, 1991.

[Bis92] Horst Bischof. Neural networks and image pyramids. In Bischof and Kropatsch [BK92], pages 249–260.

[Bis93] Horst Bischof. Learning pyramids. accepted at Neural Networks and Artificial Intelligence at Substance Identification Technologies, 1993.

[BK74] M.L. Baird and M.D. Kelly. A paradigm for semantic picture recognition. *Pattern Recognition*, 6:61–74, 1974.

[BK92] Horst Bischof and Walter Kropatsch, editors. *Pattern Recognition 1992*, volume 62 of *OCG-Schriftenreihe*. Oldenbourg, 1992.

[BK93a] Horst Bischof and Walter G. Kropatsch. Hopfield networks for irregular decimation. accepted at 17th ÖAGM, Graz, 1993.

[BK93b] Horst Bischof and Walter G. Kropatsch. Neural Networks versus Image Pyramids. Technical Report PRIP-TR-7, Dept. for Pattern Recognition and Image Processing, TU Wien, 1993.

[BK93c] Horst Bischof and Walter G. Kropatsch. Neural networks versus image pyramids. In R.F. Albrecht, C.R. Reeves, and N.C. Steele, editors, *Artifical Neural Networks and Genetic Algorithms*, pages 137–144. Springer Verlag, 1993.

[BM87] Bir Bhanu and J.C. Ming. Recognition of occluded objects: A cluster-structure algorithm. *Pattern Recognition*, 20(2):199–211, 1987.

[BP92a] Horst Bischof and Axel Pinz. The invariance problem for hierarchical neural networks. In *Neural and Stochastic Methods in Image and Signal Processing*, volume 1766 of *Proceedings of the SPIE*, pages 118–129, 1992. SPIE Symposium San Diego 92.

[BP92b] Horst Bischof and Axel Pinz. Neural networks in image pyramids. In *IJCNN'92*, volume IV, pages 374–379, 1992.

[BR80] R. Bajcsy and D. A. Rosenthal. Visual and conceptual focus of attention. In S. Tanimoto and A. Klinger, editors, *Structured Computer Vision*, pages 133–149. Academic Press, 1980.

[BR92] S. Baugher and A. Rosenfeld. Corner detection and localization in pyramid. In *Proc. Computer Vision and Image Processing*, pages 103–121, 1992.

[Bre89] Thomas M. Breuel. Adaptive model base indexing. In *Proc. Darpa IU Workshop*, Palo Alto Calif., May 1989.

[Bro81] R.A. Brooks. Symbolic reasoning among 3-d models and 2-d images. *Artificial Intelligence*, 17:285–348, 1981.

[Bur88a] Peter J. Burt. Attention mechanisms for vision in a dynamic world. In *Proc. 9th International Conference on Pattern Recognition*, pages 977–987, Rome, Italy, November 1988. IEEE Comp.Soc.

[Bur88b] P.J. Burt. Smart sensing within a pyramid vision machine. *Proceedings of the IEEE*, pages 1006–1015, 1988.

[Bur89a] W. Burger. On interactive programming of computer vision systems. In A. Pinz, editor, *Wissensbasierte Mustererkennung*, pages 23–35. R. Oldenbourg, Vienna 1989.

[Bur89b] W. Burger. Towards interactive development of object models for machine vision. In A. Pinz, editor, *Wissensbasierte Mustererkennung*, pages 49–64. R. Oldenbourg, Vienna 1989.

[Bur92a] W. Burger. Image warpings, fixations, and structural recognition. In H. Bischof and W. Kropatsch, editors, *Proc. 16th OAGM Workshop*, pages 199–213, Vienna, May 1992. R. Oldenbourg.

[Bur92b] W. Burger. *Structural Object Recognition under Space-Variant Resolution.* PhD thesis, Dept. of Systems Science, Johannes Kepler University, Linz, Austria, 1992.

[BWR89] M. Boldt, R. Weiss, and E. Riseman. Token-based extraction of straight lines. *IEEE Trans. Systems, Man, and Cybernetics*, 19:1581–1595, 1989.

[CB87] J. H. Connell and J. M. Brady. Generating and generalizing models of visual objects. *Artificial Intelligence*, 31(2):159–183, 1987.

[Cer86] J. Cerella. Pigeons and perceptrons. *Pattern Recognition*, 19(6):431–438, 1986.

[CP84] J. L. Crowley and A. Parker. A representation of shape based on peaks and ridges in the difference of low–pass transform. *IEEE Trans. Pattern Analysis and Machine Intelligence*, PAMI–6:pp.156–170, 1984.

[CVK91] J. Cooper, S. Venkatesh, and L. Kitchen. Early jump-out corner detectors. In *Conf. on Computer Vision and Pattern Recognition*, pages 688–689, 1991.

[DP88] Leila De Floriani and Enrico Puppo. Constrained delaunay triangulation for multiresolution surface description. In *Proc. 9th International Conference on Pattern Recognition*, pages 566–569, Rome, Italy, November 1988. IEEE Comp.Soc.

[DW89] J. Dolan and R. Weiss. Perceptual grouping of curved lines. In *Proc. DARPA Image Understanding Workshop*, pages 1135–1145, 1989.

[Ett88] G.J. Ettinger. Large hierarchical object recognition using libraries of paramterized model subparts. In *Conf. on Computer Vision and Pattern Recognition*, 1988.

[FE73] M.A. Fischler and R.A. Elschlager. The representation and matching of pictorial structures. *IEEE Transactions on Computers*, 22(1):67–92, 1973.

[Fer89] Cornelia Fermüller. Hierarchisches Vergleichen von Konturen. Master's thesis, Technische Universität Graz, 1989. Diplomarbeit.

[FH88] P. Fua and A.J. Hanson. Extracting generic shapes using model-driven optimization. In *Proc. DARPA Image Understanding Workshop*, pages 994–1001, 1988.

[FHW86] R. Freytag, W. Haettich, and H. Wandres. Development tools for a model directed workpiece recognition system. *Pattern Recognition*, 19(4):267–278, 1986.

[FK89] Cornelia Fermüller and Walter G. Kropatsch. Hierarchische Kontur-Beschreibung durch Krümmung. In Axel Pinz, editor, *Wissensbasierte Mustererkennung*, OCG-Schriftenreihe, Österr. Arbeitsgemeinschaft für Mustererkennung, pages 171–187. Oldenbourg, 1989. Band 49.

[FK92a] Cornelia Fermüller and Walter G. Kropatsch. Hierarchical Curve Representation. In E. Backer and E.S. Gelsema, editors, *Proc. 11th International Conference on Pattern Recognition*, volume III, pages 143–146. IEEE Comp.Soc., 1992.

[FK92b] Cornelia Fermüller and Walter G. Kropatsch. Multi-resolution shape description by corners. to appear in Shape in Picture, 1992.

[GB87] G. Reynolds G. and J.R. Beveridge. Searching for geometric structure in images of natural scenes. In *Proc. DARPA Image Understanding Workshop*, pages 257–271, Los Angeles, CA, 1987.

[GD91] G. Giraudon and R. Deriche. On corner and vertex detection. In *Conf. on Computer Vision and Pattern Recognition*, pages 650–655, 1991.

[Goa83] TC. Goad. Special purpose programming for 3d model-based vision. In *Proc. Proc. Darpa IU Workshop*, Arlington, Virginia, June 1983.

[Gra91] S. Grabner. Das Hough Verfahren zur Detektion linearer Strukturen mit hierarchischer Realisierung im Bildraum. Master's thesis, Johannes Kepler University, Linz, Austria, 1991.

[Gri90] W.E.L. Grimson. The combinatorics of object recognition in cluttered environments using constrained search. *Artificial Intelligence*, 44:121–165, 1990.

[Gri91] W.E.L. Grimson. The combinatorics of heuristic search termination for object recognition in cluttered environments. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 13(9):920–935, 1991.

[Her87] Harald Herbst. Die Auswirkungen verschiedener Reduktionsfunktionen auf Bildeigenschaften in Radar–Bildpyramiden. Master's thesis, Technische Universität Graz, 1987. Diplomarbeit.

[Heu92] M. Heumel. Ein nicht-linearer Operator zur Detektion fleckförmiger Strukturen in Bildern. Master's thesis, Johannes Kepler University, Linz, Austria, 1992.

[HH88] C. Hansen and T.C. Henderson. Towards the automatic generation of recognition strategies. In *Proc. International Conference on Computer Vision ICCV-88*, pages 275–279, Tampa FL, Dec. 1988.

[Ike87] K. Ikeuchi. Generating an interpretation tree from a cad model for 3-d object recognition in bin-picking tasks. *Intl. J. Computer Vision*, 1(2):145–165, 1987.

[JM92] Jean Michel Jolion and Annick Montanvert. The adaptive pyramid, a framework for 2D image analysis. *Computer Vision, Graphics, and Image Processing: Image Understanding*, 55(3):339–348, May 1992.

[KH88] Walter G. Kropatsch and H. Herbst. Effiziente Störungsreduktion in der Gauß-Pyramide. In Franz Pichler and Axel Pinz, editors, *Statistik und Mustererkennung*, OCG-Schriftenreihe, Österr. Arbeitsgemeinschaft für Mustererkennung, pages 22–47. Oldenbourg, 1988. Band 42.

[KJ86] T.F. Knoll and R.C. Jain. Recognizing partially visible objects using feature indexed hypotheses. *IEEE Journal of Robotics and Automation*, 2(1):3–13, 1986.

[KM90] R. Kuhn and R. De Mori. A cache-based natural language model for speech recognition. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 12(6):570–583, 1990.

[KM91] Walter G. Kropatsch and Annick Montanvert. Irregular versus regular pyramid structures. In U. Eckhardt, A. Hübler, W. Nagel, and G. Werner, editors, *Geometrical Problems of Image Processing*, pages 11–22, Georgenthal, Germany, March 1991. Akademie Verlag, Berlin.

[KM92] Walter G. Kropatsch and Annick Montanvert. Irregular pyramids. Technical Report PRIP-TR-5, PRIP, TU Wien, 1992.

[KP87] Walter G. Kropatsch and G. Paar. Aufbau einer Pyramide auf Radarbildern. In G. Pernul and A Min Toja, editors, *Berichte aus Informatikforschungsinstitutionen*, volume Band 37 of *OCG–Schriftenreihe*, pages 221–232, Wien–München, 1987. R. Oldenburg.

[KR84] L. Kitchen and A. Rosenfeld. Scene analysis using region-based constraint filtering. *Pattern Recognition*, 17(2):189–203, 1984.

[Kro85a] Walter G. Kropatsch. Hierarchical curve representation in a new pyramid scheme. Technical Report TR-1522, University of Maryland, Computer Science Center, June 1985.

[Kro85b] Walter G. Kropatsch. A pyramid that grows by powers of 2. *Pattern Recognition Letters*, Vol. 3:pp.315–322, 1985.

[Kro86a] Walter G. Kropatsch. Curve representations in multiple resolutions. In *Proc. Eighth International Conference on Pattern Recognition*, pages 1283–1285. IEEE Comp.Soc., 1986.

[Kro86b] Walter G. Kropatsch. Ein Konzept mit zwei sich ergänzenden Pyramiden. In Walter G. Kropatsch and P. Mandl, editors, *Mustererkennung'86*, OCG–Schriftenreihe der österreichischen Arbeitsgemeinschaft für Mustererkennung, pages 158–171. R. Oldenburg, 1986. Band 36.

[Kro86c] Walter G. Kropatsch. Grauwert und Kurvenpyramide, das ideale Paar. In G. Hartmann, editor, *Mustererkennung 1986*, Informatik Fachberichte 125, pages 79–83. Springer Verlag, 1986.

[Kro86d] Walter G. Kropatsch. Kurvenrepräsentation in Pyramiden. In Walter G. Kropatsch and P. Mandl, editors, *Mustererkennung'86*, OCG-Schriftenreihe, Österr. Arbeitsgruppe für Mustererkennung, pages 16–51. Oldenbourg, 1986. Band 36.

[Kro87] Walter G. Kropatsch. Elimination von "kleinen" Kurvenstücken in der $2 \times 2/2$ Kurvenpyramide. In E. Paulus, editor, *Mustererkennung 1987*, Informatik Fachberichte 149, pages 156–160. Springer Verlag, 1987.

[Kro88a] Walter G. Kropatsch. Preserving contours in dual pyramids. In *Proc. 9th International Conference on Pattern Recognition*, pages 563–565, Rome, Italy, November 1988. IEEE Comp.Soc.

[Kro88b] Walter G. Kropatsch. Pyramid research for image analysis. In Dmitrij Csetverikov, editor, *The 2nd Hungarian Workshop on Image Analysis*, Studies 206, pages 115–118, Budapest, Hungary, June 1988. MTA, Computer and Automation Institute, Hungarian Academy of Sciences.

[Kro88c] Walter G. Kropatsch. Rezeptive Felder in Bildpyramiden. In H. Bunke, Olaf Kübler, and P. Stucki, editors, *Mustererkennung 1988*, Informatik Fachberichte 180, pages 333–339. Springer Verlag, 1988.

[Kro90] Walter G. Kropatsch. Some structural properties of irregular pyramids. In preparation, 1990.

[Kro92] Walter G. Kropatsch. Irregular Pyramids. In Peter Mandl, editor, *Modelling and New Methods in Image Processing and in Geographical Information Systems*, volume 61 of *OCG-Schriftenreihe*, pages 39–50. Oldenbourg, 1992.

[KRWW] Walter G. Kropatsch, Christian Reither, Dieter Willersinn, and Günther Wlaschitz. The dual irregular pyramid. submitted to CAIP'93, Budapest.

[KT89] Walter G. Kropatsch and H. Tockner. Detecting the straightness of digital curves in $\mathcal{O}(n)$ steps. *Computer Vision, Graphics, and Image Processing*, Vol. 45(No. 1):pp.1–21, January 1989.

[KW] Walter G. Kropatsch and Dieter Willersinn. Parallel line grouping in irregular curve pyramids. submitted to CVPR'93, New York.

[KW92a] Walter G. Kropatsch and Dieter Willersinn. Curve representation in different grids. In Reinhard Klette and Walter G. Kropatsch, editors, *Theoretical Foundations of Computer Vision*, volume 69 of *Mathematical Research*, pages 141–154. Akademie Verlag, 1992.

[KW92b] Walter G. Kropatsch and Dieter Willersinn. Irregular Curve Pyramids. Proc. of the Workshop 'Shape in Picture', in press, 1992.

[KW92c] Walter G. Kropatsch and Dieter Willersinn. Representing Curves in Irregular Pyramids. In Bischof and Kropatsch [BK92], pages 233–248.

[LE92] T. Lindeberg and J.O. Eklundh. Scale-space primal sketch: Construction and experiments. *Image and Vision Computing*, 10:3–18, 1992.

[Leo93] Ales Leonardis. *Image Analysis Using Parametric Models*. PhD thesis, Faculty of Electrical Engineering and Computer Science, Univ. of Ljubljana, Slovenija, 1993.

[Lin86] R. Linsker. From basic network principles to neural architectures (series). *PNAS*, 83:7508–7512,8390–8394,8779–8783, 1986.

[Lin88] R. Linsker. Self-organization in a perceptual network. *IEEE Computer*, 21:105–117, 1988.

[LMB+90] Y. Le Cun, O. Matan, Boser B., Denker J.S., Henderson D., Howard R.E., Hubbard W., Jackel L.D., and Baird H.S. Handwritten ZIP Code Recognition with Multilayer Networks. In *Proc. of the 10.ICPR*, pages 35–40. IEEE Computer Society, 1990.

[Low85] David G. Lowe. *Perceptual Organisation and Visual Recognition.* Kluwer, 1985.

[Low87] David G. Lowe. Three-dimensional object recognition from single two-dimensional images. *Artificial Intell.*, 31:355–395, 1987.

[LR87] N.B. Lehrer and G. Reynolds. Initial hypothesis formation in image understanding using an automatically generated knowledge base. In *Proc. Proc. Darpa IU Workshop*, page 521 ff., Los Angeles, Calif., Feb. 1987.

[LW88] Y. Lamdan and H.J. Wolfson. Geometric hashing: A general and efficient model-based recognition scheme. In *Proc. ICCV-88*, pages 238–249, Tampa, FL, Dec. 1988.

[Mal89] Stephane G. Mallat. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI–11(No. 7):pp. 674–693, July 1989.

[Mee89] Peter Meer. Stochastic image pyramids. *Computer Vision, Graphics, and Image Processing*, Vol. 45(No. 3):pp.269–294, March 1989.

[MH87] D.M. McKeown and W.A. Harvey. Automating knowledge acquisition for aerial image interpretation. In *Proc. Proc. Darpa IU Workshop*, pages 205–221, Los Angeles, Calif. 1987.

[MJBR88] Peter Meer, Song-Nian Jiang, Ernest S. Baugher, and Azriel Rosenfeld. Robustness of image pyramids under structural perturbations. *Computer Vision, Graphics, and Image Processing*, Vol. 44(No. 3):pp.307–331, December 1988.

[MMR89] Annick Montanvert, Peter Meer, and Azriel Rosenfeld. Hierarchical image analysis using irregular tesselations. Technical Report TR-2322, University of Maryland, Computer Science Center, September 1989.

[MMR90] Annick Montanvert, Peter Meer, and Azriel Rosenfeld. Irregular tesselation based image analysis. In *Proc. 10th International Conference on Pattern Recognition*, pages 474–479, Atlantic City, New Jersey, USA, June 1990. IEEE Comp.Soc. Vol. I.

[MZ92] Joseph L. Mundy and Andrew Zisserman. *Geometric Invariance in Computer Vision.* MIT Press, 1992.

[Oja82] E. Oja. A simplified neuron model as a priniple component analyzer. *J. Mathematical Biology*, 15:267–273, 1982.

[Paa87] Gerhard Paar. Beibehaltung von Objekträndern in Bildpyramiden. Master's thesis, Technische Universität Graz, 1987. Diplomarbeit.

[PK88] G. Paar and Walter G. Kropatsch. Hierarchical cooperation between numerical and symbolic image representations. In R. Mohr, Theo Pavlidis, and A. Sanfeliu, editors, *Structural Pattern Analysis*, pages 113–130. World Scientific Publ. Co., 1988.

[RB84] P. Rummel and W. Beutel. Workpiece recognition and inspection by a model-based scene analysis system. *Pattern Recognition*, 17(1):141–148, 1984.

[RHW86] D.E. Rumelhart, G.E. Hinton, and R.J. Williams. Learning representations by back-propagating errors. *Nature*, 323:533–536, 1986.

[RM86] David E. Rumelhart and James A. McClelland. *Parallel Distributed Processing*, volume 1. MIT Press, first edition, 1986.

[Ros84] Azriel Rosenfeld, editor. *Multiresolution Image Processing and Analysis*. Springer, Berlin, 1984.

[RS88] A. Rosenfeld and A.C. Sher. Detection and delineation of compact objects using intensity pyramids. *Pattern Recognition*, 21(2):147–151, 1988.

[Sar92] R. Sara. The surface structure of 3-D curved objects. In H. Bischof and W. Kropatsch, editors, *Proc. 16th OAGM Workshop*, pages 129–138, Vienna, May 1992. R. Oldenbourg.

[Sar93] R. Sara. The impact of smooth surface perception capability to the structure of a vision system. *Cybernetics and Systems*, 1993. (to appear).

[Sau90] E. Saund. Symbolic construction of a 2-d scale-space image. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 12(8):355–395, 1990.

[Sch91] F. Schneeweiß. Effiziente Realisierung der Hough-Transformation durch hierarchische Verfeinerung des Parameterraums. Master's thesis, Johannes Kepler University, Linz, Austria, 1991.

[Seg85] J. Segen. Learning structural descriptions of shape. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pages 96–99, 1985.

[SF85] L. Shastri and J.A. Feldman. Evidential reasoning in semantic networks. In *Proc. IJCAI-85*, pages 465–474, 1985.

[SF91] Thomas M. Strat and Martin A. Fischler. Context-based vision: Recognizing objects using information from both 2-D and 3-D imagery. *IEEE Trans. Patt. Anal. and Mach. Intell.*, 13(10):1050–1065, 1991.

[SM86] R. E. Stepp and R. S. Michalski. Conceptual clustering of structured objects: A goal-oriented approach. *Artificial Intelligence*, 28(1):43–69, 1986.

[SMS87] L.G. Shapiro, R.S. McDonald, and S.R. Sternberg. Ordered structural shape matching with primitive extraction by mathematical morphology. *Pattern Recognition*, 20(1):75–90, 1987.

[SR89] R. Sitaraman and A. Rosenfeld. Probabilistic analysis of two stage matching. *Pattern Recognition*, 22(3):331–343, 1989.

[Tan76] Steven L. Tanimoto. Pictorial feature distorsion in a pyramid. *Computer Graphics and Image Processing*, Vol. 5:pp.333–352, 1976.

[Tan86] Steven L. Tanimoto. Paradigms for pyramid machine algorithms. In S. Levialdi and V. Cantoni, editors, *Pyramidal Systems for Image Processing and Computer Vision*, volume F25 of *NATO ASI Series*, pages 173–194. Springer-Verlag Berlin, Heidelberg, 1986.

[TK80] Steven L. Tanimoto and A. Klinger, editors. *Structured Computer Vision: Machine Perception through Hierarchical Computation Structures*. Academic Press, New York, 1980.

[TNM87] R.A. Volz T. N. Mudge, J.L. Turney. Automatic generation of salient features for the recognition of partially occluded parts. *Robotica*, 5:117–127, 1987.

[Toc87] Helmut Tockner. Syntaktische Untersuchung und Erkennung von Geraden im RULI Chain Code. Master's thesis, Technische Universität Graz, 1987. Diplomarbeit.

[Tou78] G.T. Toussaint. The use of context in pattern recognition. *Pattern Recognition*, 10:189–204, 1978.

[Uhr92] L. Uhr. Cycling logarithmically converging networks that flow information to behave (perceive) and learn. In Wechsler H., editor, *Neural Networks for Perception*, volume 2, pages 147 – 172. Academic Press, Inc., 1992.

[YK89a] Toru Yamaguchi and Walter G. Kropatsch. Structural implications of the performance of a neural network. In Axel Pinz, editor, *Wissensbasierte Mustererkennung*, OCG-Schriftenreihe, Österr. Arbeitsgemeinschaft für Mustererkennung, pages 140–148. Oldenbourg, 1989. Band 49.

[YK89b] Toru Yamaguchi and Walter G. Kropatsch. A vision by neural network or by pyramid. In *Proceedings of the 6th Scandinavian Conference on Image Analysis*, pages 104–111, Oulu, Finland, June 1989.

[YT77] M. Yachida and S. Tsuji. A versatile machine vision system for complex industrial parts. *IEEE Transactions on Computers*, 26(9):882–894, 1977.

Chapter 3

Information Fusion and Physical Models for Image Understanding and Navigation

Project coordinator:

Univ. Doz. D.I Dr. Werner Schneider,
Institute for Surveying and Remote Sensing,
University for Natural Resources, Vienna.

