

Pattern Recognition and Image Processing Group  
Institute for Computer Graphics and Algorithms  
Vienna University of Technology  
Favoritenstr. 9/1863  
Phone: +43(1)-58801-18661  
Fax: +43(1)-58801-18697  
email: sec@prip.tuwien.ac.at

## PRIP Technical Reports, September 23, 2013

The list includes TR-number, author(s), title, year and abstracts of all technical reports<sup>1</sup>.

[PRIP-TR-001] Walter G. Kropatsch. Digitales Sehen mit Bildpyramiden. 1991.

**Abstract:** Digital Computers are nowadays capable to acquire and to process visual information. The research area of 'Computer Vision' has as its goal to teach the machine abilities comparable to the human eye. At first we shall consider characteristic performance achieved by natural vision in order to derive constraints for digital vision. Then the analysis of digital images is classified into different classes of processing steps ordered according to the different degree of abstraction of the involved descriptions. An image pyramid is presented as a hierarchical data structure which is able to contain both images of various resolutions as well as image descriptions of various degrees of abstraction. Several realizations of pyramidal systems do exist and have shown that simple tasks of image analysis can be performed in real time. In dual pyramids a curve pyramid complements a pyramid of images. Some examples of this concept indicate the potential that the model of the pyramid still offers for computer vision.

[PRIP-TR-002] Walter G. Kropatsch. Image Pyramids and Curves - An Overview. 1991.

**Abstract:** This technical report represents the first part of my habilitation (University of Innsbruck, Austria, 1990). It contains the quintessence of my scientific work in the years 1984-1989. The second part of the habilitation consists of a series of my own publications to the subject. It was not reproduced here since the papers already appeared separately. The present report offers a condensed survey over the international literature in this branch, arranged and commented according to a personal weighting in order to relate it to my works on the one hand and on the other hand, to put the significance of the single publications in the greater context of the current state of the art in computer vision. Pyramids are important structures in the processing of digital images. Invented as an ordered collection of images at multiple resolutions they have developed into efficient data and processing structures. They give us the hope to achieve acceptable performance in vision tasks that have to process millions of bytes in extremely short time. Curves represent shapes at an intermediate level. They describe either the boundary of a region or the central axis of an elongated region in the image. Hierarchical curve representations aim at a stepwise reduction of the data to 'significant parts' like corners or curvature extrema while preserving the major property of a curve: its connectivity.

[PRIP-TR-003] Walter G. Kropatsch. Integration of SAR and DEM data - Geometrical Considerations. 1992.

---

<sup>1</sup>document file /pub/prip/prippapers/priptrsd.\*

**Abstract:** General principles for integrating data from different sources are derived from the experience of registration of SAR images with DEM data. The integration in our case consists of establishing geometrical relations between the data sets that allow to accumulate information from both data sets for any given object point (e.g. elevation, slope, backscatter of ground cover, etc.). Since the geometries of the two data are completely different they cannot be compared on a pixel by pixel basis. The presented approach detects instances of higher level features in both data sets independently and performs the matching at the high level. Besides the efficiency of this general strategy it further allows the integration of additional knowledge sources: world knowledge and sensor characteristics are also useful sources of information. The SAR features layover and shadow can be detected easily in SAR images. An analytical method to find such regions also in a DEM needs in addition the parameters of the flight path of the SAR sensor and the range projection model. The generation of the SAR layover and shadow maps is summarized and new extensions to this method are proposed.

[PRIP-TR-004] Horst Bischof, Werner Schneider, and Axel Pinz. Multispectral Classification of Landsat-Images using Neural Networks. 1991.

**Abstract:** Recent progress in neural network research has demonstrated the usefulness of neural networks in a variety of areas. In this work we report the application of three-layer backpropagation networks for classification of Landsat TM data on a pixel by pixel basis. The results are compared to Gaussian maximum likelihood classification. It is shown that the neural network is able to perform better than the maximum likelihood classifier. In an extension of the basic network architecture it is shown that textural information can be integrated into the neural network classifier without the explicit definition of a texture measure. The usage of neural networks for postclassificational smoothing is examined.

[PRIP-TR-005] Walter G. Kropatsch and Annick Montanvert. Irregular Pyramids. 1992.