3.0 General Problem Statement

3.0.1 Current Situation and Motivation

The following paragraph was published in the Computer Vision part of the 'Report on Workshop on High Performance Computing and Communications for Grand Challenge Applications: Computer Vision, Speech and Natural Language Processing, and Artificial Intelligence', November 1992 [HPC92]. Participants from the Computer Vision Area were: Thomas Huang, Yiannis Aloimonos, Ruzena Bajcsy, Dana Ballard, Charles Dyer, Tomaso Poggio, Edward Riseman, and Steven Tanimoto. The paragraph is entitled 'Integration of vision modules' (italic emphasis by us). Since this paragraph is a perfect motivation for our research goals and we fully agree with it, we place it at the beginning of this section:

"A vision system should have the ability to create representations of its visible environment if it is designed to autonomously interact with its environment and perform various tasks related to navigation, manipulation and recognition. Such representations are needed for several properties of the environment like shape and motion, as well as color and other material properties. *Extraction of these geometric and physical properties of the environment on the basis of image information is an inverse problem that requires modeling of the image-formation process as well as the development of geometric and physical constraints relating 3-D information to image properties or cues.*

Research during the decade of the 1980's demonstrated that these inverse problems are undetermined or ill-posed, i.e., their solution does not exist, or is not unique and does not depend continuously on the data. For example, recovering 3-D shape using cues, such as shading, texture or contour, are ill-posed problems. The same holds true for recovering properties such as image motion and discontinuities.

It is becoming clear from work of the past few years that if a vision system integrates information from several cues, such as motion, stereo, texture, and contour, then several inverse problems become well posed and stable, simply because more information is taken into account. *A grand challenge for computer vision research in the 1990's is, therefore, the development of a sound framework for integrating vision modules, with the goal of creating robust environmental descriptions.* Although several alternatives could be outlined, it is not yet obvious how one should proceed towards the solution of this problem. However, it is known that a solution requires a vast amount of computational power so that multiple processes can cooperatively utilize various image cues and exchange information in high-speed links."

Project III combines the expertise of three Austrian groups covering several aspects of the above. At Technical University of Vienna, Dr. Pinz performs research in knowledge based vision, representation, control, and information fusion. His group will contribute an overall framework, a consistent terminology, and algorithms for information fusion at all levels of abstraction. In addition, a medical application will be carried out as a significant application of the fusion framework. At the University of Natural Resources (Bodenkultur), Doz. Schneider investigates problems of image formation, radiometric calibration, and interpretation of Remote Sensing images. He aims at the development of physical models of remote sensing image acquisition and methods of image analysis based on the inversion of these models. At Joanneum Research, Dr. Pölzleitner has broad know-how in the fields of 3D vision, real time applications, stereoscopy, and control theory. He will focus at a complex problem of autonomous navigation depending on successful integration of different sensors and visual models. Visual navigation is also addressed by Dr. Fermüller, who is currently at the Univ. of Maryland, cooperating with Yiannis Aloimonos in research on active, purposive, and qualitative vision.

3.0.2 Task Overview

Project III consists of 4 individual tasks corresponding with sections 1–4 of this proposal:

1. Information Fusion in Image Understanding
2. Physical Models for Remote Sensing Image Understanding
3. 3D Navigation and Reconstruction Based on Multiple Views
4. Visuo-Motor Coordination

Task 1 serves as the theoretical 'glue', supporting the overall framework of fusion and a consistent terminology. Furthermore, specific algorithms and representations for fusion will be developed and applied to a medical problem.

Task 2 concentrates on the formulation of a physical model of remote sensing image aquisition (accentuating the radiometric aspect) and on methods to analyse these images by model inversion.

Task 3 and Task 4 deal with navigation and motion aspects of image understanding. The main goal of task 3 is the development of an autonomous navigation system based on the 3D reconstruction from multiple views and robust hierarchical feature tracking. Task 4 studies the following capabilities of a navigational system: independent motion detection, estimation of time to collision, target pursuit, homing, and hand-eye coordination.

3.1 Information Fusion in Image Understanding

Task leader:

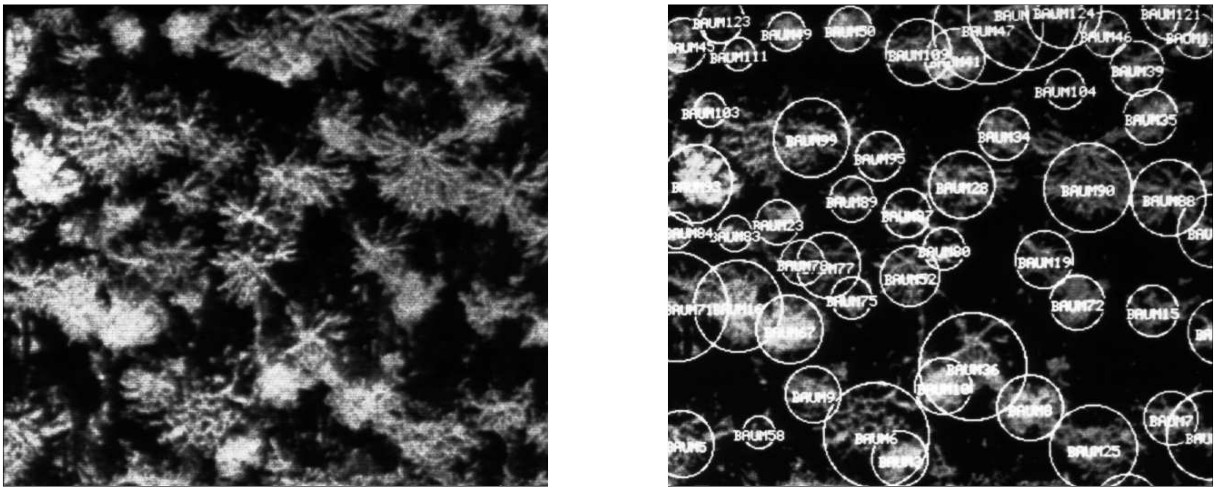
Dr. Axel Pinz,
Institute for Computer Graphics,
Technical University, Graz.

3.1.1 Problem Statement

Current Situation and Motivation

Most of our motivation has already been stated in the general problem statement, since information fusion is proposed to support an overall theoretical framework for the other tasks in this project. The integration of visual modules as an extension of the Marr paradigm [Mar82] is addressed in [AS89]. The ill-posed problem of inverting the world to image mapping is described in detail in [CY90]. The number of possible solutions of ill-posed problems in vision can be reduced by the application of constraints [CY90], by active vision [AWB87, Alo93], or by information fusion [CY90, AS89, Sze89, Hag90]. A general framework for information fusion in image understanding has not been developed yet, so that it seems to be an ambitious research goal within this program. In the following, a brief summary of our previous research activities in image understanding and information fusion is given.

- **The Vision Expert System VES:** The Vision Expert System was designed to find trees in aerial photographs, a tool of interes for the Austrian forest damage inventory developed at Universität für Bodenkultur, Vienna. Fig. 3.1.a shows an original input image, Fig. 3.1.b a result of the VES tree finder. In a dense forest image more than 90% of the trees are correctly identified. The design of the system [Pin86b, Pin86a] was awarded with the 1985 ÖAGM-Preis (ÖAGM = Austrian Association for Pattern Recognition) and its implementation led to the PhD thesis of Dr. Pinz [Pin88]. Details about the architecture of the system, its multilevel knowledge representation, and its control structure are published in [Pin89].
- **The Vision Station VS:** In contrast to the VES, the high level vision part written in LISP was isolated from the rest of the system and implemented on a dedicated LISP workstation, the Vision Station VS, controlling the rest of the image processing environment via network. Important features of this system are: Object representation in Common Lisp Object System (CLOS), process control using a graph based on a deterministic automaton, and a global blackboard [Bar89, BP90].
- **Information Fusion:** Our current research is motivated by the observation that many computer vision and image understanding processes are not very robust. Small changes in exposure parameters or in internal parameters of algorithms can lead to significantly different results. A combination (fusion) of these results is, under many aspects, profitable. An extended fusion concept dealing with different sources of information at external and internal levels was introduced and the process of fusion was defined [PB92a]. In the past, we restricted ourselves to find solutions for fusion at isolated levels. In our remote sensing application, we fuse sources at the image level [PB92b] and at the image description level [BP92b, Pin91]. In our medical application, fusion is performed at the image description level [PB92b, PD92a]. The fusion framework and future research issues are outlined in more detail below.
- **Applications:** There exists experience in two major fields of applications, both related to interpretation tasks:
  1. Remote sensing: Finding trees in aerial photographs (e.g. [Pin91]), and classification of aerial and satellite images (e.g. [BSP92, BP92a]). The fruitful cooperation with Univ. f. Bodenkultur (Doz. Schneider, DI Bartl) will be continued in this project (see also task III.2).



(a) Original Input Image (b) VES Result

Figure 3.1: The Vision Expert System VES Tree Finder

2. Medical image interpretation of the fundus of the human eye (e.g. [DPP+92, PD92b]). This application is based on a cooperation with 1.Univ. Augenklinik (Dr. Datlinger) and is outlined in more detail in sections 3.1.1 and 3.1.2.

Problem Definition

The research proposed here aims at the development of a framework for Information Fusion in Image Understanding which combines knowledge about visual sources (physical models, sensor models), several levels of abstraction (representation and matching), and classes of image understanding processes. In this framework the process of fusion is able to actively select the sources and processes of interest (i.e. to control a fusion based image understanding system) so that we term it *active fusion*. Fig. 3.2 gives an overview of our framework. We discuss each of the process and representation icons as well as their direct interrelations in more detail:

- **real world:** The real world is infinite in 4 dimensions (3-D space and time), therefore we can only observe small portions of it (scene selection). Processing of the images of these portions leads to a world description either in the computer representation or in the human brain in the case of human image understanding.

- **scene selection:** This process is related to the two questions: '*What* to look at?' and 'At which moment to observe?' (*when*). If an active sensor (robot head, moving camera, moving robot,...) is available, it can be controlled to actively select and look at the portion of the real world that is of interest in the current stage of processing. If more than one scene is selected for a subsequent interpretation and fusion process, the fusion is complicated significantly (e.g. displacement of moving objects in time series). Obviously, a selection of different scenes for fusion makes sense only, if portions of the scene are overlapping in 3-D space or in time, or if an identical object is present in several scenes.

- **scene:** We define a scene as a 3-D portion of the world at a certain moment. A scene is selected by the scene selection process and exposed by a sensor. There are constraints for the 3 coordinates

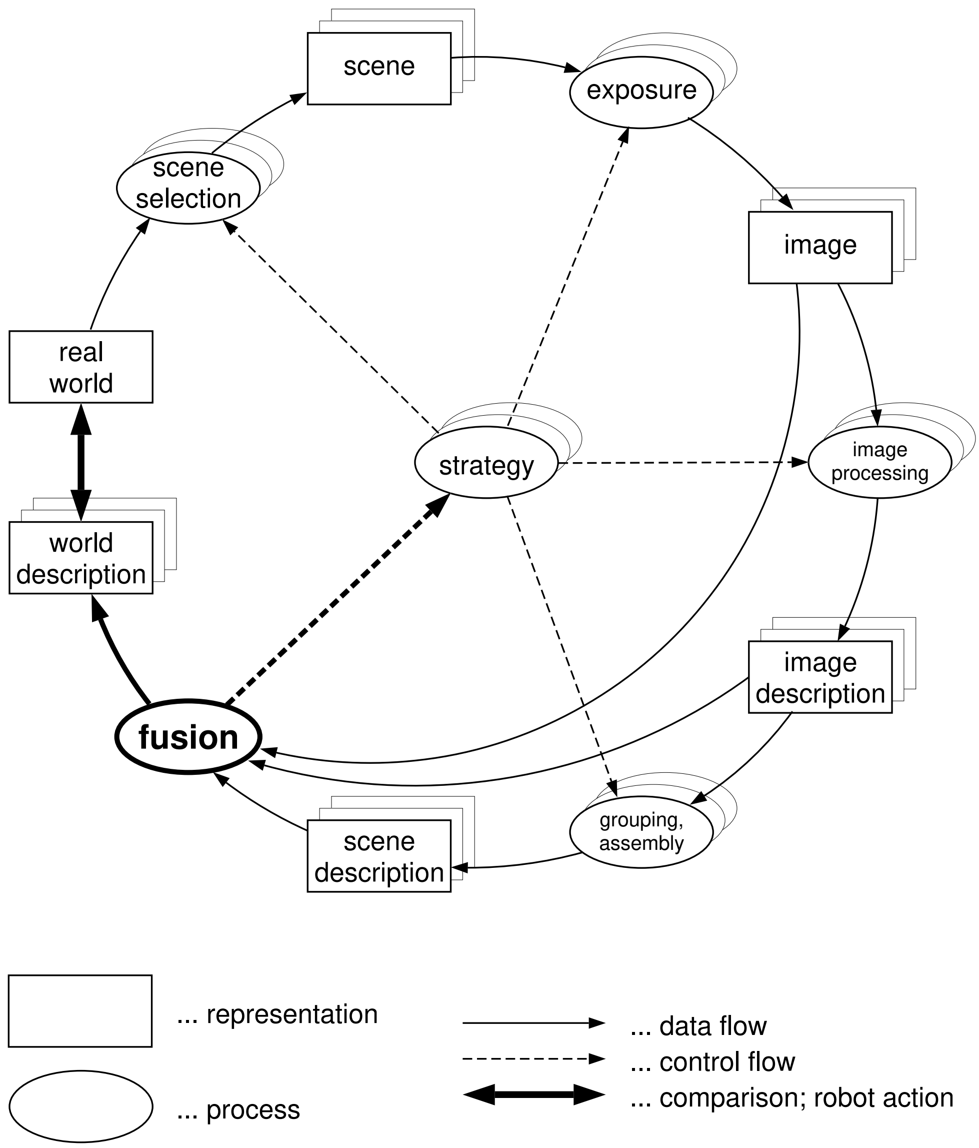


Figure 3.2: Our Framework of Active Fusion

in space due to parameters of the sensing device. A scene has important *physical properties* (scene parameters) which have to be modeled and represented explicitly for the subsequent interpretation processes. *Physical models* which fulfill these requirements have to be developed.

- **exposure:** The scene is exposed to sensing devices. Properties and parameters of the sensors have to be modeled (*sensor models*, exposure parameters). Here, the question '*How* to look at a scene?' is answered, so that the process of exposure can also be viewed as a sensor selection process. Besides the control of an active sensor, there are several other possibilities of 'quasi active' sensor selection. In remote sensing, an interpretation system can select the most appropriate out of several sources of information (satellite image, aerial photograph, radar image, scanner image, certain spectral band). In our medical application (see section 3.1.1), there are many different sources of information about the fundus of the human eye, and, depending on the current questions in the course of a diagnostic interpretation, information has to be extracted from different sources.

- **image:** As a result of the exposure process, an image is a 2-D representation of the 3-D scene. Thus the inverse problem of the recovery of the scene often is underdetermined or ill-posed. In our framework we try to find a unique solution by *combining (fusing) information* from several different sources. Scene and exposure parameters that lead to an image have to be represented explicitly together with the image itself to enable proper access by the subsequent fusion process.

- **image processing:** At this stage of processing, many different processes and sequences of processes are applied to the input images. This leads to a segmentation of the images into tokens. Features are extracted and assigned to the tokens, resulting in a kind of 'intermediate symbolic representation' called image description. To control the complete image understanding environment, these processes and sequences of processes have to be *grouped into categories* and their *properties* have to be *represented explicitly.*

- **image description:** In our fusion based image understanding approach, the control strategies allow multiple processes at all levels resulting in multiple outputs. Even in the most simple case of only one input image, we can extract many different image descriptions by applying different sequences of processes to it. In general, we need a representation that enables us to *compare* these descriptions (e.g. *match* certain image objects from different image descriptions), by comparing features of objects (shape, color, texture, size, location,...) and relations between objects (spatial, temporal). All features and relations in image descriptions are 2-D. Scene and exposure parameters are part of the image description.

- **grouping, assembly:** Grouping and assembly processes work at a higher level of abstraction than image processing. They use 2-D objects and their features to derive 3-D scene descriptions. Like for image processing, the grouping and assembly processes must be categorized and their properties must be represented. In this terminology, the term *image analysis* can be viewed as a superconcept of image processing and grouping and assembly processes.

- **scene description:** Everything that has been said about image descriptions applies here as well. The major difference is that 3-D object models as well as 3-D features and relations have to be represented.

- **fusion:** So far, we have described four different levels of processing and of representation. Diversity is possible and even desired at all of these levels, so that multiple processing and multiple representations occur at all levels of abstraction. The purpose of the process of fusion in our framework is twofold: Strategy selection, and combination of information from several representations. The *strategy selection* is required to limit the number of parallel paths, thereby controlling the image understanding system and avoiding well known problems of computational complexity and combinatorial explosion. The core problem of fusion, the *combination of information* from several representations, can be approached in many different ways. To be able to select the most valuable portions of information from the many sources of input, quality measures on the basis of probabilities for all elements of all representations are required.

- **strategy:** A strategy consists of at least one *sequence of processes* (scene selection, exposure, image processing, grouping, assembly). It is selected and controlled by the fusion process in a top down manner.

- **world description:** As a final result of processing and representation at all levels, the process of fusion comes up with a current description of the world as it was sensed by the system. From the point of view of the fusion process, this world description is better than any single one of the sources used as fusion input, and the best new output that can be found with respect to some quality measures. This world description can be related with the real world situation, either directly by having a robot interact with the real world, or indirectly by a human expert performing the comparison. If there is a robot action changing the real world, the whole circle of processing may start again.

Medical Application

Besides being a strong support for the other tasks in this project, the theoretical framework of information fusion developed here will be tested and refined using a concrete medical application: *The fusion of essential features of the human retina based on Scanning Laser Ophthalmoscope (SLO) images.* There has been an intensive scientific cooperation between Dr. Axel Pinz (TU Wien) and Dr. Peter Datlinger (1.Univ. Augenklinik, Univ. Wien) during the past three years. We started with the examination of age related macular degeneration, a disease of the fundus of the human eye, using fundus photographs and angiography photographs as our data sources [DPP$^+$92]. Currently, we use a *Scanning Laser Ophthalmoscope (SLO)* [Nas91] to acquire a variety of fundus data:

- Infrared reflexion images for layer-by-layer fundus analysis,
- Argon-blue image for the detection of the foveal structure,
- Fluorescein angiography to reveal location of subretinal leakage and fluid,
- Indocyanine-green angiography in cases of ill-defined neovascular membranes,
- Static scotometry to measure scotoma (macular areas with a loss of visual function).

Conventionally, these data are 'fused' in the brain of the physician, forming a *map of pathological changes*, which is the basis for an accurate laser photocoagulation treatment of the patients.

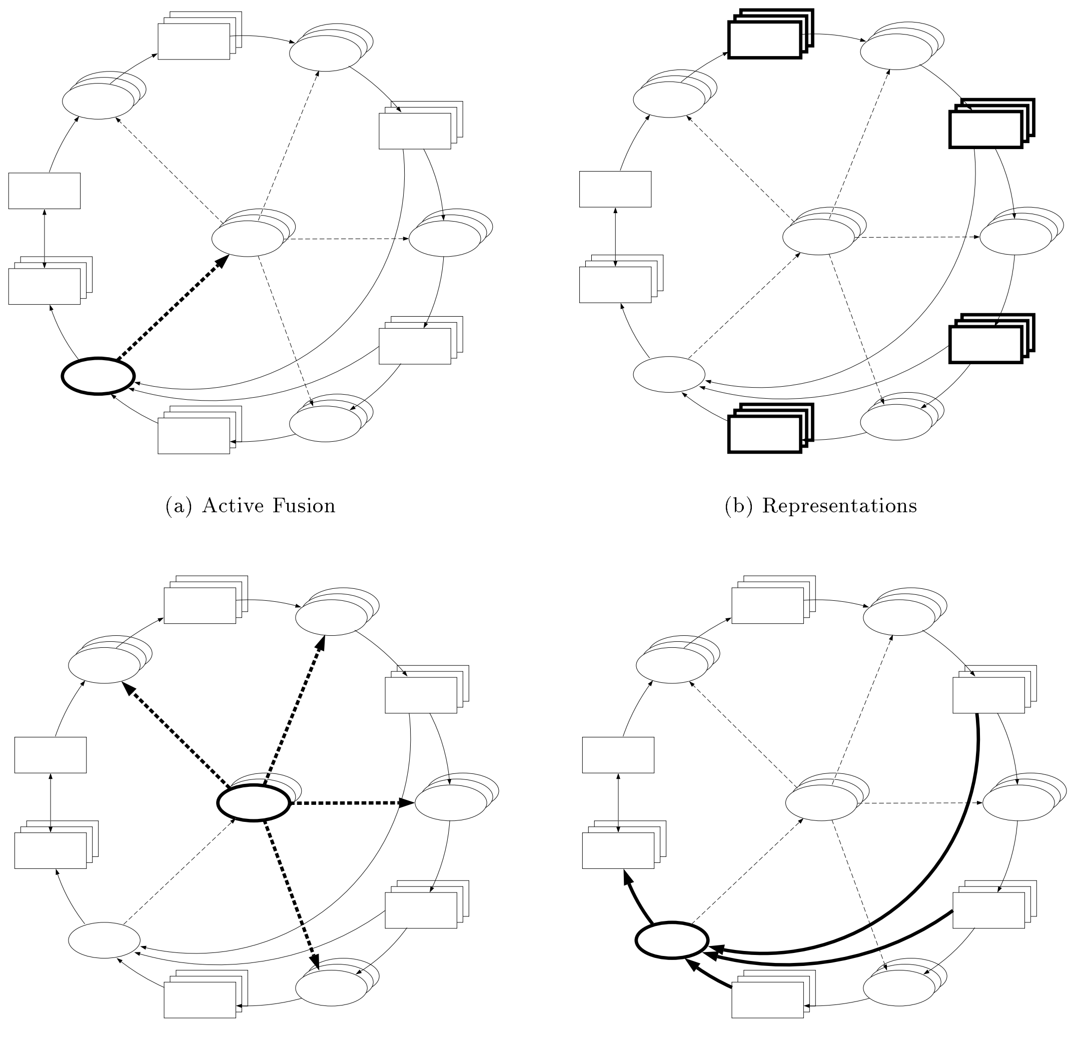
In our previous work, we have focused on the manual registration of these source images to identical geometry, and on the extraction of essential retinal features [PD92b, DP92]. In [PD92a], we describe a fusion of already registered features, and in [PB92b], we view the SLO image analysis in our framework of information fusion in image understanding and line out a possible approach to automatic registration.

Summing up, the SLO examinations are taken by different sensors. Most of the sensors produce 2-D images, but the scotometry data are already in an intermediate symbolic format. Physical properties of the fundus of the eye and sensing properties are very well known. Each of these sources is used to reveal some essential retinal features. Many of the features are only present in one source, a few ones (e.g. blood vessels) can be detected in most of the sources and therefore be used to establish the necessary spatial relations for a subsequent fusion process. The 'map of pathological changes', which can be the fusion output, is required by the physicians for accurate diagnosis and succesful treatment of the disease. From these facts, the SLO application is complicated and at the same time sufficiently well defined and constrained to be an excellent test applicaton for our framework of active fusion.

3.1.2 Research Goals and Methods

We plan to concentrate our research on the four areas listed below with the main interest in the emphasized topics (in Fig. 3.3 the most relevant components for each topic are highlighted):

- Active Fusion: Fusion model, source and process selection, framework (see Fig. 3.3.a)
- Representations, levels of abstraction, matching (see Fig. 3.3.b)
- Control, grouping of classes of processes, hierarchies of processes, (see Fig. 3.3.c)
- Combination of information from several sources (see Fig. 3.3.d)



(a) Active Fusion (b) Representations

(c) Control (d) Combination of Information

Figure 3.3: Proposed Research Topics: Fusion

Active Fusion: Fusion model, Source and Process Selection, Framework

The fusion model (Fig. 3.2) has to be defined in detail. To be able to actively select sources, strategies, and processes, a first *evaluation of the current situation* is required. Important questions to answer at this point are: What is the goal of the vision task (e.g. which kind of objects to find)?, What is the sensing situation (physical properties of the scene, available sensors, advantages and disadvantages of each possible source)?, Which strategies are applicable?. After this evaluation, the strategies are *activated and controlled* by the fusion process in a top down manner. The question is, how to activate several sources/sensors/strategies so that there is sufficient diversity for a successful combination, and at the same time, how to efficiently constrain the space of possible processes and representations to avoid unacceptable complexity.

At the University of Massachusetts, an exhaustive bottom up search of the tree of possible strategies and processes has been investigated in a similar image understanding framework. In [DHR93], the UMass Schema Learning System (SLS) for acquiring knowledge-directed object recognition strategies from training images is described. In contrast to our model, this system first performs a search on training images (training phase), then selects one precompiled strategy in the application case, and finally applies only this *one* strategy. Besides wanting to avoid the exhaustive search, we believe, that there is no single strategy to solve the ill-posed problem, so that we have to *activate several strategies* and fuse the results. This proposed research on active fusion and the fusion model will continue our previous work on a general model of information fusion in image understanding [PB92a, PB92b, BP92a].

Representations, Levels of Abstraction, Matching

In [PB92a] we define the levels of abstraction of our fusion model. These are different levels of representation (scene level, image level, image description level, scene description level, and world description level). As it is shown by the multiple boxes of Fig. 3.3.b, many different representations can occur at each level. Besides making a decision how to represent information at each level in our system (e.g. pixel images, tokensets, sets of object frames), a definition of all explicit information that is necessary to compare and match elements from different representations is required. Our previous work on object representation [Pin88, Pin89, Bar89] resulted in a frame-based [Min75] representation. This representation will be extended to capture all levels of abstraction and it will be embedded into the KBVision system (this knowledge based vision system was acquired in FWF project P8785PHY).

Control, Grouping of Classes of Processes, Hierarchies of Processes

In the same way as there are levels of representation, one can identify levels of processes. Besides this 'natural' grouping of processes into levels, a grouping into classes of processes at each level is required. A class could consist of similar processes, or of processes which have similar representations as input and as output. To control sequences of processes by strategies, we need to introduce a hierarchy of processes, where a process at a higher level can *select, invoke*, and *control* several processes at lower levels. In our framework, the active fusion process resides on top of this hierarchy of processes. Similar approaches have been proposed by Matsuyama (hypergraph [Mat88]) and by Burt (pattern tree [Bur88]). Our previous work led to a graph-based control structure for the Vision Station VS [Bar89, BP90]. The knowledge module of the KBVision system includes 'strategies' which are represented in schemas. On the basis of these KBVision 'strategies', representation and control structures covering our hierarchy of processes will be developed.

Combination of Information From Several Sources

Finally, when all the framework is set up and working, the fusion process has to handle inputs from all representational levels. Information from these multiple sources has to be selected, matched, and combined. Many theoretical approaches to the problem of combination of information have been published. In [AS89], Aloimonos discusses two methods, information compression and Markov random fields. Markov random fields for fusion have been investigated by many authors (e.g. [CY90, Sze89]). Bayesian models with their prior, sensor and posterior models seem to be a widely used and well suited approach. However, in most cases the authors deal only with fusion conditions where identical geometry can be assumed for the different sources of the fusion process (e.g. [CGH92, Toe89, Toe90]). We need *algorithms to combine sources of different geometry and at high (symbolic) levels of abstraction.* At the lower levels of our framework, the generalized evidence processing theory of Thomopoulos [Tho90] might be applicable, but a theory for all levels of abstraction is still an open problem to be addressed in our future research.

Medical Application

Our research goals directly related to diagnostic improvements are:

- **Extraction of essential retinal features:** Starting from our previous results ([PD92b, PD92a]), we will include anatomical and pathological knowledge to get more robust results for a variety of patients and features. Examples of such knowledge are the diameter of the human fovea and the distance between fovea and papilla.
- **Automatic registration:** Features which are stable and available in all source images have to be found, located, and used as control points for registration. Junctions of blood vessels, which have a similar, but not identical appearance in the source images will be processed, resulting in comparable tokens in image and scene descriptions. A fusion process which is able to match these descriptions can also be used to generate pairs of control points.
- **Map of pathological changes:** Finally, having extracted all essential retinal features and established the geometric relations between all sources, the result of the fusion process will be visualized (superimposition of all essential features). This graphic can either be used for efficient operation planning or as an overlay of previous examination results (e.g. pre-operative) over current SLO images (e.g. post-operative examination).

Besides these 'direct' research goals, this application will support a continuous verification and refinement of the theoretical fusion framework under development.

3.2 Physical Models in Remote Sensing Image Understanding

Task leader:

Doz. Dr. Werner Schneider,
Institute for Surveying and Remote Sensing,
University for Natural Resources, Vienna.

3.2.1 Problem Statement

Introduction: Image Understanding Seen as Inversion of Image Generation

In the following discussion, the terminological framework introduced in task III.1, section 1.1.2 is used: The image of a scene is generated in the exposure process using a sensor. Detailed knowledge on the physics of this process is available in principle and can be formulated in a physical model of image generation (= exposure in Fig.2, task III.1). The task of image understanding with the aim of producing a scene description from the image ('image processing' and 'grouping, assembly' in Fig.2, task III.1) can then be seen as the *inversion of the image generation process.* It should be possible to accomplish this task or at least to support its solution by inverting the physical model of image generation.

The Importance of Image Understanding by Physical Model Inversion for Remote Sensing Image Analysis

Among the various potential applications of image understanding based on the inversion of a physical model of image generation, optical remote sensing image analysis is most promising and well-suited for research in this field, for the following reasons:

1. Compared with other applications, the average remote sensing image acquisition situation is relatively simple as far as geometry is concerned (uniform view direction, objects (=terrain) can be taken as flat in many cases). The analysis of remote sensing images is therefore not further complicated by geometrical problems.

2. The physics of remote sensing image acquisition is

   - known in principle,
   - essential for detailed interpretation of the images,
   - too complicated, however, to be used explicitly in 'conventional' automatic image analysis.

   The physical mechanism of image acquisition is well described in the literature [SD78, Sla80, Asr89, Ree90]. It starts from the terrain illumination by direct solar radiation and by atmospherically scattered diffuse radiation, comprises the (directionally and spectrally varying) reflexion at the terrain surface, includes the atmospheric path radiance, and takes into consideration sensor parameters such as the spectral sensitivity of the detectors in the different channels etc. – . The first textbook on this topic in German is coauthored by the principle investigator of this project [KS88].

   The human interpreter, analysing remote sensing images visually or interactivly on an image processing system, uses knowledge about the physical mechanism (including geometrical models) of the image generation process. The main advantage of this is the fact that scene-independent knowledge (e.g. on spectral reflectance characteristics of surfaces) can be used in the interpretation process, and that disturbing factors (e.g. atmospheric influences) can be compensated in a quantitative manner.

   In automatic analysis this knowledge up to now is used, if at all, in a very coarse and implicit way only. For example, land use classification of an optical satellite image may be based on the vegetation index being defined as the ratio of an infrared and a red channel. In this case, one makes use of 2 types of knowledge: (a) that there are disturbing multiplicative influences from the atmosphere and from the illumination on uneven terrain - which cancel out by taking the ratio - , and, (b) that differences in the terrain vegetation cover can be recognized in terms of differences in the ratio of infrared and red reflectance.

3. Fully automatic analysis of remote sensing image data is highly desirable because of the large data volumes, the high expenditure in visual or semiautomatic, interactive interpretation and the shortage of expert interpreters.

4. The parameters of a physical model of remote sensing image acquisition are of interest not only as input parameters for performing an interpretation. Some of these parameters may also emerge as output parameters from the analysis procedure (e.g. atmospheric parameters unfluencing the spectral signatures). They may subsequently be used in the evaluation of images from a different sensor. Exchange of model parameters in this sense is a sort of information fusion.

5. The experience obtained from the use of physical models in optical remote sensing image understanding may be of even higher importance in the analysis of other image data (e.g. Radar images), where the interpreter cannot refer to visual practical knowledge at all, but has to rely on physical explanations completely.

Outline of a Physical Model of Image Generation

The formulation of a *physical model* is possible only for restricted classes of sensors and exposure situations. The following model proposed for this work is formulated for remote sensing image acquisition in the visible, near-infrared and middle-infrared part of the electromagnetic spectrum with photographic or electrooptic image sensors.

The various *data components* of the model and the *data flow* are shown in Fig. 3.4. The data flow from the scene to the digital image corresponds to the exposure process in Fig.2, task III.1. The attempt to invert the model means that one wants to deduce the scene description from the digital image. Scene and scene description in this sense coincide in this model (compare [MS92]).

For every component, the data representation is given in Fig. 3.4. The equations describing quantitatively the relationships between the components are not given here, their formulation (at various degrees of detail) is, however, straightforward [KS88] with some exceptions (e.g. scene description). For these, restrictions for the range of application will have to be made.

- The *scene* is modelled as a set of objects $\{O_m\}$ characterized by their geometric properties and their surface materials.
- A data base of spectral *reflectance* curves $\rho_m(\lambda)$ for the different materials m as a funcion of wavelength λ is given.
- The spectral *irradiance of the scene objects*, $E(\xi, \eta, \lambda)$, depends on the *radiation source* (e.g. sun) with its spectral radiant intensity $I(\lambda)$, and the *radiant transfer* between radiation source and scene. For many applications, the atmosphere and its absorption and scattering parameters will determine this radiant transfer. The relevant parameters in this case are the spectral volume scattering function $\beta_\vartheta(\lambda)$ and the spectral extinction coefficient $\beta_{ext}(\lambda)$ [KS88]. This irradiance depends also on the scene geometry. ξ and η are local coordinates on the scene object surfaces.
- The *geometric sensor parameters* are specified in a notation used in photogrammetry (see e.g. [KW89]): x_0, y_0 and c represent the *interior orientation* and the coordinates X_0, Y_0 and Z_0 together with the angles ω, φ and κ are the parameters of the *exterior orientation* of the sensor (e.g. photographic camera).

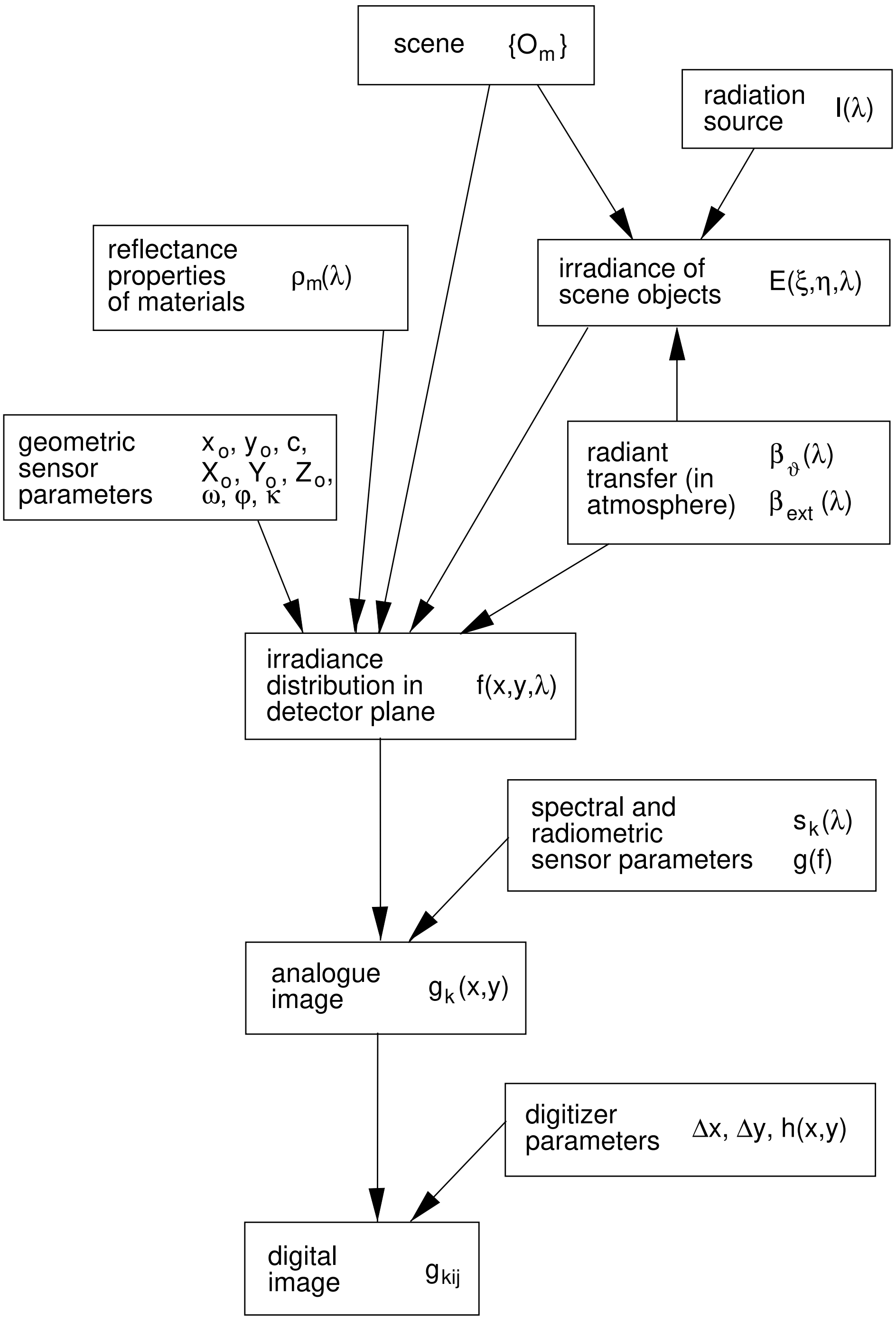


Figure 3.4: Physical Model of Image Generation

- The spectral *irradiance distribution in the detector plane* of the sensor, $f(x, y, \lambda)$, can be calculated from the data components discussed up to now.
- Introducing the spectral sensitivity functions $s_k(\lambda)$ of the spectral channels k of the sensor and taking into consideration detector nonlinearities $g(f)$, one obtains the *analogue image* $g_k(x, y)$ for each channel.
- Digitizing the analogue images will finally yield the *digital images* g_{kij}. The *digitizer parameters* are the pixel spacings Δx, Δy and the *sampling function* $h(x, y)$.

A number of simplifications and reductions have been made when formulating this model. (E.g.: The reflectance curves $\rho_m(\lambda)$ cannot give a complete description of the reflectance properties of the scene objects. Bidirectional reflectance distribution functions would be necessary for this.) Various adaptations and extensions of this model can be envisaged.

Geometrical models

A special problem in the quantitative analysis of optical satellite data are the *'mixed-pixel'* spectral signatures: The spectral signature of every pixel covering regions of diverse land use categories are a mixture of the spectral signatures of the categories concerned. A special module for the formation of mixed pixel signatures based on a geometric model therefore should be included in the image generation model. Inversion of such a geometric mixed-pixel model makes it feasible to classify land use with subpixel accuracy.

Preliminary experiments with *geometric subpixel models* have been performed at the Institute for Surveying and Remote Sensing [Sch93]. In these investigations, every pixel of an image is analyzed within the context of its 8 neighbouring pixels. Certain models about the scene pattern within this cell of 3×3 pixels are assumed, e.g.

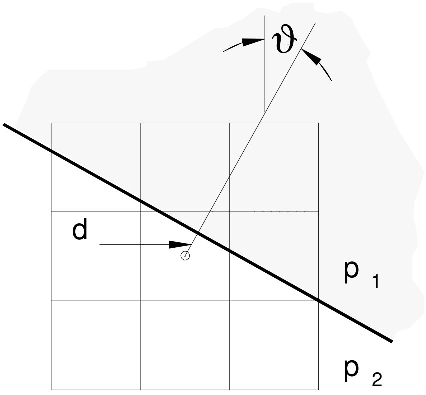
- model (pattern) no.1: The scene is composed of 2 homogeneous areas separated by a straight boundary. 4 parameters are necessary to describe this model (Fig. 3.5.a).
- model (pattern) no.2: The scene is composed of 3 homogeneous areas delimited by 2 straight boundaries at right angles. 6 parameters are necessary to describe this model (Fig. 3.5.b).
- An arbitrary number of further models can be defined. The maximum number of parameters of any model must not exceed 8.

For every model, the parameters are computed from the 9 given pixel values by least squares adjustment. The sum of squares of the residuals is a criterion for the appropriateness of the model. The model most appropriate in this sense is finally selected. The (central) pixel is replaced by a raster of smaller pixels (e.g. 3×3 pixels of 1/3 of the original pixel size), whose values are computed from the selected model. If no model could be found (i.e. if the residuals were too large for all models), then the new, small pixels are all set equal to the one original large pixel.

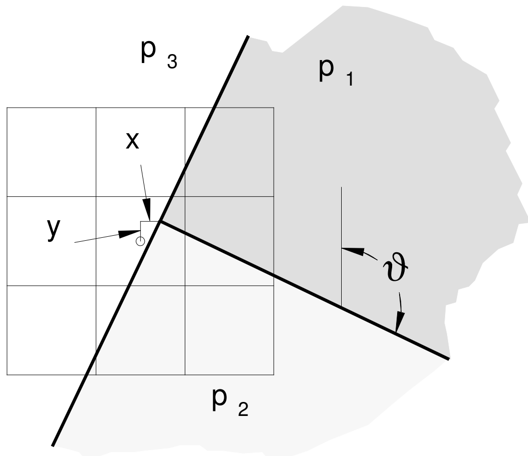
Following this procedure, LANDSAT TM scenes of agricultural regions could be produced with 'resampled' pixel size of 10m. Detail rendering in these images is impressively improved and the number of mixed pixels is greatly reduced [Sch93] (see Fig. 3.6). The inclusion of this geometric subpixel model in an image understanding system should enhance the capabilities of the system considerably.

Implementation Problems

The main difficulties in implementing an image understanding system based on a physical model of image generation are seen in:



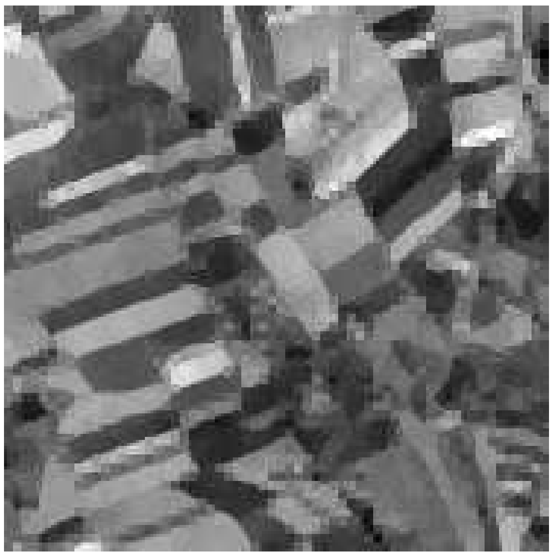
(a) Model No.1, Described by4 Parameters



(b) Model No.2, Described by6 Parameters

Figure 3.5: Geometric Subpixel Models

(a) Original Image



(b) Subpixel-Resolution Image

Figure 3.6: Landsat TM Image, Band 4

1. the representation of the different types of data in the model: For some of the components of the model, data representation is straightforward (e.g. radiation source: The spectral radiant intensity of the sun, $I(\lambda)$, is constant and well known). For other components, data representation is extremely complicated and cannot be implemented in full generality (e.g. the scene description itself). For some types of data, *fuzzy* information is available from the start, which can be refined in the process of running the model (e.g. reflectance properties of materials: It is known, that the spectral reflectance of vegetation exhibits a small relative maximum in the green part of the visible spectrum, that the reflectance of bare soil increases monotonously with increasing wavelength, and that there are no discontinuities in these spectra. The exact shapes of the spectra, however, may vary from object to object.).

2. the detailed quantitative formulation of some parts of the model, e.g.

   - the geometric object description and the determination of the irradiance of the scene objects $E(\xi, \eta, \lambda)$ as a function of their geometry for man-made objects, to be used for describing the image formulation process for large-scale (photographic) images, and
   - bidirectional reflectance characteristics of terrain cover.

   It seems to be necessary to restrict the scope of a physical/geometrical model with regard to these problems.

3. the control strategy of the system: Various methods exist in principle for the solution of inverse problems, in particular for model parameter estimation (e.g. [Tar87]). In applying these methods for the image analysis task under discussion here, the problem of combinatorial explosion [Tso90] is of major concern. To alleviate this problem, an approach based on a symbolic description of the image contents ('image description' in Fig.2, task III.1) is advantageous.

State of the Art and Previous Experience of the Investigators

Many attempts to develop image understanding systems realizing various aspects of the ideas presented here are reported in the literature(e.g. [Oht85, RJ88, Nie91]. Keywords to this field are *knowledge based image analysis*, *computer vision*, *image understanding* and *model inversion*. The general characteristics of these efforts are the attempt to enhance the flexibility of the system

- by explicit formulation of knowledge,
- by separation of knowledge from control,
- by using a scene-independent representation of spectral knowledge, etc.

The investigations apply to various applications, in particular to those where the diversity of the possible scenes is restricted, e.g. in industrial automation with a limited number of simple object prototypes, controlled illumination, etc. [BN86, Kan87, LE89].

The attempts to apply model inversion and image understanding methods to remote sensing image analysis are still limited and they are often restricted to special problems [Wha89, MH90, FJLS92, PH92]. A general image understanding system based on physical (radiometric) model inversion in the form outlined here, to our knowledge has not yet been implemented.

At the Institute for Surveying and Remote Sensing experience from previous work relevant for this project can be summarized as follows:

- Image Understanding: The activities in this field are described in more detail in section 1.1.1 of the task III.1, as they have been initiated by Dr. Pinz working at the Institute for Surveying and Remote Sensing previously. The work has started with the development of the *Vision Expert System*

(VES) [Pin86b, Pin86a] and has continued with the *Vision Station (VS)* [Bar89, BP90]. Current experiments concerning the fusion of remote sensing images [BP92b, BS93, BPS93] are performed within the KBVision environment.

- Information Fusion: From the beginning, the theoretical work in the field of fusion has been accompanied by investigations on remote sensing data at the Institute for Surveying and Remote Sensing. Attempts to improve the accuracy of land use classifications by fusion of results of a maximum likelihood classifier [Wag91] and a neural net classifier [BSP92] on a pixel-by-pixel basis are documented in [PB92b] and [BP92b].

  In a recent application [BSP93] a Landsat TM image and an image taken during the AUSTROMIR mission by the photographic satellite camera MKF6 have been fused. The images from these two sensors differ in scale as well as in orientaton. From both sources, forests have been extracted independently and represented as regions with features describing, among others, the morphological characteristics of the regions. These features are compared to find candidates for pairs of corresponding regions. In another approach [BS93, BPS93] the spatial relations between pond regions (within clusters) are compared to find correspondences. Based on the established correspondences, the scaling factor, the rotation angle and the translation vector between the images are estimated. This fusion of symbolic information has the advantage that the amount of data to be processed is reduced and that differences due to diverse sensing devices are diminished.

- Physics of remote sensing image acquisition, radiometric calibration: A comprehensive programme is pursued at the Institute for Surveying and Remote Sensing since more than 10 years for developing tools for quantitative, computer-assisted thematic analysis of colour infrared aerial photos and other optical remote sensing images for applications in the various disciplines represented at the University for Natural Resources (Universität für Bodenkultur) which are agriculture, forestry, water management, regional planning, landscape planning, environmental monitoring etc.