**Abstract:** A pyramid is a stack of images with exponentially decreasing resolutions. Many image processing algorithms run on this hierarchical structure in  $O(\log n)$  parallel processing steps where  $n$  is a side of the input image. Perturbations in the structure may disturb the originally regular neighborhood relations and also the stability of the results. On the other hand, biological vision is based on piecewise regular patches in the retina, e.g. of a monkey or a human. P. Meer's stochastic pyramid is such an irregular structure. The parallel generation of the structure is governed by two "decimation rules" that also characterize a maximal independent set on the neighborhood graph of the image pixels. In general, the number of neighbors in the decimated graph may increase. It is shown that the decimation  $G'(V', E')$  of any neighborhood-graph  $G(V, E)$  preserves the degrees in the corresponding dual graphs. However the dual of  $G'$  is not always a "good" decimation of the dual of  $G$ . Investigating in parallel dual decimations of regular graphs, one finds unique solutions that have interesting properties for image pyramids. Besides the above theoretical motivation for irregular structures, we can find similar structures in the retinas of monkeys (and also of humans).

[PRIP-TR-006] Renate Bartl. Information Fusion in Remote Sensing. 1992.

**Abstract:** Information fusion deals with the combination of different sources of information at external (world, scene, image) and internal (image description, scene description) levels. Each of these levels has specific properties and requires its own quality

measures and fusion techniques. In this remote sensing application fusion is performed at two levels: at the image level classification results obtained by several classification from one image algorithms are improved; at the image description level classifications from different images are fused to get a more comprehensive result which permits to infer relations between the original images. This information extracted by fusion can be used profitably by other image analysis tasks and/or remote sensing tasks.

[PRIP-TR-007] Horst Bischof and Walter G. Kropatsch. Neural Networks versus Image Pyramids. 1993.

**Abstract:** Neural networks and image pyramids are massively parallel processing structures. In this paper we exploit the similarities as well as the differences between these structures. The general goal is to exchange knowledge between these two fields. After introducing the basic concepts of neural networks and image pyramids we give a translation table of the vocabulary used in image pyramids and those used in neural networks. In the following sections we compare neural networks and image pyramids in detail. We show how a modified Hopfield network can be used for irregular decimation. We examine the type of knowledge stored and the processing performed by pyramids and neural networks. In the case of numerical information, so called 'numerical pyramids' are rather similar to neural networks. But also for 'symbolic pyramids' we show how to implement them by neural networks. In particular we present a neural implementation of the  $2 \times 2/2$  curve pyramid. We derive some general rules for implementing symbolic pyramids by neural networks. Finally we briefly discuss the role of learning in image pyramids.

[PRIP-TR-008] Andreas Dörsam, Walter Kropatsch, and Axel Pinz. Adaptive Image Compression using Fractals and Pyramids. 1991.

**Abstract:** Image pyramids and IFS-Codes have been used independently for image processing and image compression. In this paper several possible combinations of both methods are described. Pyramids could be used to segment an image into regions that are well suited for IFS-Coding and regions that can be better described by other methods. IFS-Codes for image segments could be calculated with the help of pyramids, and combinations of IFS-Codes and Laplacian Pyramids could achieve very high compression rates.

[PRIP-TR-009] Horst Bischof. Modular, Hierarchical, and Geometrical Neural Networks. 1991.

**Abstract:** In order to apply neural networks to large scale, real world tasks, several obstacles have to be overcome. One main deficiency is long learning time. A closer look at the brain reveals that the topology of the brain is considerably different from current neural network models. We show that modular and hierarchical topologies (also common in the brain) offer a potential solution to speed up learning. To build modular and hierarchical network topologies, knowledge from conventional.

[PRIP-TR-010] Robert Sablatnig, Christian Menard, and P. Dintsis. A Preliminary Study on Methods for a Pictorial Acquisition of Archaeological Finds. 1991.

**Abstract:** In this paper we propose a system that could help the archaeologist in his work. First we present an overview of our complete project, of which this preliminary study is the first step to find out whether the process of archiving archaeological

finds could be automated. Next, we present two different acquisition methods, stereo-acquisition and structured-light-acquisition to get the 3D-surface representation ( a so called 3D-object model) of a sherd. Based on the output from preliminary processing, we present the possibilities which can be applied on the object model. The final result of our system is the automatic generation of a profile section of a sherd.