  An important prerequisite for the quantitative thematic analysis of remote sensing images is their radiometric calibration: An image is termed *radiometrically calibrated* if the algorithm for converting pixel values to (directional) spectral terrain reflectance values is known. In this context, extensive know-how in terrestrial spectroradiometric reference measurements and in the use of these data for radiometric calibration of both photographic and optoelectronic image data based on physical models was obtained [SP84, Sch86a, Sch91]. The essence of these experiences is documented in a two-volume German textbook on Remote Sensing [KS88, Kra90]. The radiometric calibration method was also applied in the project FEM-Austromir (a remote sensing experiment carried out on the occasion of the flight of an Austrian cosmonaut on board of the Russian space station MIR in October 1991 [KSEL92, KKS+92]).

- 'Conventional' digital remote sensing image analysis: These activities started with the project 'Rechnerunterstützte objektivierte Luftbildinterpretation' (computer-assisted objective aerial photo interpretation) financed by the FWF (P4489). A digital image processing system was installed and comprehensive software for the preprocessing and analysis of remote sensing images was developed in this and in follow-on projects ('Schwerpunkt Fernerkundung' of the FWF (S38)) [Sch81, SHP83, SHH84, Sch86b, Sch89a]. More recently, modules to be used and controlled by the systems VES and VS (see 'image understanding', above) as well as for spectral and textural image analysis with simulated neural networks [BSP92] were added. A system for combined visual-digital analysis of colour-infrared aerial photos for forest damage assessment was developed to operational maturity. For this purpose, an analytical stereoplotter has been equipped with a CCD-camera for real time photo digitizing. The colour information thus obtained is also analyzed in real time, simultaneously with the stereo-shape interpretation performed by the human interpreter. Radiometric standardization as a prerequisite for time series evaluation to assess forest damage dynamics is also provided for in this method [Man92, Sch89b].

3.2.2 Research Goals and Methods

The overall research goal of this project is to obtain knowledge and experience in the development of an image understanding system for automatic analysis of remote sensing images. The system should imitate as far as possible the reasoning of a human image interpreter and, by fully exploiting the radiometric image information in a quantitative way, in some respect even surpass the capabilities of a human interpreter.

The goal will be pursued by

- theoretical studies on the detailed formulation of a physical model of remote sensing image acquisition and on the problems of knowledge/data representation and control in an image understanding system based on inversion of the physical model, including a geometric model for subpixel resolution of mixed pixel signatures,
- software implementation of such a system for remote sensing image analysis for land use classification, and
- explorative use of the system with different types of image data (airborne and spaceborne, photographic and electrooptic), including the compilation of a-priori knowledge for these problems.

Due to considerable interrelations between these subtasks, they cannot be treated in a purely sequential way. Instead, after some initial theoretical investigations, software implementation, explorative use and refinement of the representation and control structures will have to be pursued in parallel.

Representation

For every component of the physical model as shown in Fig. 3.4, knowledge representation has to be found. These representations will be chosen with a varying degree of detail, starting with rather coarse concepts and refining the representations later depending on the experiences obtained in the practical implementation and the explorative use:

The scene object representation will be defined on the basis of the data types provided for in the KB-Vision system. Additional provisions will have to be made to allow for terrain height information as well as for the representation of man-made objects in large-scale (photographic) images. The incorporation of subpixel resolution information into the object representation will be investigated.

Suitable representation for spectral knowledge will have to be found, allowing for varying spectral resolution as well as for*qualitative* spectral information (e.g.: 'Reflectance in the near infrared is much higher than in the visible part of the spectrum'). The basic question on the possibility of and the need for detailed description of directional reflectance effects will be addressed.

A certainty measure will have to be associated with each element of knowledge.

The compilation of *a priori* knowledge (e.g. catalogue of reflectance properties of surface materials) will also be part of this subtask.

Control

Theoretical investigations will be performed to obtain rough estimates of the computational expenditure for image analysis by inversion of the physical model of image generation

- for different inversion strategies (e.g. simulated annealing),
- for different levels of abstraction of the image description (e.g. pixel level, token level),
- for varying degrees of detail of the individual data components (e.g. allowing for directional reflectance effects),

- for a varying number of unknown or uncertain parameters (e.g. atmospheric conditions known and fixed, or variable?).

Based on these estimates, decisions on the control strategies to be implemented in software will be made. Practical experiments should then lead to improved and refined control structures.

Software Implementation

The previously developed concept for representation and control (Vision Station, [Bar89], [BP90]) will be adapted for connection to the KBVision System. Representations of image descriptions and scene descriptions are based on the Common Lisp Object System (CLOS). This tool will also be used for representations concerning the special requirements of the physical model. In the control graph the use of parameters describing the physical properties in the image generation process will be added.

A special module for subpixel resolution image analysis according to the geometric models as described above in section 3.2.1 will be developed and integrated in the system.

Remote Sensing Test Data and Explorative Use of Developed Methods

The investigations and tests will be performed using image data of the following types:

- colour infrared aerial photos (3 channels) of different image scale (1:4.000 up to 1:32.000)
- colour infrared satellite photos (2 channels) of type KFA1000 from Russian space station MIR, image scale approx. 1:400.000
- multispectral satellite photos (6 channels) of type MKF6 from Russian space station MIR, image scale approx. 1:3.200.000
- SPOT images, pixel size 20m $\times$ 20m (3 channels) and 10m $\times$ 10m (1 channel)
- LANDSAT TM images, pixel size 30m $\times$ 30m (6 channels).

For most of these images, spectroradiometric reflectance data of test sites on the ground as well as spectroradiometric and other data characterizing the state of the atmosphere at the time of image acquisition are available.

For a start, the scene description will be simplified to a plane thematic map of land use. 10 to 20 categories of land use will be defined and characterized in terms of spectral (and directional ?) reflectance properties. Reference data (from a geographic information system (GIS), e.g. existing thematic map, digital terrain model) will be incorporated in the analysis procedure.

Connection with Fusion

The parameters of the physical model of image generation, which are determined in the model inversion (= image analysis) process, can be considered as part of the image description. Fusion of images of different geometry from different sensors, which are not registered on a pixel-by pixel basis, has to be performed at a higher level of abstraction, i.e. on the image description level. This fusion process can therefore also benefit from the physical model parameters. Examples are:

- the use of atmospheric parameters derived from the analysis of images from one sensor in the model inversion process for images of another sensor

- the transfer of geometric parameters extracted from high-spatial-resolution, but low-spectral-resolution images to the subpixel analysis procedure of low-spatial-resolution, but high-spectral-resolution images, etc.

Using the remote sensing image data material available for this project as described in 3.2.2, various fusion strategies based on this idea of model parameter exchange will be examined.

It seems advanteageous to perform these investigations within a wider scope of research on fusion applied to remote sensing images, in order to be able to assess the potential of the model-based fusion approach in comparison to other methods. The main goals of our further research on this topic (to be pursued in close cooperation with task III.1) are:

- **selection of morphological and spatial features** which can be used to establish correspondences between several sources: Their sensitivity against translation, rotation and scaling will be investigated especially in dependence of the level of abstraction of the features.
- **measures of quality** relevant for remote sensing classifications: Parameters describing *reliability*, *robustness*, and *completeness* have to be defined within this context.
- **time series studies** by fusion of data from different years of acquisition, with the aim to obtain information on land use changes.

3.3 3D Navigation and Reconstruction Based on Multiple Views

Task leader:

Dr. Wolfgang Pölzleitner,
Institute for Digital Image Processing,
Joanneum Research, Graz.

3.3.1 Problem Statement

Navigation and imagery used by autonomously operating devices (robots, vehicles, submarines, industrial automation systems) require methods that are able to derive the elevation map (DEM) or 3D structure of a surface to be inspected, measured, or traversed. One approach to acquire this information would use robust automatic computer vision methods and various types of sensors (lasers, interferometers, cameras). The development of novel stereo vision techniques for this task is strongly required to achieve a consistent surface description. Navigation based on the tracking of artificial and natural landmarks on the recognized surface structure in a further step can lead to fully autonomous behavior as necessary for any scenario that does not allow human interaction.

The task defined in the following sections is geared to develop prototype system consisting of mechanical components, electronic hardware, and software for a demonstrator system that can solve the following 3 generic tasks in a complex three-dimensional scene:

- 3D modeling (reconstruction) of the surfaces in the scene
- Tracking
- Recognition (comparison to a prestored representation of models)
- Navigation

These items are ordered with increasing complexity. For certain industrial applications not all steps will be required. The modules should be generic in that they should be rather independent of the sensors used (gray-level cameras, interferometry, etc.).

DEM Generation

Steps of DEM Generation
The automatic generation of DEMs can be split into several steps to be carried out either in succession or using feedback methods:

1. **Definition of Camera Geometry**
2. **Preprocessing**
3. **Matching**
4. **Error Detection**
5. **DEM derivation from Disparity Map**

[ABG89] describes a set of camera geometries including both the *lateral* and the *axial* model of stereo image acquisition. For the lateral (side by side) geometry, all important relationships between camera position, offsets, search space, and resolution are listed there. The aim of the matching step is to find pixel coordinates in one image that can be located in the other stereo partner describing the same real world location. Section 3.3.1 lists several currently used matching techniques. To avoid errors in the disparity map, methods are available that are either contained in the matching process (correlation threshold, local consistency check, threshold in plausibility image), based upon an intermediate step (back-matching [Han89]), or check the final result after the matching process (edge detection and thresholding on the DEM, [Ols90]). To detect these errors it is necessary to establish assumptions about the scene to be mapped. Several constraints have been defined (e.g. *smoothness assumption, order constraint, rigidity assumption, condition of linear variation, assumption of uniqueness and continuity*). See [Gri81] for further explanation of these principles. Having exact knowledge about geometrical parameters describing the sensing system, the derivation of an elevation model is straightforward. Several photogrammetric methods are known from literature for the lateral and axial camera geometry [ABG89] as well as for the more complex case of remote line scanners [RBM89].

Matching techniques

The different matching techniques can be split into 2 parts each of them following special constraints on the mapped scene:

Grey-level matching

In a recently implemented spatial correlation system [RA91], product moment correlation is used for matching. [Han89] uses normalized cross-correlation [GGB84], in [BNL90] it is used for refinement. [Ehl83] contains a detailed description and comparison of the most important spatial correlation techniques used.

Feature-Based Matching

The major part of the feature-based matching techniques uses edges and/or linked edges (polygons), correspondence of edge orientations [Ols90], zero crossings of Laplacians [Gri85], angle at vertices [GS89], corner points [CB89], regions [LL90] or chain code [GS89], [RA90]. Local phase properties at different spatial resolutions are used in [Jen91] and [JJ89] for disparity measurement. One suggestion for combination of feature and grey-level based matching methods deals with the idea that edge pixels will give a first net of disparity values to be filled and corrected using the disparities derived from grey-level correlation. A suggestion similar to this technique is made in [PM90] and [PHS89].

Multiresolution

Hierarchical methods allowing the access to and information transfer between different image resolutions of one scene are an efficient tool also for matching purposes. A grey-level pyramid method is described in [BNL90] using a hexagonal pyramid, and in [Han89], using a spiral top-down search. The feature pyramid in [Gri85] is built using Laplacians as edge operator, but multiresolution is only simulated using differently sized kernels. A more sophisticated, well described approach is [PHS89], where edges are segmented.

Autonomous Navigation

Position Initialization

In the area of autonomous navigation many approaches have been published using image sequences to get information on the motion parameters. These methods have in common that they relate a world-coordinate system with a camera-coordinate system. This is also called extrinsic calibration. A key to this procedure is that a minimum number of 3D points must be known in world coordinates. This prerequisite can be achieved through external sensors (laser range finders [MM91, Paa92]). To avoid external sensors in time critical situations, however, we propose a method that uses precomputed elevation models and gray-level images to estimate the robot position in a first step. This initial estimate is followed by a refinement using known iterative schemes, e.g., Kalman filter-based methods and tracking.

Landmark Tracking
Purpose for landmark tracking is surface relative navigation, i.e., navigation without knowing the exact absolute position, and hazard avoidance by obstacle detection. The system must have the ability to

- Select new landmarks autonomously since resolution is changing and landmarks disappear from the field of view
- Track the selected landmarks in a robust manner (use Kalman filtering)
- identify the accuracy of the robot position in order to decide if obstacles are possibly hazardous or not.

3.3.2 Research Goals and Methods

The main goal of this project part is the development of an autonomous navigation system based on 3D reconstruction from multiple views and robust hierarchical feature tracking. These two items lead to a partition of our project part into the following subproblems:

1. 3D Elevation Modeling from Multiple Views
2. Landmark Tracking for Navigation
3. Object Recognition

The latter task is strongly dependent on its robust behavior. This leads us to the decision to provide two different approaches for this problem which import a high degree of functional redundancy into the system.

Investigations are planned that lead to the development of an integrated system for the tasks described above. The following considerations gather the expertise of the research scientists involved in the study. They should serve as entry point into the project and consolidate the planned activities:

Hierarchical Feature-Vector Matching (HFVM)

Almost all published matching techniques deal with just one or, at most, two different local properties of an image. A natural extension of this property based matching philosophy is to use a *combination of many of these features* that would lead to a significant improvement of the stereo matching step, especially in terms of robustness. Such a method combines the advantages of several local image features, whereas the particular disadvantages are equalized by the large variety of features. This new approach of stereo matching is based upon the idea of creating a **feature vector for each pixel** (-surrounding) and comparing these features in the images to be registered [PP91a, PP91b, PP92, PPC92]. The principles can be described as follows:

1. Create a set of feature images for both the reference and the search image. These features are derived from local properties in the surrounding of each pixel. The size of the windows depends on the feature to be calculated. Proper features could be local filters, local variance, moments, Fourier and Gabor coefficients, edge properties, or local orientation and frequency. The feature images have the same geometries as their input images (reference and search image, respectively). Hence, the contents of these feature images describes a feature vector for each pixel in both the reference and search image.

2. Compare feature vectors of left and right image to get homologous points using the Euclidean distance as metric.
3. Remove errors and interpolate undefined disparities.

This algorithm has two advantages: First, time consumption on sequential hardware is smaller, as the feature images have to be calculated only once, and the number of features and the amount for vector comparison is usually one order less than the number of pixels for 2-D correlation.

Second, it is not strictly dependent on a pre-registration step. Even stereo pairs that are rotated relative to each other or scaled differently can be used for the matching, if features are used that are invariant to these parameters.

Robust Behavior by Multiresolution

Several approaches in stereo vision deal with multiresolution for the generation of the disparity map. In our case, a 3×3 / 4 Gauss pyramid is built for each of the two images. The number of levels depends on the maximum expected disparity. For each of these levels the above mentioned features are generated. Beginning with the top level, a disparity map is created for this resolution. For each of the next higher resolutions, the disparity derived above is used to begin the search in the right image for the best correspondence on each pixel. Median filtering is carried out on the disparity map to remove salt and pepper noise. To provide backmatching (left-to-right and right-to-left), matching is done in both directions. The result of this process is again a set of undefined pixels on the disparity map to be interpolated.

Hierarchical Feature-Vector Tracking

The tracking problem can be split into at least five different steps:

1. **Position estimation, and calculation**
2. **Selection of interest points** for tracking
3. **Prediction** of 2D interest point coordinates in next frame
4. **Tracking**
5. **Position verification and error detection.**

Interest Operator

The selection of tracking points is greatly dependent upon the matching algorithm used. In case of Feature-Vector Matching (FVM), the idea is to select points in the image whose surrounding shows a large variation in all features of the FVM. This section gives an example of such an interest operator. Local peaks on the multiplied local variances of the feature images are used as interest points:

1. Calculate features as used for feature-vector matching
2. Calculate local 5×5 variance on each feature
3. Multiply the variance images
4. Calculate local 3×3 average over the result
5. Calculate local peaks. They are good tracking candidates since the variation of the used features is large in their surrounding which minimizes ambiguities.

One crucial task is the initialization of the tracking process. Three possibilities can be listed:

1. HFVM, if no initial data about robot direction of motion, speed and distance of the observed surface is known. In this case, only a rough estimate of the maximum possible disparity must be known.

2. Extension of the search space in the second frame. In this case initial values for the above mentioned parameters have to be known with defined uncertainty.

3. If initial values are unknown, but some very distinctive objects exist, the first estimates of these values can be derived relying on these objects only.

The first case is combined with high computational costs compared with the rest of the tracking process. However, HFVM is robust against all sources of distortion (geometrical: rotation, shape change by different viewing angle, different scale; radiometric: Viewing angle dependent albedo difference, noise) and should be used if the second proposed tracking initialization method leads to insufficient results due to ambiguities caused by the search space extension.

Tracking Point Matching

The selection of a matching algorithm within the tracking process has to take into account, that local properties of tracking point surroundings show large variations from one frame to the next, especially at locations far away from the focus of expansion [BB90]. Simple correlation methods fail for that reason. As a result, also for tracking the use of FVM is proposed, both for robustness, and for compatibility with DEM creation.

As first step, the locations of the tracking points have to be estimated, as well as the possible position variations in the 2D images. This can be done using the 3D position of each tracking point and estimated motion parameters, or using the optical flow vectors of the points in the previous frames.

Having estimations for locations of the tracking points in the current frame, the 2D position of each tracking point is calculated using FVM within the respective search space.

3D Position Estimation - Calculation - Verification

Having disparities and initial 3D coordinates of some tracking points together with calibration parameters for the camera system, photogrammetric methods can be used to calculate the position and pointing parameters of the image acquisition platform. With these parameters and the respective parameters from previous frame acquisitions, an estimation (*Kalman Filter*, [MR88]) can be made for the next frame tracking point positions. This data is fed into the matching for position calculation of this frame acquisition (see above).

Together with the tracking, some frames could be used as stereo pairs (as well as a real stereo setup for rover imagery). The calculation of a world coordinate DEM can be done additionally to the tracking, without the real time constraint. With a reasonable delay of several frames, this DEM can be used to verify and correct the position information derived by the tracking steps carried out while the DEM was created. Additionally, the use of auxiliary sensors like laser range finders is used to calculate the position of the image acquisition platform as well as the initial DEM (frame 0).

One result of the position verification is a set of incorrectly matched tracking points that have to be removed from the current set. To provide enough points for further tracking, some new interest points have to be identified at each tracking step. The prediction of their flow vectors can be carried out using several constraints applied to stereo reconstruction methods (rigidity $\longrightarrow$ linear 3D interpolation, or smoothness $\longrightarrow$ 2D Spline interpolation).

Associative Tracking/Matching

Purpose and Assumptions
In this section we describe a concept of a vision system that is used to navigate the robot using a syntactic description of the surface and tracked landmarks. The final goal is to find the position of the robot and guide it on its path.

Navigation is based on vision by identification and tracking of natural objects or parts thereof (landmarks). Such objects are restricted to rocky peaks, ridges, valleys, and areas of constant slopes. Among these landmarks slopes and valleys are the most important candidates for natural terrain.

For the purpose of natural feature tracking we assume that a *global high resolution digital elevation model (DEM)* is available from the orbit phase. It is resampled to the current resolution of the imaging sensor on the robot. In the context of the control of robot relative to the surface it is important to distinguish between the two goals of obtaining an accurate model of the environment, and determining the current position within it. It has been widely agreed [DG88], [PPM90] that this process of dynamic vision can be decomposed into at least two important modules:

1. The automatic generation and update of an highly accurate elevation model (scene description, or map). This necessitates the combination of multiple-position views of the terrain to give complete and robust data); Novelty detection, i.e., objects (or potential hazards) that have not been seen previously (due to occlusion or insufficient resolution).

2. The use of visual tracking to provide navigation signals for the robot. This tracking procedure uses landmarks since it cannot fully rely on external sensors.

In this context the problem is that of dynamic stereo tracking used to compute a path through the 3D environment on the surface based on a subset of objects selected as landmarks. The basic strategy is as follows:

1. After an initialization phase where a full resolution DEM is analyzed, a set of landmarks (primitive objects) is selected.

2. The processing of subsequent frames is performed on the level of the existing objects. Here any available path information from external sensors is used in a feedback loop between landmark tracking and geometrical correspondence of objects to give an updated estimate of their 3D location as well as that of the robot.

Algorithmic Choices
In terms of the algorithms to be implemented we have two distinct choices. The first computes the DEM at each frame and tracks 3D primitive locations, the second only computes the DEM once and then tracks on the 2D gray-level images. A mixture of the two approaches is also possible (as will be obvious below) through the use of a pyramid representation of the DEM as well as the gray-level information.

1. Single frame processing: Here the full DEM generation, recognition and object abstraction process is executed in each frame. For recognition no information from previous frames is used. This mode of operation is feasible when ≈ 1 sec of time delay between frames is acceptable. (Transputer implementation currently going on has shown the functionality on selected types of images, and has also provided performance data). Tracking is based on the high-level part of objects, where the segmentation and recognition parts are repeated at each frame. The advantage of this method is the avoidance of error propagation that may occur in the dynamic tracking procedure described below.

2. Interframe (spatio-temporal) coherence: This concept uses feature extraction as described in 1. only as a starting point. Following an initial recognition in full resolution, tracking is performed on the 2D primitive objects without computing the elevation model. The problem here is to select the best level of abstraction for tracking.

Levels of Abstraction

In this section, we describe a major part of the system which is the segmentation of the *2D images* or the *DEM* into a "background region" and "regions of interest" that may contain various primitives. We present a brief review, details can be found in [Pöl89]. This segmentation is based on a 3-level hierarchy. First, pixel-level operations are used to classify pixels locally into five model classes. This is done by Gradient analysis in terms of thresholding and analysis of sign changes. The output is a symbolic pixel image. Second, this information supports a grouping of pixels into texture elements. These texture elements are found by assigning symbolic classes to a group of 4×4 pixels. The symbolic classes are derived by 1st and 2nd-order statistics on the symbolic and pixel image derived in the first step. The classes assigned to each texture element and the symbols assigned are summarized in textureelements. Third, the symbolic description is used to segment the image syntactically.

Combination Choices

The above mentioned levels of representation as well as the two choices (tracking on 2D images and tracking on DEM) for the algorithms must be combined to get an efficient real-time system. Note that for initial object detection it is not necessary to compute peaks as such, but the pieces thereof as primitives. This basically determines the break point in a hierarchy H1. The eventual determination of the break points can be found from experimentation.

Matching and Tracking

In an elliptical approximation step each primitive region (i.e., area of approximately constant slope) is described by a feature vector. The feature set used include the length of the minor and major axis of the best-fitting ellipse, the orientation of the slope, the size of the region, the mean height of the region, etc. Also shape related features can be computed during connected component labeling, such as second order moments, elongation, and orientation. For matching this representation is used to achieve efficient refinement of an initial position estimate. Matching using this representation is basically a graph isomorphism problem and fault tolerant procedures have been extensively studied. A full set of algorithms to solve the matching and tracking problem has been implemented elsewhere [LEAB91].

A second choice for matching the new elevation model to the reference elevation model (map) would be to use the symbolic representation on the level of texture elements or even pixels. Very fast and efficient matching algorithms exist for this type of representation. The drawback of this choice is the bad performance of matching algorithms in cases where the new DEM and the reference are in different rotation or scaling. In any case, the best level in the gray-level and DEM pyramids have to be selected to get the optimum compromise between accuracy and speed of computation.

Modules

Low-level part This is the primitive object extractor module (which was shown to run in real time). It contains the following steps:

- Low-level processing by pixel-level feature detectors
- Grouping of pixels into texture elements in a feature pyramid
- Fast syntactic segmentation into objects

High-level parts This level deals with the objects (landmark fragments as ridges, valleys, areas of homogeneous slope) and outputs (1) a classification of these primitives, and (2) tracks the primitives.

- Numeric feature vectors on object level
- Symbolic/numeric decision network for classification
- Initial matching, which is relaxation-based when not only a classification is output, but also the corresponding area in the map should be located.
- Dynamic matching to provide tracking. This part consists of both a predictor module (Kalman filter), and a matching module.

A REALISTIC PROCEDURE
A feasible configuration is given in the following procedure:

1. Get DEM (Pyramid-based feature-vector matching)

   The by-product of DEM computation is a pyramid of DEMs
2. Select appropriate level in the pyramid depending on the required resolution. The level should have an optimum detail/noise ratio.
3. (Possibly omitted) Describe the DEM by symbolic texture elements. This will provide fast tracking and recognition rates.
4. Combine symbolic texture elements to objects (ridges, valleys, slopes) by connected component analysis
5. Use feature vectors to describe and classify objects.
6. Do the same with the prefabricated DEM map model.
7. Use the high-level representation of the DEM and map to implement fast matching and correspondence finding.
8. When the correspondence between the map and the new DEM is found the position estimate of the robot can be computed. This estimate is then input into a prediction-update cycle based on Kalman filtering techniques.

The first 7 steps of the above procedure can be viewed as a bootstrap mechanism that supplies the initial estimates of landmarks and positions. The following step of dynamic tracking was well studied in [LEAB91]. The tracking system as implemented in this study can fully be used as the module to follow steps 1 – 7.

Recognition using Associative Memories
The indexing problem, i.e., memory access by contents rather than by address, plays a central role in pattern recognition systems. Specifically, a signature of a pattern is presented to the memory in order to output the classification of the pattern. A suitable model to solve the indexing problem is provided by the Distributed Associative Memory (DAM) [Koh88]. This model, conceptually close to classical conditioning [Heb49], builds up the memory by associating pairs of key stimulus and response vectors via a matrix projection operator. Memory retrieval is performed by multiplying (projecting) the unknown key vector with (on) the memory matrix. The resulting output vector (the *recall* or *recollection*) is either a particular stored data vector or an approximation of it. The recognition device belongs to the class of linear associative memories extensively studied and compared ([Che], [CT90], and [TC90]). A

detailed description on how the DAM can be used for pattern recognition can be found in [PW90]. In the application described here, the problem is to locate a specific reference object in the test DEM acquired by the robot. The specific properties leading to the application of the particular recognition system are its invariance to scale and rotation.

The complete procedure of recognizing objects in a DEM is as follows:

Recognition of Objects in a DEM

Input: A reference DEM containing objects of interest and a test DEM in which the objects should be recognized. Both DEMs are segmented by symbolic preprocessing as described above.

Output: For each object in the test DEM the recognition histogram is provided by the associative memory.

Method:

Step 1. In both the reference and test DEMs the largest k objects are selected. This limits the fragmentation problem as described below.

Step 2. The objects in the reference DEM are used to set up the training base in the associative memory.

Step 3. The test DEM is now tested with the memory to output the recognition histogram.

In Step 2 of the above algorithm the training stimulus vectors are stored in the memory. These vectors are the numeric height values at each location in the object. A later extension would include the use of the gray-level information of one (or both) stereo partners, and low-level and medium-level featural representation.

3.4 Visuo-Motor Coordination

Task leader:

Dr. Cornelia Fermüller,
Center for Automation Research,
University of Maryland, USA.

3.4.1 Problem Statement

Visual Navigation / Manipulation constitutes a problem of considerable practical and scientific interest, and in general it refers to the performance of sensory mediated movement. Visual Navigation / Manipulation is defined as the process of motion control based on an analysis of images. A system with navigational capabilities interacts adaptively with its environment. The movement of the system is governed by sensory feedback which allows it to adapt to variations in the environment and does not have to be limited to a small set of predefined motions as it is the case, for instance, with cam-activated machinery.

In the past research on navigation has mainly been based on the reconstruction paradigm described by Marr [Mar82]. All navigational tasks have been considered as applications of the general *structure from motion problem*. The idea was to recover the relative 3D-motion and the structure of the scene in view from a given sequence of images taken by an observer in motion relative to its environment. The way the problem has been addressed, was to solve it in two computational stages. First accurate image displacements between consecutive frames have been computed, either in form of point correspondences [Ull79, Fau92] or as dense motion fields (optical flow fields) [BT80, HS81, Hil93, Nag83, AW85]. Then, in a second step the 3D-motion and the structure has been computed from the equations relating it to the 2-D image velocity [BH83, Adi85a, Hor90, LH81, LHP80, WKS87, SA88, TH84]. This computational theory has been uncritically accepted and as a result most studies on visual motion perception are to be found at the algorithmic level of the problem; addressing either the estimation of image motion or the recovery of 3-D motion and structure.

Of course, if structure and motion can be computed, then various subsets of the computed parameters provide sufficient information to solve many practical navigational tasks. However, in general the estimation of optical flow (or correspondence) is an ill-posed problem and additional assumptions must be made in order to estimate it. If these assumptions hold, as in the case of some model-based approaches [BBHK89, KDTN92], optical flow can be computed accurately and in the sequel correct 3-D motion can be derived. In the general case, however, any algorithm will produce imperfect output (erroneous output, if the assumptions do not hold). Furthermore, recovering 3-D motion from noisy flow fields has turned out to be a problem of extreme sensitivity with researchers reporting very large amount of errors in the motion parameter estimation under small perturbations in the input. Even optimal algorithms [SA88, DN90] perform quite poorly in the presence of moderate noise. Although a formal proof is still lacking, it has been argued [Adi85b], that the estimation of 3-D motion from image motion is itself ill-posed, because it does not continuously depend on the input. Therefore, although a great deal of effort has been spent on the subject, the problem of structure from motion still remains unsolved for all practical purposes.

The main reason why the approach to vision suggested by Marr has not led to the development of successful artificial systems is that vision was studied in a vacuum, i.e. its utilization was completely ignored. However, in no system is vision purposeless; thus it needs to be studied in conjunction with the task the system is involved in. From this viewpoint, understanding vision means understanding a system possessing visual capabilities.

In general, if our goal is to study (or more precisely formulated, analyze in order to design) a system, we are advised by engineering considerations to follow some common principles in system design and address a set of basic questions: What is the functionality of the system? What are the autonomous subsystems (modules) the system is divided into? What is the relationship of the modules to each other?

What is the representation of information within the subsystems, and how do the modules communicate with each other? Finally, we have to ask: what is the most efficient and effective way to design the individual modules? Having these questions in mind a new approach for studying visual systems is proposed, and this approach is applied to study the interpretation of image sequences in the context of visual navigation and manipulation.

3.4.2 Research Goals and Methods

The approach for studying navigational systems proposed here, takes it for granted that the observer (the system) possesses an active visual apparatus. Since, furthermore, it is inspired by evolutionary, neuroethological considerations, we call it the synthetic (evolutionary) approach [Fer93c]. It is basically substantiated by two principles. The first principle is related to the overall structure of the system and how it is modularized, what are the problems to be solved and in which order they ought to be addressed. The second principle is concerned with the way the individual modules should be realized.

As a basis for its computations a system has to utilize mathematical models, which serve as abstractions of the representations employed. The first principle of the synthetic approach states that the study of visual systems should be performed in a hierarchical order according to the complexity of the mathematical models involved. Naturally, the computations and models are related to the class of tasks the system is supposed to perform. A system possesses a set of capabilities which allow it to solve certain tasks. The synthetic approach calls for first studying capabilities whose development relies on only simple models and then going on to study capabilities requiring more complex models. Simple models do not refer to environment- or situation-specific models which are of use in only limited numbers of situations. On the contrary, each of the capabilities requiring a specified set of models can be used for solving a well-defined class of tasks in every environment and situation the system is exposed to. In other words, the assumptions used have to be general with regard to the environment. The motivation for this approach is to increasingly gain insight into the process of vision, which is of such high complexity. Therefore the capabilities which require more complex models should be based on "simpler", already developed capabilities. The complexity of a capability is thus given by the complexity of its assumptions; what has been considered a simple capability might require complex models, and vice versa. For example, as shown in [FA93b, FA92, Fer93a], the celebrated capabilities of egomotion estimation and estimation of an object's motion do not require, as has been believed, complex models about the geometry of the scene in view or the time evolution of the motion, but only a simple rigid motion model. Thus, these two capabilities form the bottom of the hierarchy of visual navigational tasks.

Next in the hierarchy follow the capabilities of independent motion detection and obstacle avoidance. Although the detection of independent motion seems to be a very primitive task, it can easily be shown by a counterexample that in the general case it cannot be solved without any knowledge of the system's own motion. Imagine a moving system that takes an image showing two areas of different rigid motion. From this image alone, it is not decidable which area corresponds to the static environment and which to an independently moving object. A motion model more complex than the rigid one has to be employed. In order to perform obstacle avoidance it is necessary to have some representation of space. This representation must capture in some form the change of distance between the observer and the scene points which have the potential of lying in the observer's path. An observer that wants to avoid obstacles must be able to change its motion in a controlled way and must therefore be able to determine its own motion and set it to known values. As can be seen, the capability of egomotion estimation is a prerequisite for developing general independent motion detection as well as obstacle avoidance mechanisms.

Even higher in the hierarchy are the capabilities of target pursuit and homing (the ability of a system to find a particular location in its environment) and the capability of hand-eye coordination. Obviously, a system that possesses these capabilities must be able to compute its egomotion and must be able to avoid obstacles and detect independent motion. Furthermore, homing requires knowledge of the space and models of the environment, whereas target pursuit relies on models for representing the operational space and the motion of the target. In order to solve hand-eye coordination models must be available to relate the changes in the 3-D scene to the motions of the robot's joints.

It is proposed to study the above described five capabilities of independent motion detection, estimation of time to collision, target pursuit, homing, and hand-eye coordination following the principles of the synthetic approach. The goal is to study the capabilities in an order defined by the complexity of the visual capabilities, such that more complex capabilities require the existence of simpler ones.

The way the capabilities are to be realized is according to the second principle of the synthetic approach, which is motivated by the need for robustness. It is this the quest for algorithms which are qualitative in nature. The synthetic approach does not have as its goal the reconstruction of the scene in view, but the development of a class of capabilities that recognize aspects of objective reality which are necessary to perform a set of tasks. The function of every module in the system constitutes an act of recognizing specific situations by means of primitives which are applicable in general environments. For example, a system, in order to avoid obstacles, does not have to reconstruct the depth of the scene in view. It merely has to recognize that the distance to a close-by object is decreasing at a rate beyond some threshold given by the system's reaction time. Recognition, of course, is much easier than general reconstruction of the scene, simply because the information necessary to perform a specific task can be represented in a space having only a few degrees of freedom [FA93a]. Moreover, in order to speak of an algorithm as qualitative, the primitives to be computed do not have to rely on explicit unstable, quantitative models. Qualitativeness can be achieved for a number of reasons: The primitives might be expressible in qualitative terms, or their computation might be derived from inexact measurements and pattern recognition techniques, or the computational model itself might be proved stable and robust in all possible cases.

Let us now explain the basic technical characteristics of our approach to solving the above described capabilities of visual navigation and manipulation. Current motion understanding techniques require the computation of exact image motion (optical flow in the differential case or correspondence of features in the discrete case). This computation, however, amounts to an ill-posed problem. As a result, existing motion analysis techniques cannot be used in general environments and thus cannot serve as the basis for navigation and manipulation tasks. The proposed approach to solving these tasks differs from existing techniques in the following fundamental issues: As input for the description of image motion only the spatio-temporal derivatives of the image intensity function, the so-called "normal flow" will be used. In existing approaches to navigation and manipulation in a first step a complete and accurate geometric description of the scene is attempted to be extracted, which is then used in the planning phase. We propose to perform motor control in a more efficient way by extracting only relevant partial information of the scene and study the relation between image motion patterns and the change of objects in the 3D world. It is proposed to provide a scientific platform for the development of visual capabilities for navigation and manipulation which provides a categorization of visual capabilities by means of the computational models involved. The idea is to relate global and local properties of normal flow fields to visual competences and study the relationship between relevant subsets of the input data and an observer's actions. It is planned to study how actions facilitate the computations by providing additional mathematical constraints. In particular, our research goals in navigation and manipulation are as follows:

Detection of Independent Motion

by a moving observer from a sequence of normal flow fields by studying discontinuities and nonrigidities:

A independently moving object in the view of a moving observer creates a motion field which is globally non-rigid. As has been described before, it can easily be shown by a counterexample that in the general case this task cannot be solved without deriving some information about the observer's egomotion. As has been shown in [Fer93b, Fer93d], the egomotion of an observer moving in a static scene can be estimated by locating in the image plane global patterns formed by the spatio-temporal derivatives of the image intensity function. If an object is moving independently these global patterns of image motion due to rigidity are destroyed. Independently moving objects give rise to discontinuities in the flow field at their boundaries. These, however, are not the only places where the flow field is not smooth; surfaces that are separated in depth also cause discontinuities. Research will be devoted to differentiating / reasoning between these two types of discontinuities. Also, successive fixations could supply information regarding independent motion by extracting quantities remaining invariant for features belonging to the static scene.

Estimation of Time-to-collision Hazard Maps

where maps at consecutive time instances are connected by a regularizing operator:

In the past work on finding shape / time-to-collision relied on local values of the image motion. Since, however, local image measurements are very hard to be computed exactly, such methods have been very unreliable. It is proposed to relate the structure of local motion fields, i.e, deformations of flow patterns, and the temporal evolution of motion fields to the underlying local shape models. Also, it is proposed to study how particular activities facilitate the extraction of the pattern deformations and how they can make the underlying computations robust.

Target Pursuit

through a synergistic application of tracking, 3D-motion parameter estimation, time-to-collision, and spatial planning:

The problem of target pursuit or prey catching is an ideal platform for studying the integration of different cognitive / perceptual capabilities in a working system. In particular, a system with prey catching capabilities must possess the more primitive competences of 3-D motion parameter estimation, tracking, estimation of time-to-contact and spatial planning. An important aspect of integrating these different competences is related to issues of control. It is proposed to study an integration of the aforementioned competences using the framework of Discrete event dynamic systems.

Homing / Docking

through learning of environmental structure and the creation of an associative mapping to establish connections between different views and motion commands:

In order to accomplish homing (i.e. the capability of a system to move from any position in the environment to a particular location (home) using visual means) the system must build a representation of the environment. The characteristics of this representation depend on the physiology and computational capacity of the system. Recent results from Psychology demonstrate that even for humans such representations are of a qualitative nature not employing detailed, elaborate models, but using direct relationships between visual features, actions and environments. Here it is proposed to study qualitative representations of space and study their relation to the amount of available memory.

Hand-eye Coordination

by associating image feature trajectories with motions of joints in a forward manner:

In the past hand-eye coordination usually was studied by creating a detailed description of the environment and then solving the inverse kinematic equations. Such an approach, however, is computationally heavy and instable. Here it is proposed to associate image features directly to the motion or forces of the robots joints. Such image features include local as well as global motion fields, distortions in the motion fields, the change of motion fields over time, and local shape features as they relate to activities such as grasping. Thus, studying hand-eye coordination in the suggested way, requires exploring relationships between an observer's physiology and spatio-temporal image features due to changes in the observer's environment.

Bibliography

[ABG89] N. Alvertos, D. Brzakovic, and R. Gonzalez. Camera Geometries for Image Matching in 3-D Image Vision. *IEEE Trans. Patt. Anal. Mach. Intell.*, 11(9):897–915, September 1989.

[Adi85a] G. Adiv. Determining 3d motion and structure from optical flow generated by. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 7:384–401, 1985.

[Adi85b] G. Adiv. Inherent ambiguities in recovering 3d motion and structure from a noisy flow field. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 70–77, 1985.

[Alo93] John Aloimonos, editor. *Active Perception.* Advances in Computer Vision. Lawrence Erlbaum, 1993.

[AS89] John Aloimonos and David Shulman. *Integration of Visual Modules. An Extension of the Marr Paradigm.* Academic Press, 1989.

[Asr89] G. Asrar, editor. *Theory and Application of Optical Remote Sensing.* John Wiley & Sons, 1989.

[AW85] P. Anandan and R. Weiss. Introducing a smoothness constraint in a matching approach for the computation of optical flow fields. In *Proc. 3rd Workshop on Computer Vision: Representation and Control*, pages 186–194, 1985.

[AWB87] John Aloimonos, Isaac Weiss, and Amit Bandyopadhyay. Active vision. In *Proc. 1st ICCV*, pages 35–54. IEEE Comp.Soc.Press, 1987.

[Bar89] Renate Bartl. Entwurf einer objektorientierten Wissensrepräsentation und einer Ablaufsteuerung für ein universelles bildverstehendes System. Master's thesis, Technical University of Vienna, 1989.

[BB90] Burger, W. and Bhanu, B. Estimating 3-D Egomotion from Perspective Image Sequences. *IEEE Trans. Patt. Anal. Mach. Intell.*, 12(11):1040–1058, November 1990.

[BBHK89] P. Burt, J. Bergen, R. Hingorani, and R. Kolczynski. Object tracking with a moving camera. In *Proc. IEEE Workshop on Visual Motion*, pages 2–12, 1989.

[BH83] A. Bruss and B. Horn. Passive navigation. *Computer Vision, Graphics, and Image Processing*, 21:3–20, 1983.

[BN86] Arthur Browne and Leonard Norton-Wayne. *Vision and Information Processing for Automation.* Plenum Press, 1986.

[BNL90] Brookshire, G., Nadler, M., and Lee, C. Automated Stereophotogrammetry. *Computer Vision, Graphics, and Image Processing*, 52:276–296, 1990.

[BP90] Renate Bartl and Axel Pinz. Allgemeines Konzept einer Ablaufsteuerung für ein bildverstehendes System. In Gustav Bernroider and Axel Pinz, editors, *Image Acquisition and Real-Time Visualization*, volume 56 of *OCG-Schriftenreihe*, pages 51–76. Oldenbourg, 1990.

[BP92a] Renate Bartl and Axel Pinz. An extension of the data fusion concept. In Peter Mandl, editor, *Modelling and New Methods in Image Processing and Geographic Information Systems*, volume 61 of *OCG-Schriftenreihe*, pages 111–122. Oldenbourg, 1992.

[BP92b] Renate Bartl and Axel Pinz. Information fusion in remote sensing: Land use classification. In *Multisource Data Integration in Remote Sensing for Land Inventory Applications, Proc Int.IAPR TC7 Workshop*, pages 9–17, 1992.

[BPS93] R. Bartl, A. Pinz, and W. Schneider. A framework for information fusion and an application to remote sensing. In S.J. Pöppl and H. Handels, editors, *Proc. 15th DAGM Symposium, Lübeck, Germany*, Informatik aktuell, pages 313–320. Springer, 1993.

[BS93] R. Bartl and W. Schneider. Information fusion at the image description level: Fusion of regions from different images. In Wolfgang Pölzleitner and Emanuel Wenger, editors, *Image Analysis and Synthesis*, volume 68 of *OCG-Schriftenreihe*, pages 147–155, 1993.

[BSP92] Horst Bischof, Werner Schneider, and Axel Pinz. Multispectral classification of Landsat-images using neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, 30(3):482–490, 1992.

[BSP93] Renate Bartl, Werner Schneider, and Axel Pinz. Information fusion in remote sensing: Combining austromir data and landsat data. In *Proceedings of the 25th ERIM Symposium, Graz*, volume II, pages 163–172, 1993.

[BT80] S. Barnard and W. Thompson. Disparity analysis of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2:333–340, 1980.

[Bur88] P.J. Burt. Attention mechanisms for vision in a dynamic world. In *9th ICPR*, volume II, pages 977–987. IEEE Comp.Soc.Press, 1988.

[CB89] Charnley, D. and Blissett, R. Surface reconstruction from outdoor image sequences. *Image and Vision Computing*, 7(1):10–16, February 1989.

[CGH92] V. Clément, G. Giraudon, and S. Houzelle. A multi-specialist architecture for sensor fusion in remote sensing. In *Proceedings of the 11th ICPR*, volume I, pages 202–206, 1992.

[Che] V. Cherkassky. Linear Algebra Approach to Neural Associative Memories. To appear.

[CT90] D. Casasent and B. Telfer. Ho-Kashyap CAAP 1:1 Associative Processors. In *Proceedings Intelligent Robots and Computer Vision IX: Neural, Biological, and 3-D Methods*, volume 1382, Boston, November 1990. SPIE.

[CY90] James J. Clark and Alan L. Yuille. *Data Fusion for Sensory Information Processing Systems*. Kluwer Academic Publishers, 1990.

[DG88] E. D. Dickmanns and V. Graefe. Dynamic Monocular Machine Vision. *Image and Vision Computing*, 1(4):223–240, 1988.

[DHR93] Bruce A. Draper, Allen R. Hanson, and Edward M. Riseman. Learning blackboard-based scheduling algorithms for computer vision. *International Journal on Pattern Recognition and Artificial Intelligence*, 7(2):309–328, 1993.

[DN90] K. Daniilidis and H. Nagel. Analytical results on error sensitivity of motion estimation from two views. *Image and Vision Computing*, 8:297–303, 1990.

[DP92] Peter Datlinger and Axel Pinz. Scanning laser ophthalmoscope and digital image analysis: Registration, blobs and curves. In *IXth Congress of Societas Ophthalmologica Europaea*, page 156, 1992.

[DPP+92] Peter Datlinger, Axel Pinz, Hubert Plank, Susanne Binder, Michaela Velikay, Ulrike Stolba, and A. Wedrich. Digitale Bildanalyse zur Darstellung des Fundus bei seniler Makuladegeneration. *Spektrum Augenheilkd*, 6(1):13–19, 1992.

[Ehl83] Ehlers, M. *Untersuchung von digitalen Korrelationsverfahren zur Entzerrung von Fernerkundungsaufnahmen.* PhD thesis, University of Hannover, Hannover, 1983. pp.123-135.

[FA92] C. Fermüller and Y. Aloimonos. Tracking facilitates 3-d motion estimation. *Biological Cybernetics*, 67:259–268, 1992.

[FA93a] C. Fermüller and Y. Aloimonos. Recognizing 3-d motion. In *Proc. International Joint Conference on Artificial Intelligence*, 1993.

[FA93b] C. Fermüller and Y. Aloimonos. The role of fixation in visual motion analysis. *International Journal of Computer Vision: Special Issue on Active Vision*, 1993.

[Fau92] O. Faugeras. *Three Dimensional Computer Vision.* MIT Press, Cambridge, MA, 1992.

[Fer93a] C. Fermüller. Global 3-d motion estimation. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 1993.

[Fer93b] C. Fermüller. Motion constraint patterns. In *Proc. IEEE Workshop on Qualitative Vision*, 1993.

[Fer93c] C. Fermüller. Navigational preliminaries. In Y. Aloimonos, editor, *Active Vision, Advances in Computer Vision.* Lawrence Erlbaum, Hillsdale, NJ, 1993.

[Fer93d] C. Fermüller. Qualitative interpretation of image derivatives. *International Journal of Computer Vision: Special issue on Qualitative Vision*, 1993.

[FJLS92] H. Füger, K. Jurkiewicz, K. Lütjen, and U. Stilla. Ein wissensbasiertes System für die automatisch Bildanalyse. In *Proc. 17th ISPRS 92, Washington, volume XXIX, part B3, commission III*, pages 167–172, 1992.

[GGB84] A. Goshtasby, S.H. Gage, and J.F. Bartholic. A Two-Stage Cross Correlation Approach to Template Matching. *IEEE Trans. Patt. Anal. Mach. Intell.*, 6(3):374–378, May 1984.

[Gri81] Grimson, W.E.L. *From Images to Surfaces.* MIT Press, Massachusetts, 1981.

[Gri85] Grimson, W.E.L. Computational Experiments with a Feature Based Stereo Algorithm. *IEEE Trans. Patt. Anal. Mach. Intell.*, 7(1):17–34, January 1985.

[GS89] Greenfeld, J.S. and Schenk, A.F. Experiments with Edge-Based Stereo Matching. *Photogrammetric Engineering & Remote Sensing*, 55(12):1771–1777, December 1989.

[Hag90] Gregory D. Hager. *Task-Directed Sensor Fusion and Planning.* Kluwer Academic publishers, 1990.

[Han89] Hannah, M.J. A System for Digital Stereo Image Matching. *Photogrammetric Engineering & Remote Sensing*, 55(12):1765–1770, December 1989.

[Heb49] D. O. Hebb. *The Organization of Behavior.* Wiley, 1949.

[Hil93] E. Hildreth. *The Measurement of Visual Motion.* MIT Press, Cambridge, Massachusets, 1993.

[Hor90] B. Horn. Relative orientation. *International Journal of Computer Vision*, 4:59–78, 1990.

[HPC92] Report on workshop on high performance computing and communications for grand challenge applications: Computer vision, speech and natural language processing, and artificial intelligence. Report of NSF Workshop, November 1992.

[HS81] B. Horn and B. Schunck. Determining optical flow. *Artificial Intelligence*, 17:185–203, 1981.

[Jen91] Jenkin, M.R.M. Techniques for Disparity Measurement. *Computer Vision, Graphics, and Image Processing*, 53(1):14–30, January 1991.

[JJ89] Jepson, A.D. and Jenkin, M.R.M. The Fast Computation of Disparity from Phase Differences. In *Proc. Conf. Comp. Vis. Patt. Rec.*, pages 398–403, San Diego, CA, June 1989. IEEE.

[Kan87] Takeo Kanade, editor. *Three-Dimensional Machine Vision.* Kluwer Academic Publishers, 1987.

[KDTN92] D. Koller, K. Daniilidis, T. Thorhallson, and H. Nagel. Model-based object tracking in traffic scenes. In *Proc. Second European Conference on Computer Vision*, pages 437–452, 1992.

[KKS+92] R. Kalliany, K. Kraus, W. Schneider, O. Lebedev, and M. Seger. FEM: Fernerkundung auf 'MIR'. In *AUSTROMIR 91, Endbericht - Ergebnisse des sowjetisch-österreichischen Raumfluges auf der Orbitalstation 'MIR' (Projekt 'AUSTROMIR')*, pages 290–306, 1992.

[Koh88] T. Kohonen. *Self-Organization and Associative Memory.* Springer-Verlag, 2nd edition, 1988.

[Kra90] K. Kraus. *Fernerkundung, Auswertung photographischer und digitaler Bilder*, volume 2. Dümmler, 1990.

[KS88] K. Kraus and W. Schneider. *Fernerkundung, Physikalische Grundlagen und Aufnahmetechniken*, volume 1. Dümmler, 1988.

[KSEL92] R. Kalliany, W. Schneider, K. Eerme, and O. Lebedev. Remote sensing experiment 'FEM' - A multisensoral dataset for radiometric and geometric analysis. In *Proc. of Satellite Symposium 4: Columbus Eight (COSY-8) Utilisation of Earth Orbiting Laboratories, from the 'International Space Year' Conference, Munich*, pages 133–138, 1992.

[KW89] K. Kraus and P. Waldhäusl. *Photogrammetie, Band 1.* Dümmler Verlag, 1989.

[LE89] Claus-E. Liedtke and Manfred Ender. *Wissensbasierte Bildverarbeitung.* Springer Verlag, 1989.

[LEAB91] Lodola, P., Eklundh,J.O., Andersson, M., and Berthod, M. Stereovision and Motion for Advanced Space Robotics. In *Proc. of First ESA Workshop on Computer Vision and Image Processing for Spaceborne Applications*, ESTEC, Noordwijk, June,10-12. 1991.

[LH81] H. Longuet-Higgins. A computer algorithm for reconstruction of a scene from two projections. *Nature*, 293:133–135, 1981.

[LHP80] H. C. Longuet-Higgins and K. Prazdny. The interpretation of a moving retinal image. *Proceedings of the Royal Society, London B*, 208:385–397, 1980.

[LL90] Lee, H-J. and Lei, W-L. Region Matching and Depth Finding for 3D Objects in Stereo Aerial Photographs. *Pattern Recognition*, 23(1/2):81–94, 1990.

[Man92] Reinfried Mansberger. *Ein System zur visuell-digitalen Zustandsbeurteilung von Baumkronen auf Farb-Infrarot-Luftbildern.* PhD thesis, Universität für Bodenkultur, 1992.

[Mar82] D. Marr. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information.* W.H. Freeman and Company, New York, 1982.

[Mat88] Takashi Matsuyama. Expert systems for image processing – knowledge based composition of image analysis processes. In *Proceedings of the 9th ICPR*, pages 125–133. IEEE Comp.Soc., 1988.

[MH90] T. Matsuyama and V. Hwang. *SIGMA a Knowledge-based Aerial Image Understanding System.* Plenum, 1990.

[Min75] Marvin Minsky. A framework for representing knowledge. In Patrick H. Winston, editor, *The Psychology of Computer Vision*, chapter 8, pages 211–277. McGraw Hill, 1975.

[MM91] Moring, I. and Myllylä, R. On some 3-d Vision Systems Based on Pulsed Time-of-Flight Laser Radar Technology. In *Proc. of First ESA Workshop on Computer Vision and Image Processing for Spaceborne Applications*, ESTEC, Noordwijk, June,10-12. 1991.

[MR88] Mermiod, B. and Rizos,C. A Surveyor's Perspective of Kalman Filtering. *Aust.J.Geod.Photogram.Surv.*, 49:1–38, December 1988.

[MS92] N. J. Mulder and K. Schutte. Knowledge engineering in R.S. and knowledge based systems in GIS. In *Proc. 17th ISPRS 92, Washington, vol. XXIX, part B3, commission III*, pages 965–968, 1992.

[Nag83] H. Nagel. Displacement vectors derived from second order intensity variations in image sequences. *Computer Vision, Graphics and Image Processing*, 21:85–117, 1983.

[Nas91] J.E. Nasemann. Scanning-Laser-Ophthalmoskopie. Prinzip und klinische Anwendung. *Augenärztliche Fortbildung*, 14:14–19, 1991.

[Nie91] H. Niemann. Knowledge based image processing. In *Proc. 43th Photogrammetric Week, Stuttgart University*, pages 215–225, 1991.

[Oht85] Yuichi Ohta. *Knowledge-Based Interpretation of Outdoor Natural Color Scenes.* Pitman, 1985.

[Ols90] S.I. Olsen. Stereo Correspondence by Surface Reconstruction. *IEEE Trans. Patt. Anal. Mach. Intell.*, 12(3):309–315, March 1990.

[Paa92] G. Paar. Technical Assistance for Comparative Digital Elevation Map (DEM) Measurements with Active and Passive Optical Sensors. Report to ESTEC purchase order 122961, Joanneum Research, Graz, Austria, October 1992.

[PB92a] Axel Pinz and Renate Bartl. Information fusion in image understanding. In *Proceedings of the 11.ICPR*, volume I, pages 366–370. IEEE Computer Society, 1992.

[PB92b] Axel Pinz and Renate Bartl. Information fusion in image understanding: Landsat classification and ocular fundus images. In *SPIE Sensor Fusion V, Boston 92*, volume 1828, pages 276–287. SPIE, 1992.

[PD92a] Axel Pinz and Peter Datlinger. Digital image analysis and scanning laser ophthalmoscope I: Superimposition of essential retinal features. In Horst Bischof and Walter G. Kropatsch, editors, *Pattern Recognition 1992*, volume 62 of *OCG-Schriftenreihe*, pages 45–55. Oldenbourg, 1992.

[PD92b] Axel Pinz and Peter Datlinger. Digital image analysis and scanning laser ophthalmoscope: Registration, blobs, curves and pyramids. In *Digital Image Processing in Medicine, Remote Sensing and Visualization of Information*, pages 105–117, 1992.

[PH92] Keith Price and Andres Huertas. Using perceptual grouping to detect objects in aerial scenes. In *Procx 17th ISPRS 92, Washington, volume XXIX, part B3, commission III*, pages 842–855, 1992.

[PHS89] Pong, T-C., Haralick, R.M., and Shapiro, L.G. Matching topographic structures in stereo vision. *Pattern Recognition Letters*, 9:127–136, February 1989.

[Pin86a] Axel Pinz. Architektur und Anwendung des bildverstehenden Expertensystems VES. In C.R. Rollinger and Werner Horn, editors, *GWAI-86 und 2. Österreichische AI-Tagung*, volume 124 of *IFB*, pages 212–217. Springer, 1986.

[Pin86b] Axel Pinz. VES, ein bildverstehendes Expertensystem - Einsatz in der Waldschadenserhebung. In Walter Kropatsch and Peter Mandl, editors, *Mustererkennung 86*, volume 36 of *OCG-Schriftenreihe*, pages 74–91. Oldenbourg, 1986. ÖAGM-Preis 1985.

[Pin88] Axel Pinz. *Ein bildverstehendes Expertensystem zur Erkennung von Bäumen auf Farb-Infrarot-Luftbildern.* PhD thesis, Technische Universität Wien, 1988.

[Pin89] Axel Pinz. Final Results of the Vision Expert System VES: Finding Trees in Aerial Photographs. In Axel Pinz, editor, *Wissensbasierte Mustererkennung*, volume 49 of *OCG-Schriftenreihe*, pages 90–111. Oldenbourg, 1989.

[Pin91] Axel Pinz. A Computer Vision System for the Recognition of Trees in Aerial Photographs. In James Tilton, editor, *Multisource Data Integration in Remote Sensing*, volume 3099 of *NASA Conference Publication*, pages 111–124. NASA, 1991.

[PM90] Perlant, F.P. and McKeown, D.M. Scene Registration in Aerial Image Analysis. *Photogrammetric Engineering & Remote Sensing*, 56(4):481–493, April 1990.

[Pöl89] W. Pölzleitner. *Algorithms for the Classification of Profiled Boards by Real-Time Computer Vision.* PhD thesis, Technische Universität Graz, 1989.

[PP91a] Paar, G. and Pölzleitner, W. Automatic 3D Elevation Modeling by Pyramid-Based Stereovision Using Remotely Sensed Data. In P. Mandl, editor, *Modelling and New Methods in Image Processing and in Geographical Information Systems*, pages 95–110, Vienna, Munich, 1991. Oesterreichische Computer Gesellschaft, R. Oldenbourg.

[PP91b] Paar, G. and Pölzleitner, W. Stereovision and 3D Terrain Modeling for Planetary Exploration. In *Proc. 1st ESA Workshop on Comp. Vision and Image Processing for Spaceborne Applications*, Noordwijk, The Netherlands, June 1991. European Space Research and Technology Centre.

[PP92] Paar,G. and Pölzleitner,W. Robust Disparity Estimation in Terrain Modeling for Spacecraft Navigation. In *Proc. 11th ICPR.* International Association for Pattern Recognition, 1992.

[PPC92] Paar, G., Pölzleitner,W., and Classen, H.J. Planetary Body High Resolution 3D Model of Terrain. Final report, European Space Research and Technology Centre, Noordwijk, The Netherlands, October 1992.

[PPM90] S. B. Pollard, J. Porrill, and J. E. W. Mayhew. Experiments in Vehicle Control Using Predictive Feed-Forward Stereo. *Image and Vision Computing*, 8(1):63–70, February 1990.

[PW90] W. Pölzleitner and H. Wechsler. Selective and Focused Invariant Recognition Using Distributed Associative Memories (DAM). *IEEE Trans. Patt. Anal. Mach. Intell.*, 11(8):809–814, August 1990.

[RA90] Rodríguez, J.J. and Aggarwal, J.K. Matching Aerial Images to 3-D Terrain Maps. *IEEE Trans. Patt. Anal. Mach. Intell.*, 12(12):1138–1149, December 1990.

[RA91] Raggam, J. and Almer, A. A Multi-Sensor Stereo Mapping Experiment. In *Proc. ACSM/ASPRS/Auto Carto Annual Convention*, pages 173–182, Baltimore, Maryland, March, 25–29 1991.

[RBM89] Raggam, J., Buchroithner,M.F., and Mansberger, A. Relief mapping using nonphotographic spaceborne imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 44(1):21–36, September 1989.

[Ree90] W. G. Rees. *Physical Principles of Remote Sensing.* Topics in Remote Sensing. Cambirdge University Press, 1990.

[RJ88] A. R. Rao and R. Jain. Knowledge representation and control in computer vision. *IEEE Expert*, pages 64–79, 1988.

[SA88] M. Spetsakis and J. Aloimonos. Optimal computing of structure from motion using point correspondence. In *Proc. International Conference on Computer Vision*, pages 449–453, 1988.

[Sch81] Werner Schneider. Digitale Auswertung von Farb-Infrarot-Luftbildern. In *Proc. AGM-Tagung 'Mustererkennung in Österreich', Ramsau*, pages 203–218, 1981.

[Sch86a] Werner Schneider. Photographic quality of CIR aerial photos as a function of atmospheric parameters. In *Progress in Imaging Sensors, Proc. ISPRS Comm.I Symposium, Stuttgart, ESA SP252*, pages 505–509, 1986.

[Sch86b] Werner Schneider. Rechnerunterstützte Interpretation von Farb-Infrarot-Luftbildern zur Erkennung und Kartierung von Waldschädigungen. In *Symposium Fernerkundung, Wien*, Mitteilungen der Forstlichen Bundesversuchsanstalt Wien, Heft 157, pages 47–53, 1986.

[Sch89a] W. Schneider. Verfahren, M"oglichkeiten und Grenzen der Fernerkundung f'ur die Inventur des Waldzustandes. FBVA Berichte, Schriftenreihe der Forstlichen Bundesversuchsanstalt, Wien, 1989. Sonderheft.

[Sch89b] Werner Schneider. Verfahren, Möglichkeiten und Grenzen der Fernerkundung für die Inventur des Waldzustandes. FBVA-Berichte, Sonderheft, 1989.

[Sch91] Werner Schneider. Problems of radiometric calibration of multilayer photographic remote sensing images. In M. Buchroithner, editor, *Proc. 11th EARSeL Symposium*, pages 5–13, 1991.

[Sch93] Werner Schneider. Landuse mapping with subpixel accuracy from Landsat TM image data. In *Remote Sensing and Global Environmental Change, Proc. of the 25th ERIM Symposium*, volume II, pages 155–161, 1993.

[SD78] Philip H. Swain and Shirley M. Davis, editors. *Remote Sensing: The Quantitative Approach.* McGraw-Hill, 1978.

[SHH84] W. Schneider, P. Heiselmayer, and H.Plank. Land use mapping in Austria with SPOT. In *SPOT Simulation Europe 82, Final Report. CEC Joint Research Centre, Ispra Establishment, SA/I.04.E2.84.11*, pages 79–90. 1984.

[SHP83] W. Schneider, P. Heiselmayer, and H. Plank. Rechnerunterstützte objektivierte Vegetationskartierung aus Farb-Infrarot-Luftbildern, dargestellt am Beispiel der Umgebung des Glocknerhauses. In *Maß-Kartenband 'Hohe Tauern', Veröffentlichungen des österr. Maß-Programms, Band 7*, pages 61–78. Universitätsverlag Wagner, Innsbruck, 1983.

[Sla80] Philip N. Slater. *Remote sensing, optics and optical systems.* Addison-Wesley, 1980.

[SP84] Werner Schneider and Hubert Plank. Correction of SPOT images for atmospheric effects. In *SPOT Simulation Europe 82, Final Report*, pages 245–255. CEC Joint Research Centre, Ispra Establishment, SA/I.04.E2.84.11, 1984.

[Sze89] Richard Szelisky. *Bayesian Modeling of Uncertainty in Low-level Vision.* Kluwer Academic Publishers, 1989.

[Tar87] A. Tarantola. *Inverse Problem Theory.* Elsevier, 1987.

[TC90] B. Telfer and D. Casasent. Ho-Kashyap Optical Associative Processors. *Applied Optices*, 29:1191–1202, March 10, 1990.

[TH84] R. Tsai and T. Huang. Uniqueness and estimation of three–dimensional motion parameters of rigid objects with curved surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:13–27, 1984.

[Tho90] Stelios C.A. Thomopoulos. Sensor integration and data fusion. *Journal of Robotic Systems*, 7(3):337–372, 1990.

[Toe89] Alexander Toet. Image fusion by a ratio of low-pass pyramid. *Pattern Recognition Letters*, 9:245–253, 1989.

[Toe90] Alexander Toet. Hierarchical image fusion. *Machine Vision and Applications*, 3:1–11, 1990.

[Tso90] J.K. Tsotsos. Analyzing vision at the complexity level. *Behavioral and Brain Sciences*, 13(3):423–469, 1990.

[Ull79] S. Ullman. The interpretation of structure from motion. *Proceedings of the Royal Society, London B*, 203:405–426, 1979.

[Wag91] R. Wagner. Verfolgung von Siedlungsentwicklung mit Landsat-TM Bildern. Master's thesis, Universit"at f"ur Bodenkultur, 1991.

[Wha89] S. W. Wharton. Knowledge-based spectral classification of remotely sensed image data. In G. Asrar, editor, *Theory and Applications of Optical Remote Sensing*, chapter 14, pages 548–577. John Wiley & Sons, 1989.

[WKS87] A. Waxman, B. Kamgar-Parsi, and M. Subbarao. Closed–form solutions to image flow equations for 3d structure and motion. *International Journal of Computer Vision*, 1:239–258, 1987.

Chapter 4

Stereovideometry and Spatial Object Recognition

Project coordinator:

o.Univ. Prof. D.I.Dr. H. Kahmen,
Institute of National Surveying and Engineering Geodesy,
Technical University, Vienna.

4.0 General Problem Statement

Goal of the subproject is 3D-reconstruction of spatial formations from 2D-images. Besides it is concerned in solution of a singular inverse problem. The singularity can be eliminated when aditional informations are given to the 2D-images. In this subproject one or more sets of resembling informations 2D-images are used. There are four tasks in the research project. Three of them are using different methods for taking picture of an object and object reconstruction. The fourth task deals with the visualization.

TASK 1

For on-line acquisition (electrooptical scanning, image data transfer, A/D conversion, image storage) and 3D-information extraction from static or moving scenes a digital stereo camera system shall be developed, where cameras can rotate about vertical and tilt axes. The rotation, orientation, and positioning of the cameras should be controlled by a computer so that scenes can be scanned producing a series of quasi identical images.

TASK 2

The main focus is on geometric accurarcy and reliability of the restituted coordinates, object features and surface elements. For obtaining the utmost accuracy, subpixel measurements are essential. Feature extraction, feature matching as well as feature modelling in conjunction with area based matching techniques will be one of the main problems which must be dealt with. The introduction of background information stored in a kind of knowledge base should support solving ambiguities, detecting and locating object details, and increasing the reliability, while geometric constraints are taken into account too.

TASK 3

The generation of "full" geometric models from complex spatial point heaps is a basic requirement for most systems working further on the sensor information (e.g. robot vision systems, recognition systems). The closer such a geometric model is to models used in already existing systems the better the embedding of sensor data will be. In Task 4 the dexel model is investigated as the model into which the heaps of points are converted. The dexel model has already proven to be very applicable in different industrial applications.

Common basis of this subproject is recovery of object features in the image for the 3D-reconstruction. For this purpose various matching methods shall be used or developed. The fourth task has to solve the problem visualization of complex spatial point heaps.

4.1 Optical 3D Sensing of Static and Moving Scenes by Computer-Controlled Rotating Cameras and 3D Information Extraction

Task leader:

o.Univ. Prof. D.I.Dr. H. Kahmen,
Institute of National Surveying and Engineering Geodesy,
Technical University, Vienna.

4.1.1 Problem Statement

Measuring with a theodolite we only get one direction at a special target on the surface of the object. Using a photogrammetric camera a bundle of directions is measured covering the total area of the image and 3D information may be extracted with much higher flexibility. If a photogrammetric system shall be used for automated measuring processes it has to be a real-time system, processing digital images. The heart of such a system consist of an image processing unit and instead of film-based cameras self-scanning solid state cameras (CCDs) are used. Comparing film-based cameras and self-scanning solid state cameras the following characteristics of CCDs may be highlighted:

- Very small image area
- Limited spatial resolution
- Fast image acquisition
- Stable sensor chip geometry
- Electronic distortions.

There are several experiments to overcome the small image format of CCD arrays:

- Design of larger formats (LF-CCDs)(still restricted to scientific applications and laboratory research) [Len89]. Recent developments have tried to solve this problem by using very large CCD chips with several million photodiodes, 5000x5000 pixels have been anounced.
- Scanning Cameras: The CCD-sensor is placed in the image plane of alarge format camera. A shifting device moves the sensor both within the image plane and perpendicular to it. A computer controlled device makes it possible to record the large format sequentially in partial images of about 5 square mm. The position of the digital partial image can be determined by the actual position of the sensor with an accuracy of about 1 μ m. This means, that the positioning accuracy of the scanner determines the measurement accuracy. Scanning systems are under development (e.g. the large format scanner UMK-Highscan) using four sensors simultaneously, each recording a quater of the image plane (166 mm x 120 mm). Scanning of one large format lasts about 10 minutes [GR93]. The position of the partial image can also be determined by a réseau. This type of scanning system is called réseau-scanning camera. A typical system is the Réseau Scanning Camera (RSC) produced by Rollei [LWE87, Rie89, Rie92].
- Robot eye-in-hand using fiber optics (used only for the inspection of small objects) [Hun90].

In this research project the disadvantages of the small format of CCD arrays shall be overcome by computer controlled rotating cameras. Cameras of this type nowadays are available in form of video theodolites. The tangent screws of these theodolites for the horizontal and vertical axis were replaced by stepper motors with mounted coars/fine gears. The telescope can automatically focused by a stepper motor driven spindle and an additional feature of the telescope is an integrated CCD camera for image acquisition.

The computer controlled motor driven video theodolites, system control software and vision software made it possible to create systems for automatic pointing, if well defined targets are given [GB77, HK88, Kah91a, Kah91b, Kah91c, Kah91d, Kah92b, Kah92a, KSS84, KS83, Kat89, Mof87]. Using the telescope of the theodolite in combination with the CCD camera enables to improve the accuracy capacity of the CCD cameras to that of the high precision theodolites, which could amount to 1/200000 [Got89].

This research project shall base on these developments but mainly deal with high accuracy dimensional measurement using mosaicking of image patches and non-targeted object features.

4.1.2 Research Goals and Methods

The goal of the project is the contactless and continuous measurement of the features (coordinates, deformation) of an object by means of two CCD cameras integrated in the telescope of the motor driven theodolites. As the CCD sensors have a limited image sensor area and therefore a limited number of picture elements (pixels), the object must be scanned by means of sequences of pictures to achieve a proper accuracy in object coordinates. Two motor driven video theodolites shall be used simultaneously for image acquisition of quasi identical images. The system works on the photogrammetric principle with the CCD cameras as imaging devices and the theodolits as precise direct orientation tools. If there are not enough informations about the scene, in a first step the cameras can be used as wide angle cameras and features in the scene are then detected roughly by image processing methods. Finally, the rough three dimensional positions can than be used to guide the telescopes in a second step for high precision point on line detection. There are various strategies possible of doing this.

The system calibration involves the determination of the relative geometric relationship between the two theodolites the determination of the camera interior parameters and the determination of the geometric relatioship between the cameras and the theodolites. The latter is stated by three positional parameters and three rotational parameters with respect to the telescope coordinate system. To avoid a drifting of these positional and rotational parameters, a calibration method has to be developed, which can be used frequently and can always be finished after a very short time.

A very fast image acquisition and proccessing system is responsible for the control of the two motor driven theodolites to get stereo picture pairs of the whole scene by means of image proccessing algorithms. The many narrow imaging bundles of the rotating cameras have to be transformed into singl imaging bundles. To improve the geometric and radiometric quality of the reconstructed images extensive image preprocessing procedures have to be developed and used. Detection and recognition - determination of the location of particular and conjugate features of the object - might prove to be one of the most difficult parts of the project, especially when targeting is not used.

As two CCD cameras are simultaneously used for image acquisition of quasi identical images, extensive quantities of data have to be processed during a very short time. For an optimal solution of the problem the vision system should comprise the following modules:

- a very fast CPU (50 MHz)
- videoboard for a very fast A/D conversion (10 MHz)
- image memories with a very short access time
- a very fast bus for connection of oll the modules
- a very fast videobus for transfer of image data between the CPU
- and the videoboard.

During the first part of the project (1991-1993) a vision system could be installed. For image acquisition in the future two motor driven video theodolites shall be used to get the orientation location of the image bundles with a very high accuracy.

To solve this measuring problem the following scientific and technical topics have to be studied and solved:

- Robot theodolite control for scanning objects and producing stereo images
- Scanning objects
- Image fusion
- Enhancement of image features
- Contactless measurement
- Image matching
- Real-time measurement
- Camera calibration
- Model development

4.2 Subpixel Target Location and Surface Matching for Precise Photogrammetric Measurement

Task leader:

Univ. Doz. Dipl. Ing. Dr. techn. Josef Jansa,
Institute for Photogrammetry and Remote Sensing,
Technical University, Wien.

4.2.1 Problem Statement

Digital image processing for photogrammetric tasks becomes increasingly important. Analogue analytical compilation methods are gradually replaced by workstations comprising analytical methods for the geometric restitution and digital image processing tools for the support of the mensuration task. The main objective of photogrammetry is the measurement of 3D coordinates of points, the extraction of linear features, and the determination of object surfaces. Currently the measurments are still done by human operators and only a few tasks are supported by automatic procedures. The aim of photogrammetric image processing must be

- maximizing the automatic support and thus avoiding erroneous or less precise measurements of a tired operator during cumbersome and time-consuming
- avoiding a loss of reliability of the results due to incorrect automatic interpretations.

High geometric precision of the results and their reliability are the most important matters. Photogrammetric image processing has to be a combination of interpretation (often pattern recognition) and precise location of object features found.

The photogrammetric mensuration task can be subdivided into the following sections [Kra93, Kra84]:

- interior and exterior orientation of the photographs (camera calibration and bundle adjustment)
- recognition and extraction of object features (interpretation and segmentation)
- stereo (or multi-image) feature classification and assignment (= matching in the broad sense)
- measurement of object details, usually isolated (possibly targeted) points, curves and surfaces in 3D space (= matching in the narrow sense)
- object reconstruction (3-D and/or 2-D mapping)
- estimation of accuracy achieved for orientation and compilation
- visualization of intermediate and final results (such as enhanced and rectified image data or data derived through image processing methods)

As far as employing digital methods in photogrammetric practice is concerned the accuracy obtained through digital image processing ought to be comparable to the accuracy achievable in conventional photogrammetric systems which are using high-resolution photographs taken by special cameras, precise optical and mechanical instruments, and well experienced human operators.

Digital photogrammetry is based on digital images, which have to be captured in an appropriate way. For precise measurements geometrically and radiometrically high-quality images are required. Basically there are two possibilities for obtaining digital image data: By using suitable cameras, which directly deliver digital image data; or by scanning conventional photographic films.

Digital photogrammetric close-range systems using CCD video or CCD digital cameras were developed for various purposes such as point determination and surface measurements in medicine and mechanical manufacturing. These systems work very well under laboratory conditions but there are still problems in real workspace caused by disadvantageous often time-variant properties (dependencies on temperature, poor resolutions, small image formats, various electrical interference); e.g. [Bey92].

The advantage of simple off-the-shelf CCD cameras is the possibility of use for real time or near real time applications. Besides this, they are rather cheap and the simultaneous usage of several cameras can be economically justified. More precise and high-resolution cameras are both still quite expensive and slow. The exposure of one image frame may take up to ten minutes or even more, and hence, they cannot be employed if short exposure time is essential (e.g. in case of moving or unstable objects or of moving camera positions). There very few digital camera for metric purposes available (e.g. Zeiss UMK, Rollei Scanning Camera [GR93]).

The utilization of scanners for digitizing films unites three major advantages: *High-resolution (and large-format) metric cameras* can be used, with *short exposure time* and *stable interior orientation.* Furthermore, the main data storage is the photographic film which is still the most effectively compressed image data format. Scanning is done when the data are needed and the resolution is selected depending on the actual application and usefulness [Bäh92, Die92]. But there is, of course, a drawback too: near real time becomes by far impossible. After comparing all possibilities for capturing digital image data the *utilization of image scanners is preferred* as it currently opens the widest field for applications in high-precision photogrammetry where real time in not an important matter.

For digital photogrammetric mensuration tasks the following preconditions are required:

- **High-quality digital image data**, i.e. high geometric and radiometric accuracy as well as good geometric resolution ("small pixels"): Usual pixel sizes are 10 microns to 30 microns (depending on the application) for pictures up to a format of 10 x 10 inches. Photogrammetry uses quite big data sets which image processing tools must be able to deal with.
- **Measurements with subpixel accuracy:** Although the pixels are small they are still big compared to the accuracy obtainable in conventional photogrammetry. Therefore, methods for subpixel-positioning must be used. Least square matching [Ack83, Ack84, Ros86, TAFS86], centroid positioning [Tri89] and other algorithms can fulfill the needs [MAM84, Han88, Zil92]. Though these methods are working very well within continuous surfaces, for targeted points (obtaining a subpixel accuracy down to 1/100 of a pixel) and under ideal conditions, they are very sensitive to any disturbance caused by poor illumination, by image differences due to perspective distortions (especially for close-range applications) and so on. Even filter methods commonly used in image processing may cause displacements.

Digital photogrammetric workstations offered by various companies (e.g. by Intergraph, Leica-Helava, and others) can more or less fulfill the needs for a highly automatic spatial restitution from digital images. On one hand they are mainly designed for supporting the compilation of aerial photographs, on the other hand they are just conventional analytical plotters based on digital images; i.e. the optical viewer is replaced by computer screens. One must not forget that currently these instruments are a first generation of digital workstations as the requirements for reliable processing are still very high as far as computing speed, data storage and robust algorithms is concerned. Often they are not ready for extensive economical utilization. The near future for image processing in digital photogrammetry seems to concentrate on interactive procedures, in which strenuous and time-consuming tasks will be replaced by automated digital methods. In close-range applications the automatic compilations are even more complicated due to great image parallaxes, to occlusions, to great variations of image scales, and last not least quite often due to a lack of texture of the object surface.

Close-range photogrammetry is the salient objective in this research project. There is a great relationship to robot vision tasks. The main difference between these two fields is the request for a reliability and precision measure in photogrammetry in contrast to near real time processing in robot vision. This on the other hand influences the strategies and the procedures used. Photogrammetric measurements are overdetermined and the geometric restitution is closely connected to matching methods by introducing geometric contraints [GB88, Bal91]. Least squares adjustment and the derivation of accuracy values is one of the most important algorithms as far as image matching and perspective restitution are concerned.

Many research institutes all over the world are investigating matching problems for point location and surface reconstruction in photogrammetry utilizing various algorithms. It becomes clear that area based matching which is widely used in photogrammtry cannot solve the problems without support by feature extraction methods. Multi-image matching techniques are employed as well as object space matching, and knowledge about the shape of the object is taken into considerations (e.g. [DPR92, DH92, För88, För91, GS89, Hei91, KBGd92, Li91, RML92, ST92a, Sch91, TS92, VF92, Wil92, Wro91]). There is no general solution available. The only possibility for finding a more universal way is the extensive cooperation between various research sites, which are specializing in certain topics such as metric problems (e.g. camera calibration, camera orientation, geometric 3D and 2D restitution), data structures, image analysis, robot vision and visualization. A research project involving photogrammetrists, mathematicians and computer scientists is the great chance for optimizing the effort of the investigators. The 3-D restitution of close-range objects is just one application in the field of photogrammetry, but it is one of the most demanding tasks. It is not intended to develop a fully automatic procedure. The focus is on creating a reliable interactive tool for high precision surveying by employing metric photographs. (Other publications with useful information about that field: [AF92, AS92, Bäh88, Fra92, Sch91, SH92, Tri87])

4.2.2 Research Goals and Methods

Work Already Done during Previous Research Projects
Preparatory work was done within previous FWF projects (e.g. Forschungsschwerpunkt "Fernerkundung" = Research Project "Remote Sensing"), within related research projects in loose collaboration with an other university (a system for surface measurements for medical purposes was developed by the main investigator of this task at the University of NSW, Australia, during a stay as a research associate), and within the pilot project of FWF which will be finished by the beginning of November 1993. The main areas of that research were:

- extraction of linear features and feature matching for the exact determination of subpixel coordinates of control points (Dr.techn. thesis in progress, almost finished);
- problems of precise camera orientation using targeted points in digital images and measurement of object surface by combining structured light segmentation and least squares area and line based matching (Research project partly finished);
- the development of a user interface and of basic image processing tools on a TRANSPUTER workstation for image processing (Two Dr.techn. theses in progress, advanced state);
- the development of a module for subpixel matching of fiducial marks in photogrammetric images and considerations about the representation of various shapes of fiducials and their automatic recognition and extraction. (Diploma thesis, recently finished);

It is not necessary to mention the expertise of the Institute of Photogrammetry and Remote Sensing in camera calibration, photo orientation by using bundle adjustment, point determination in close-range and aerial applications, DEM modelling, enhancement and classification and rectification of various remote sensing imagery. This expertise is the foundation for solving problems (in particular the metric and geometric ones) which will occur within this project.

Work to be Done
The goal of this research project is to provide basic, robust and reliable modules for photogrammetric mensuration tasks (especially for applications in close-range industrial and architectural surveying). The

methods should work automatically as far as possible without operator interaction (even though operator interaction might not be completely avoided). The precision and reliability of the measurements (i.e. the determination of the 3D object features) should be guaranteed by geometric constraints derived from camera properties and camera orientation (e.g. epipolar geometry). There is a close relation to task 4.2. The installation of feedback loops (= applying previously calculated results) can be advantagous for detecting (partly) occluded features, for increasing the accuracy, for eliminating erroneous measurements, and for support of stereo matching. As conventional stereo photogrammetry is very often replaced by multi-image photogrammetry, the digital compilation methods must also be able to take into account more than two images. Multi-image processing yields higher reliability, robustness and higher accuracy. On the other hand occlusions can be detected more easily and uncertainties may be eliminated.

One of the main problems is the segmentation process and the automatic recognition of prominent object features which deliver important details for the determination of the object surface (corners, edges and their sub-pixels restitution [Sch93] (see also task 3.2) etc.). If the objects are small enough the "structured light" approach is a proper, fast and rather reliable way for both segmentation and matching (e.g. [SW90]). Large objects such as buildings cannot be surveyed in that way. One of the possibilities to solve these problems seems to be the introduction of a knowledge base if appropriate information about the object and object details is available. The contents of the knowledge base may be information whether certain features may occur within the object, or (e.g. in case of inspection tasks in mechanical or civil engineering) the data base may contain all details of the original design. Edge extraction methods alone cannot solve this problem independent of how sophisticated the algorithms are. Shadows and large perspective distortions are disturbing both the classification of the relationship of edges within the various pictures and the assignment to the correct object features. A knowledge base may help to avoid misassignments and mismatches, nevertheless this task will be one of the crucial parts of the research project. The combination of pattern recognition, feature description and knowledge base is also a subject of other tasks within this research project, thus an extensive cooperation with tasks of the project groups 1 and 2 (in particular 1.2, 1.3, 2.1, 2.2) are expected to provide basic algorithms at least at a later state of the research work. The geometric problems of all projects within group 4 (especially those of surface reconstruction, i.e. 4.1, 4.2, 4.4) are similar which automatically leads to an extensive cooperation within this latter group.

The following overview lists the main problems of this task:

- *Handling of large data sets* up to 15000 x 15000 pixels and more as fast as possible (note: real time processing is not the main goal). Image pyramids (see project group 2.1) can provide an appropriate tool (e.g. [KSW92, Li91]). Geometric constraints derived from orientation parameters, on the the other hand, can help in reducing the effort at looking for corresponding image details (see also task 4.3). Besides this, simplier and thus faster matching algorithms can be applied.

- *Recognition of well known features* (stored in a feature data base) and measurement of image coordinates with subpixel accuracy. Least squares matching might be the right method for many cases (especially for continuous sufaces) but subpixel feature matching seems to be the only way for objects with poor texture and for non-continuous surfaces. In some (very close-range) cases projecting an artificial pattern onto the object surface will be the best way for matching uniform areas.

- *Location and measurement of targeted points* which may be used for the geometric orientation of the photographs or which are the object features to be measured. Using retro-reflecting targets is an ideal way for increasing the target-background contrast and for making target identification easier, more reliable, and more accurate. This is very important for the orientation task, as a good camera orientation is the precondition for high-precision measurments.

- *Identification of recognized points* (=number assignment) by support through a knowledge base or multi-image constraints [CC92]. Investigation about the design of number plates and their automatic recognition, although there is already some work done by other institutions (e.g. [vK92]). Is it possible to design a number plate which is readable within a wide range of scales without degrading the pointing accuracy?

- *Extraction and modeling of linear features* and their processing for orientation purposes of the digital images and for surface generation.

- *Coping with perspective distortions and occlusions* and effects caused by shadows which are one of the main problems in image matching and interpretation of close-range applications. A knowledge base and multi-image processing can probably solve many of these problems.

- *Investigation of the influence of scanner resolution and image processing (enhancement) techniques* on the precision of the measurements. Investigations showed that the quality of digital images cannot be increased by digitizing a photograph beyond a certain high resolution due to the lack of information of the original photo as well as to the deterioration of the signal-to-noise ratio when using smaller pixels (e.g. [Die92, TAFS86]). Besides that fact, the amount of data can hardly be handled. The solution of this problems seems to be the digitization using pixel sizes as big as possible on the one hand and the improvement of subpixel matching to fractions of a pixel as small as possible on the other hand. As the accuracy of the location is intended to go down to 1/10 (or even to 1/100) of a pixel one must be very careful in enhancing the original digital data. Changing the relationship of the greyvalues within a neighbourhood means changing the precondition for subpixel positioning. Therefore, the influence of enhancement methods (such as filtering) on the accuracy of the positioning algorithm must be investigated. Briefly: What does an accuracy of e.g. 1/10 of a pixel really mean? Is it just a theoretical value or can it be obtained in practice? Consequently: Which pixel sizes for digitizing are the best to attain a specified geometric matching quality at the smallest amount of digital data?

- *Combination of feature and area based matching methods* for the determination of object surfaces. We know that area based matching alone cannot deliver correct and accurate results in all cases. Features (such as edges, corners or just isolated points) are very often more important information about the object than texture or pattern of the surface. The main problem for feature matching is

  - finding corresponding features in particular in areas with great perspective distortion and
  - developing a reliable algorithm for sub-pixel matching

- Elimination or *reduction of the influence caused by variations of the object illumination* (an important problem for subpixel location). Discrimination of shadows versus real object features.

- *Usage of background information (knowledge base)* to support feature extraction, feature matching and to increase the reliability of the results. Background information can be provided for typical features of photogrammetric pictures (e.g. fiducial marks) and for many surveying tasks in the field of civil engineering, architecture and mechanical manufacturing, in particular in the field of quality check and "reverse engineering".

- *Derivation of accuracy and reliablity values* for the surface model by using geometric properties and values describing the matching quality. Of course the results of a surveying task must be reliable, but they must also be accurate and the final user should know how accurate the data are which he is going to deal with. A proper accuracy model in conjunction with the 3D data model seems to be the only way for providing reasonable information about the quality of the surveying task.

- Appropriate *visualization of 3D models* derived from the images comprising isolated points, linear features, and irregularly distributed surface points. Visualization must be able to combine raster and vector data, and possibly, as a further dimension, accuracy values. This part will be done in close collaboration with project 4.4.

- Investigations for the *utilization of parallel algorithms* for feature extraction and surface matching as well as for visualizing surfaces. Matching is a very time-consuming procedure. One way of optimization might be parallel algorithms. Currently Transputer networks are not the most sophisticated and up-to-date computer technology, particularly for image processing. But they are still a very good and fairly simple tool for developing and checking parallel algorithms. Knowledge of parallel processing will be very useful independent of the future development of processors. This task will be of minor importance as parallel computing is the main topic of project 1.4.

It is unlikely that all points mentioned above can be investigated by one team within the proposed research project. Because of the tight cooperation between the various institutes not all problems need to be solved within this task, thus it might be useful to concentrate on a few topics which seem to be the most important.

The main topics of the research project have been summarized:

- The investigations of the geometric matching quality in conjunction with the *optimization of the digitization parameters* and the *influence of image enhancement techniques* on the matching accuracy.
- The usage of geometric *contraints of multi-image arrangments* and background information for support of *feature extraction*, and feature and area based *matching* respectively.
- *Matching of targeted points* (for orientation purposes and precise point measurement) and surface elements with an accuracy as high as possibly, the design of an *optimum shape and size of targets*, and highly automated procedures for assignment of point numbers.
- The derivation of *accuracy and reliability values* and their *visualization* together with the surface data measured.

4.3 Object Identification Based on Spatial Enumeration Techniques

Task leader:

Dr. S. Stifter,
Research Institute for Symbolic Computation,
University of Linz.

4.3.1 Problem Statement

A novel approach to the problem of (robot) vision is that of introducing the *active perception paradigm* extensively into the vision system. Active processes use feed-back loops, among others a loop between the recognition module and the modeling (or representation) module, [PS91]. Task 3 in Subproject 3 extensively uses this idea. However, this idea is not limited to this special application. It rather suggests to do "lazy evaluation" of the sensor data, which is applicable in many situations. The modeling module in such a vision system lives from the generation of a geometric model of the object to be recognized from sensor data. The sensor data can be assumed to come from a stereoscopic vision system with two cameras. The closer the relation of the geometric model and the sensor data is, the more effective the feed-back loops can be implemented and the better the whole vision system will work.

For geometric models there are several paradigms that have to be satisfied in order to have "correct" geometric models. For the comparison of models with the goal of identifying objects, the following two properties are inevitable:

- *Uniqueness,* i.e. each object has only one model.
- *Unambiguity,* i.e. each model has only one realization, i.e. represents only one object.

These two properties entail that two models are identical if and only if the respective objects (which are determined uniquely) are identical. So the two paradigms have especially to be satisfied for a geometric model used in a robot vision system.

These paradigms have been studied extensively in the area of geometric modeling. The necessity of real time visualization and animation in computer graphics required the development of geometric models that represent objects as a description "how they can be seen" rather than as a description "how they can be built". The tendency in this area is to represent objects by models that contain the information for visualizing them as directly as possible, but that also allow to perform boolean operations and to answer geometric questions about the object. For example, [MS89], [Mor89], [Sti91], consider these topics from different points of view.

Up to now, not so much effort has been undertaken to make the results from geometric modeling and computer graphics also available for (robot) vision systems. A bibliography on this subject is [TM92].

Spatial enumeration techniques are (besides purely polyhedral models) more and more invented for visualization, graphical simulation and animation purposes. Also within robot vision these classes of geometric modeling techniques have already achieved some attention, [JA84]. The basic principle of spatial enumeration techniques is to partition the space into "small" regions (often called cells) and to store for each of the regions whether it is inside the object, outside the object or just intersects the object under certain considerations. Typical instances of spatial enumeration techniques are octree-models, [SW88], dexel-models, [vH86], and layer-models, [HM90].

Similar as the application puts certain requirements on the geometric model, the sensor system and the specific approach to a vision system put certain restrictions on the geometric model. The information that the two cameras can provide and the intended incorporation of feed-back loops have to be used in such a way that the provided information is made avaiable to the application in the most effective and efficient way. The following necessities are most fundamental to the system:

- The model should be *hierarchical:* In order to be able to compare two models on different levels of accuracy within a feed-back loop, a hierarchical modeling scheme is best suited.
- The sensor data must be *convertable* to the model: It is clear that a model that cannot be obtained from the sensor data has no advantage for a vision system.
- The model should be *"independent" of the view direction:* Since usually the view direction (i.e. the orientation of the cameras relative to the object to be identified) is not known, an adequate modeling scheme should be as independent of this view direction as possible.

One of the main operations on a geometric model used in vision is just the opposite of visualization: Images, i.e. optical information has to be converted to an (internal) geometric representation. It seems to be near at hand that this conversion process is easier if the desired representation, the model, is close to the sensor information. The sensor information itself is often sufficient to visualize the considered object with respect to a specific view point depending on the positions of the cameras. However, also other operations are required for the models. So the geometric models used for visualization and graphical simulation should be also adequate for robot vision systems. Note, however, that in robot vision systems the view point is fixed in space, not with respect to the objects to be recognized. This is comparable to simulation systems, but differs from pure visualization purposes.

After these considerations it is not astonishing that in image processing representation techniques have been developed and used for many years that are not so far away from spatial enumeration techniques, although, these techniques are not formulated in full generality and are often based on heuristic results and observations. We refer, for example, to "Hoehenmodelle", see e.g. This also relates this part of the proposal especially with Tasks 1, 2, and 3 of this Subproject, with Tasks 1 and 5 of Subproject 2, and with Tasks 1, 2, and 3 of Subproject 3.

4.3.2 Research Goals and Methods

In the frame of the project "Stereovideometry and Spatial Object Recognition" (supported by FWF as project P8786) we already started to investigate the applicability of the dexel model to (robot) vision and built a first vision system based on the results obtained. Within the proposed project, this system will be generalized and improved, and made available to more classes of objects. Especially, also sparse sets of points will be considered and the vision system will be embedded into a feed-back loop.

The overall goal is to develop a software system that efficiently can solve the following tasks:

- Given:

  A set of points in 3D coming from two cameras together with (possibly) some additional information.

  A model, say A, of an object.
- Find:

  A "suitable" geometric model, say B, for the sensor data, and
- Decide:

  whether the object represented by B is identical with the object represented by A.

Instead of considering just one object A for comparison, also a whole set of objects could be considered. I.e. it has to be decided whether B represents the same object as any model in the set. The objects may also be parameterized, e.g. their size. “Objects” can be technical objects as they appear in robot work cells, but also natural objects as they appear in national surveying. The intention is to also bring results and experience from geometric modeling into the area of image processing.

In the construction process of a geometric model from the data of the sensors, no relevant information must be lost. Especially, if the sensors are mounted on moving parts, the geometry and kinematics of the handling devices have to be taken into account too.

The segmentation of a scene in so-called primitive elements plays a major role for the recognition part and the correct 3D modeling. Segmentation is usually viewed as a major constituent of recognition. First the scene is broken down into meaningful parts, which in turn are recognized. We emphasize, however, that also the reverse is true: without recognition, segmentation is hardly possible. There is no answer yet to the question what should be done first, recognition or segmentation. The most probable answer is a feed-back loop between both processes.

Keeping all this in mind, the development of the vision system is planned in two major components:

1. Processing dense sets of points:

   If the processed sensor information is a dense set of points in 3D, i.e. the distances between points is sufficiently smaller than the required accuracy, then this information should be sufficient to generate a geometric model, e.g. a dexel model.

2. Processing sparse sets of points:

   If the processed sensor information is a sparse set of points in 3D, i.e. the distances between points is larger (or not sufficiently smaller) than the required accuracy, then some additional information will be needed to generate a geometric model, e.g. a dexel model.

For both of these two components, the recognition of an object, i.e. the comparison of the model generated from sensor data and a model obtained previously is crucial. The development of the vision system is planned in three steps that are distinguished by the classes of rotations that are allowed for the objects to be recognized.

1. No rotations are allowed.
2. Rotations in the plane normal to the view (encoding) direction are allowed.
3. Arbitrary rotations are allowed.

The dexel modeling scheme has first been sketched by [vH86]. In keeping the convention with the names pixel and voxel, he introduced the name *dexel* as an abbreviation for depth element. Dexel models represent objects in the following way: A set of parallel and equidistant rays in intersected with the object to be encoded. For each ray the points of intersection with the object are stored. Two points defining a line segment that is totally inside the object make up a dexel. All dexels of a ray are sorted and concatenated to a dexel list. These lists are organized in a dexel matrix.

The conversion of a dense set of points into a dexel model has already been considered in the project “Stereovideometry and Spatial Object Recognition” up to some extent. A vision system based on the algorithmic results has been implemented and will be tested within the next few months on some classes of examples.

The major topics that have been considered within the project “Stereovideometry and Spatial Object Recognition” are:

1. Development and implementation of a module converting dense sets of points into dexel models.
2. Accuracy considerations for the dexel model.

3. Impacts of the accuracy considerations on the use of the dexel model for vision systems.
4. First steps into comparing dexel models that come from sensor data.

The following papers contain the details of these results: [Eis92], [Eis93], [Hei93], [JLMW93b], [JLMW93a], [Sti92], [Sti93b], [Sti93a], [ST92b].

The major topics to be considered within the proposed project are:

1. Refinement of the module converting dense sets of points into dexel models.
2. Further tests of the vision system based on dense sets of points in close cooperation with the other subtasks:
   - Using data from national surveying.
   - Using data from robot vision.
3. Processing sparse sets of points:
   - Development a concept for representing additional information.
   - Conversion of sparse sets of points plus addtional information into dexel models.
4. Embedding of the two modules (for dense and sparse sets of points, respectively) into feed-back loops.

   One key idea of the active perception paradigm is to embed the segmentation/recognition process into a feed-back loop. This will be realized within the planned vision system.
5. Comparison of models.

   This will be done in the three steps, taking different levels of rotations into accout.
6. Comparison of the system based on dexel models with other systems developed within the frame of the proposed project.

For Items 1 and 2, by the investigations in the current project, it is guaranteed that they can be realized in the suggested way. Items 3 and 5 are advanced research goals. The extent up to which they can be realized within the proposed project is still open. However, by the work done in the current project it is ensured that they can be realized at all, the problems that remain may be stability or efficiency problems.

Bibliography

[Ack83] F. Ackermann. High precision digital image correlation. In *Proceedings der 39. Photogrammetrischen Woche*, Stuttgart, 1983.

[Ack84] F. Ackermann. Digital image correlation: Performance and potential application in photogrammetry. *Photogrammetric Record*, 11, 1984.

[AF92] C. Adamos and W. Faig. Hough transform in digital photogrammetry. In L.W. Fritz and J.R. Lucas, editors, *International Archives of Photogrammetry and Remote Sensing, Vol XXIX, Part B3, Commission III*, pages 250–254, Washington D.C., 1992. ISPRS, Committee of XVII International Congress for Photogrammetry and Remote Sensing.

[AS92] R. Al-Tahir and T. Schenk. On the interpolation problem of automated surface reconstruction. In L.W. Fritz and J.R. Lucas, editors, *International Archives of Photogrammetry and Remote Sensing, Vol XXIX, Part B3, Commission III*, pages 227–232, Washington D.C., 1992. ISPRS, Committee of XVII International Congress for Photogrammetry and Remote Sensing.

[Bäh88] H.-P. Bähr. Measures for geometric resolution of digital cameras. In *International Archives of Photogrammetry and Remote Sensing, Vol XXVII, Part B9, Commission I*, pages 61–74, Kyoto, 1988. ISPRS, Committee of XVII International Congress for Photogrammetry and Remote Sensing.

[Bäh92] H.-P. Bähr. Appropriate pixel size for orthophotography. In L.W. Fritz and J.R. Lucas, editors, *International Archives of Photogrammetry and Remote Sensing, Vol XXIX, Part B1, Commission I*, Washington D.C., 1992. ISPRS, Committee of XVII International Congress for Photogrammetry and Remote Sensing.

[Bal91] E.P. Baltsavias. Multiphoto geometrically constrained matching. Technical Report Mitteilungen Nr.49, Institut für Geodäsie und Photogrammetrie an der Eidgenössischen Technischen Hochschule, Zürich, 1991.

[Bey92] H. Beyer. Geometric and radiometric analysis of a ccd-camera based photogrammetric close-range system. Technical Report Mitteilungen Nr.51, Institut für Geodäsie und Photogrammetrie an der Eidgenössischen Technischen Hochschule, Zürich, May 1992.

[Bra90] G. Brandstaetter. On the importance of projective geometry for analytical and digital photogrammetric restitution. *ISPRS Arcive*, 28(6), 1990.

[Bra91a] G. Brandstaetter. Notizen zur voraussetzungslosen gegenseitigen orientierung von messbildern. *Oesterr. Zeitschr. f. Vermessungswesen*, 79(4), 1991.

[Bra91b] G. Brandstaetter. Voraussetzungslose relative orientierung mittels bilddrehung durch bildkorrelation. *Oesterr. Zeitschr. f. Vermessungswesen*, 79(4), 1991.

[Bra91c] G. Brandstaetter. Zur relativen orientierung projektiver buendel. *Zeitschr. f. Fernerkundung u. Photogrammetrie*, 59(6), 1991.

[Bra92] G. Brandstaetter. Notes on the direct projective transformation of general stereo pairs into the rigorous normal case by image correlation. *ISPRS Archives of Photogrammetry and Remote Sensing*, XXIX(B3):701–706, 1992.

[Bra93a] G. Brandstaetter. On the orientation of axial stereo pairs by projective image correlation. *25th Intern. Symposium on Remote Sensing and Global Environmantal Change*, in edition, 4 1993.

[Bra93b] G. Brandstaetter. On the transformation to the normal case of general projective stereo pairs. *2nd Conference on Optical 3-D Measurement Techniques Zuerich*, in preparation, 10 1993.

[CC92] J. Chen and T.A. Clarke. The automatic recognition, location and labelling of targets in digital photogrammetric engineering measurements. In L.W. Fritz and J.R. Lucas, editors, *International Archives of Photogrammetry and Remote Sensing, Vol XXIX, Part B5, Commission V*, pages 686–693, Washington D.C., 1992. ISPRS, Committee of XVII International Congress for Photogrammetry and Remote Sensing.

[DH92] H. Diehl and C. Heipke. Surface reconstruction from data of digital line cameras by means of object based image matching. In L.W. Fritz and J.R. Lucas, editors, *International Archives of Photogrammetry and Remote Sensing, Vol XXIX, Part B3, Commission III*, pages 287–294, Washington D.C., 1992. ISPRS, Committee of XVII International Congress for Photogrammetry and Remote Sensing.

[Die92] H. Diehl. Optimal digitization steps for usual film material. In L.W. Fritz and J.R. Lucas, editors, *International Archives of Photogrammetry and Remote Sensing, Vol XXIX, Part B1, Commission I*, Washington D.C., 1992. ISPRS, Committee of XVII International Congress for Photogrammetry and Remote Sensing.

[DPR92] S.J. Dickinson, A.P. Pentland, and A. Rosenfeld. 3-d shape recovery using distributed aspect matching. *IEEE Transactions on PAMI*, 14(2):174–198, February 1992.

[Eis92] R. Eisenstöck. Visualization of algebraic surfaces using the dexel model. Master's thesis, Diploma Thesis, RISC-Linz, Johannes Kepler University, A-4040 Linz, Austria., 1992.

[Eis93] R. Eisenstöck. Visualization of algebraic surfaces using the dexel model. *Workshop on Computational Geometry, Hagen Germany*, March 1993.

[För88] W. Förstner. Model-based detection and location of houses as topographic control points in digital images. In *International Archives of Photogrammetry and Remote Sensing, Vol XXVII, Part B11, Commission III*, pages 505–517, Kyoto, 1988. ISPRS, Committee of XVII International Congress for Photogrammetry and Remote Sensing.

[För91] W. Förstner. Statistische Verfahren für die automatische Bildanalyse und ihre Bewertung bei der Objekterkennung und -vermessung. Technical Report Reihe C, Heft 370, Deutsche Geodätische Kommission, München, 1991.

[Fra92] C.S. Fraser. Photogrammetric measurement to one part in a million. *Photogrammetric Engineering and Remote Sensing*, 58(3):305–310, March 1992.

[GB77] R. Gottwald and W. Berner. Electronic theodolite-sensor systems for real-time photogrammetry. In Gruen, editor, *Proc. of the ISPRS Intercommision Conference on Fast Processing of Photogrammetric Data*, Interlaken, 1977.

[GB88] A.W. Grün and E.P. Baltsavias. Geometrically constrained multiphoto matching. *Photogrammetric Engineering and Remote Sensing*, 54(5):633–641, May 1988.

[Got89] R. Gottwald. Kern e2-se, a new instrument for industrial surveying. Technical report, Kern Swiss Industrial Metrology, 1989.

[GR93] R. Goldschmidt and U. Richter. Neue Hardwarekomponenten für die digitale Nahbereichsphotogrammetrie. *ZPF*, 60(2):71–75, March 1993.

[GS89] J.S. Greenfeld and A.F. Schenk. Experiments with edge-based stereo-matching. *Photogrammetric Engineering and Remote Sensing*, 55(12):1771–1777, December 1989.

[Han88] M.J. Hannah. Digital stereo image matching techniques. In *International Archives of Photogrammetry and Remote Sensing,Vol XXVII, Part B3, Commission III*, pages 280–293, Kyoto, 1988. ISPRS, Committee of XVII International Congress for Photogrammetry and Remote Sensing.

[Hei91] J. Heikkilä. Use of linear features in digital photogrammetry. *The Photogrammetric Journal of Finland*, 12(2):40–56, 1991.

[Hei93] J. Heinzelreiter. *Simulation and algorithmic verification of NC machining.* PhD thesis, RISC-Linz, Johannes Kepler University, Linz, Austria, 1993. PhD Dissertation, In preparation.

[HK88] W. Huep and O. Katowski. Theodolitmesssysteme fuer industrielle geodaetische Messungen. *Tehnische Rundschau*, 39, 1988.

[HM90] J. Heinzelreiter and H. Mayr. Solid modeling of dynamic objects using spatial enumeration techniques, a survey. Technical report, RISC-Linz, Johannes Kepler University, Linz, Austria, 1990.

[Hun90] V.D. Hunt. *Understanding Robotics.* Academic Press INC, New York, 1990.

[ISP93] ISPRS. Isprs congress reports. *ISPRS Journ. of Photogr. and Rem. Sens.*, 48(2):26–48, 1993.

[JA84] C.H. Chien J.K. Aggarwal. A volume/surface octree representation. In IEEE, editor, *Proc. 7th Int. Pattern Recognition Conference*, pages 817–820, 1984.

[JLMW93a] St. Jais, R. Leitner, M. Marko, and W. Wolf. Rovi: program documentation. Technical report, RISC-Linz, Johannes Kepler University, Linz, Austria, 1993.

[JLMW93b] St. Jais, R. Leitner, M. Marko, and W. Wolf. Rovi: Project description. Technical report, RISC-Linz, Johannes Kepler University, Linz, Austria, 1993.

[Kah91a] H. Kahmen. Polare Ortungsverfahren in der Hydrographie. In *27. DVW Woorkshop*. Oldenbourg, Wien, 1991.

[Kah91b] H. Kahmen. Robot techniques and their impact on surveying systems. In *I. International Symposium on Apllications of Geodesy to Engineering*, Stuttgart, 1991.

[Kah91c] H. Kahmen. Roboter und ihr Einfluss auf Vermessungssysteme. *Vermessungswesen und Raumordnung*, 53(7):310–320, 1991.

[Kah91d] H. Kahmen. Satellitengestuetzte Positionierugsverfahren und Messroboter fuer die Absteckung von Konstruktionen im Bruecken- und Hochbau. *Geodetski List Zagreb*, 68(45), 1991.

[Kah92a] H. Kahmen. Hochpraezise Zielerfassung durch Robothertheodolite. In Matthias/Gruen, editor, *Ingenieurvermessung 92*, Bonn, 1992.

[Kah92b] H. Kahmen. Hybrid measurement robots in engineering surveys. In *ISPRS Congress*, Washington, 1992.

[Kat89] O. Katowski. Deformationsmessung an Bauwerken mit dem automatischen Theodolit-Messsystem ATMS. In Gruen/Kahmen, editor, *Optical 3-D Measurement Techniques*, Karlsruhe, 1989. Herbert Wichman Verlag.

[KBGd92] O. Kölbl, U. Bach, D. Gasior, and K. de Laporte. Multi-templet-matching for the automation of photogrammetric measurements. In L.W. Fritz and J.R. Lucas, editors, *International Archives of Photogrammetry and Remote Sensing, Vol XXIX, Part B3, Commission III*, pages 540–548, Washington D.C., 1992. ISPRS, Committee of XVII International Congress for Photogrammetry and Remote Sensing.

[Kra84] K. Kraus. *Photogrammetrie*, volume 2. Dümmler, Bonn, 1984.

[Kra93] K. Kraus. *Photogrammetry*, volume 1. Dümmler, Bonn, 4. edition, 1993.

[KS83] H. Kahmen and H. Suhre. Ein Lernfaehiges tachymetrisches Vermessungsystem zur Ueberwachung kinematischer Vorgaenge ohne Beobachter. *Zeitschrift f. Vermessungswesen*, 108(8):345–351, 1983.

[KSS84] H. Kahmen, H. Schwaeble, and H. Suhre. Ein "intelligentes" polares Vermessungssystem fuer die Beobachtung statischer Punktfelder und kinematischer Vorgaenge. *Zeitschrift f. Vermessungswesen*, 109(11):553–563, 1984.

[KSW92] B. Kaiser, M. Schmolla, and B.P. Wrobel. Application of image pyramid for surface reconstruction with FAST vision (=FAcet Stereo Vision). In L.W. Fritz and J.R. Lucas, editors, *International Archives of Photogrammetry and Remote Sensing, Vol XXIX, Part B3, Commission III*, pages 341–345, Washington D.C., 1992. ISPRS, Committee of XVII International Congress for Photogrammetry and Remote Sensing.

[Len89] R. Lenz. Image data acquisition with ccd cameras. In Gruen/Kahmen, editor, *Optical 3-D Measurement Techniques*, Karlsruhe, 1989. Herbert Wichman Verlag.

[Li91] Mengxiang Li. Hierarchical multi-point matching with simultaneous detection and location of breaklines. *Transactions of the Royal Institute of Technology, Department of Photogrammetry*, ?(55), July 1991.

[LWE87] T. Luhman and W. Wester-Ebbinghaus. Digital image processing by means of reseau-scanning. In *Proceedings Photogrammetrische Woche*, Stuttgart, 1987.

[MAM84] E. Mikhail, M. Arkey, and O. Mitchell. Detection and sub-pixel location of photogrammetric targets in digital images. *Photogrammetria*, 39:63–83, 1984.

[Mof87] N. Moffit. Kern system for positioning and automated coordinate evaluation - a real-time system for industrial measurement. *Tehnical Papers 1987, ACSM-ASPRS Annual Convention*, 1987.

[Mor89] M.E. Mortenson. *Geometric Modeling.* J. Wiley & Sons, 1989.

[MS89] H. Mayr and S. Stifter. Off-line generation of error-free robot/nc code using simulation and automatic programming techniques. In *Proc. Int. Conf. on CAD/CAM and AMT in Israel, North Holland*, December 1989.

[PS91] W. Pölzleitner and S. Stifter. Active sensing using associative processes and 3d geometric modeling. In *Proc. 14. ÖAGM Arbeitstreffen "Image Acquisation and Real-Time Visualization"*, pages 139–151, Salzburg, Austria, May 1991. Oldenbourg Verlag.

[Rie89] W. Riechmann. Hochaufloesende digitale Objekterfassung mit Hilfe einer Résau-Scanning Camera. In Gruen/Kahmen, editor, *Optical 3-D Measurement Techniques*, Karlsruhe, 1989. Herbert Wichman Verlag.

[Rie92] W. Riechmann. The accuracy of the reseau-scanning camera (rsc) in comparision with theodolite and photogrammetric systems. In *International Archives of Photogrammetry and Remote Sensing,Vol XXIX, Part 5*, Washington D.C., 1992. ISPRS, Committee of XVII International Congress for Photogrammetry and Remote Sensing.

[RML92] R.L. Russell, R.A. McClain, and B.P. Landell. A 3-d model extraction system. In L.W. Fritz and J.R. Lucas, editors, *International Archives of Photogrammetry and Remote Sensing,Vol XXIX, Part B3, Commission III*, pages 446–451, Washington D.C., 1992. ISPRS, Committee of XVII International Congress for Photogrammetry and Remote Sensing.

[Ros86] Dan Rosenholm. Accuracy improvement in digital matching. Technical Report 52, Transactions of the Royal Institute of Technology, Department of Photogrammetry, 1986.

[Sch91] C.-T. Schneider. Objektgestütze Mehrbildzuordnung. Technical Report Reihe C, Heft 375, Deutsche Geodätische Kommission, München, 1991.

[Sch93] W. Schneider. Land use mapping with sub-pixel accuracy from LANDSAT TM image data. In *The 25th International Symposium, Remote Sensing and Global Environmental Change; Presented Paper*, Graz, Austria, 4-8 April 1993.

[SH92] F. Schneider and M. Hahn. Matching images of different geometric scale. In L.W. Fritz and J.R. Lucas, editors, *International Archives of Photogrammetry and Remote Sensing,Vol XXIX, Part B3, Commission III*, pages 295–302, Washington D.C., 1992. ISPRS, Committee of XVII International Congress for Photogrammetry and Remote Sensing.

[ST92a] T. Schenk and Ch.K. Toth. Reconstructing small surface patches from multiple images. In L.W. Fritz and J.R. Lucas, editors, *International Archives of Photogrammetry and Remote Sensing,Vol XXIX, Part B3, Commission III*, pages 255–258, Washington D.C., 1992. ISPRS, Committee of XVII International Congress for Photogrammetry and Remote Sensing.

[ST92b] S. Stifter and St. Taferner. The impact of geometric modeling on robot vision. In *Proc. 16th ÖAGM Meeting on Pattern Recognition*, Vienna, Austria, May 1992. Oldenbourg Verlag.

[Sti91] S. Stifter. Simulation and verification of nc machining using symbolic computation techniques. In *Proc. IMACS "91 Congress on Computation and Applied Mathematics*, Dublin, Ireland, July 1991.

[Sti92] S. Stifter. Modeling of objects that may change its shapes. In *Proc. 4th Int. Symp. on Systems Analysis and Simulation*, Berlin, Germany, August 1992. Elsevier Publ. Comp.

[Sti93a] S. Stifter. Modeling of objects that may change their shapes: accuracy considerations for dexel models. *Journal Systems Analysis – Modeling – Simulation*, 1993. To appear.

[Sti93b] S. Stifter. Simulation of nc machining based on the dexel model: a critical analysis. *Int. Journal of Advanced Manufacturing Technology*, 1993. To appear.

[SW88] H. Samet and R.E. Weber. Hierarchical data structures and algorithms for computer graphics, part i: fundamentals. *IEEE Computer Graphics and Applications, vol. 8, no. 3*, pages 48–68, 1988.

[SW90] Th. Stahs and Fr. Wahl. Oberflächenvermessung mit einem 3D-Robotersensor. *ZPF*, 58(6):190–202, June 1990.

[TAFS86] K. Torlegard, U. Almroth, J.-O. Fallvik, and J. Stokes. Digital matching of stereo images. Technical Report 51, Transactions of the Royal Institute of Technology, Department of Photogrammetry, 1986.

[TM92] St. Taferner and R. Mittermaier. Robot vision: articles summary. Technical report, RISC-Linz, Johannes Kepler University, Linz, Austria, 1992.

[Tri87] J.C. Trinder. Measurements on digitized hardcopy images. *Photogrammetric Engineering and Remote Sensing*, 53(3):315–321, March 1987.

[Tri89] J.C. Trinder. Precision of digital target location. *Photogrammetric Engineering and Remote Sensing*, 55(6):883–886, June 1989.

[TS92] Yi-Hsing Tseng and T. Schenk. A least-squares approach to matching lines with fourier descriptors. In L.W. Fritz and J.R. Lucas, editors, *International Archives of Photogrammetry and Remote Sensing, Vol XXIX, Part B3, Commission III*, pages 469–475, Washington D.C., 1992. ISPRS, Committee of XVII International Congress for Photogrammetry and Remote Sensing.

[VF92] R. Vaillant and O.D. Faugeras. Using extremal boundaries for 3-d object modeling. *IEEE Transactions on PAMI*, 14(2):157–173, February 1992.

[vH86] T. van Hook. *Real-time shaded NC milling display*, volume 20, pages 15–20. ACM Siggraph, 4 edition, 1986. In: D.C. Evans, R.J. Athay, (eds.), Computer Graphics.

[vK92] F.A. van den Heuvel and R.J.G.A. Kroon. Digital close-range photogrammetry using artificial targets. In L.W. Fritz and J.R. Lucas, editors, *International Archives of Photogrammetry and Remote Sensing, Vol XXIX, Part B5, Commission V*, pages 222–229, Washington D.C., 1992. ISPRS, Committee of XVII International Congress for Photogrammetry and Remote Sensing.

[Wil92] A. Wilkin. Robust 3d-object representation by linear features. In L.W. Fritz and J.R. Lucas, editors, *International Archives of Photogrammetry and Remote Sensing, Vol XXIX, Part B5, Commission V*, pages 659–666, Washington D.C., 1992. ISPRS, Committee of XVII International Congress for Photogrammetry and Remote Sensing.

[Wro91] B.P. Wrobel. Least-squares methods for surface reconstruction from images. *Photogrammetry and Remote Sensing*, 46(2):67–84, April 1991.

[Xu93a] X. Xu. A feature based signal relational tree matching technique used in photogrammetry. *3rd Intern. Conference for Young Computer Scientists Peking*, in edition, 7 1993.

[Xu93b] X. Xu. Hybrid methods for digital image matching. *Mittlg. d. geodaet. Inst. d. TU Graz*, in preparation, 1993.

[Xu93c] X. Xu. A hybrid stereo algorithm for digital photogrammetric image matching. *2nd Conference on Optical 3-D Measurement Techniques*, in edition, 10 1993.

[Zil92] O. Zilberstein. Relational matching for stereopsis. In L.W. Fritz and J.R. Lucas, editors, *International Archives of Photogrammetry and Remote Sensing, Vol XXIX, Part B3, Commission III*, pages 711–719, Washington D.C., 1992. ISPRS, Committee of XVII International Congress for Photogrammetry and Remote Sensing.