[PRIP-TR-011] Axel Pinz and Peter Datlinger. Digital Image Analysis of Retinal Fundus Images in Age Related Macular Degeneration. 1991.

**Abstract:** Digital Image Analysis and Pattern Recognition methods were applied to fundus photographs and fluorescein angiograms of a patient with age-related macular degeneration taken before and after laser photocoagulation treatment. These analysis techniques provide contrast and brightness enhancement (interactive manipulations), direct comparison of several images in various stages of the disease (image registration), and the extraction of essential retinal structures (blobdetection by image pyramids, edge detection).

[PRIP-TR-012] Axel Pinz, Walter Kropatsch, Alois Hinterleitner, and Johann Petrak. Evaluation of Image Processing Systems for Application in Research and Education. 1991.

**Abstract:** In this report we assess image processing hardware and software from a university institute's point of view. After giving a brief sketch of the overall situation and a review of related work on computer vision systems and image analysis benchmarks, we proceed with a description of typical image analysis tasks. There are basic-, mid- and high-level tasks in education, research and consulting. Three types of image processing systems are defined: general, hardware oriented, and software oriented. We conclude that "the appropriate system" does not exist, so that a heterogeneous network of several different image analysis components constitutes the best solution.

[PRIP-TR-013] Michael A. Neuhauser and Irene J. Leitgeb. Iterated Function Systems; A Direct Discrete Approach with Pyramids. 1992.

**Abstract:** Iterated Function Systems (IFS) are sets of contractive transformations. They define a unique attractor which can be interpreted as a binary image. Since IFS with few transformations can generate very complex images, they can be used for image compression. The difficulty lies in finding an IFS that approximates a given image well; this is known as the inverse problem. We show a new way of computing the discrete attractor of an IFS directly for a specific screen resolution. The run time efficiency of this algorithm is improved by the use of image pyramids. Furthermore, some ideas for approaching the inverse problem from a new direction are presented. We discuss the 1D case with the intention of using the so gained experience in 2D.

[PRIP-TR-015] Axel Pinz and Horst Bischof. Neural Network 'Surgery': Transplantation of Hidden Units. 1992.

**Abstract:** We present a novel method to combine the knowledge of several neural networks by replacement of hidden units. Applying neural networks to digital image analysis, the underlying spatial structure of the image can be propagated into the network and used to visualize its weights (WV-diagrams). This visualization tool helps to interpret the behaviour of hidden units. We notice a process of specialization of certain hidden units, while others remain apparently useless. These units are cut out of one network and replaced by units taken from other networks trained for the same

task using different parameters. We achieve better prediction accuracies for the new, combined network than for any of the two original ones. This constitutes a special kind of information fusion in image understanding. We give an application example from the field of remote sensing, where neural networks are used to interpret the species of trees in aerial photographs. The interpretation accuracy is raised from 85% to 90%.

[PRIP-TR-016] Alois Eder Hinterleitner. Rekonstruktion der Form von Graeben fuer Archaologische Magnetische Prospektion. 1993.

**Abstract:** Magnetic prospection is a very important instrument for preparing an archeological excavation. Apart from this, it enables new information, which can mostly not be obtained by the excavation, because a prospection can cover, in contrast to an excavation, the whole site. To get quantitative information about the subsurface of the earth, it is necessary to determine the shape of the objects producing the magnetic anomalies. In this diploma theses we try to reconstruct the shape of filled ditches. First we describe the measuring method and the image preprocessing steps to visualize the data. Then a magnetic model for filled ditches is presented, which allows the determination of the shape of the ditch by estimating the depth of the humus-like soil at the measuring points. The reconstruction of the depth is done by the Least-Square-Method using a simulated annealing schedule. To get archeologically plausible ditches a regularization of the depth is needed. The Least-Square terms are weighted to make the algorithm more robust. The weights are calculated by using all the information about the measuring procedure. At last the goals of the diploma theses are presented.

[PRIP-TR-017] Michael A. Neuhauser. Diskrete Iterierte Funktionensysteme. February 1993.

**Abstract:** Iterated function systems (IFS) are sets of contractive functions. They define a unique fractal attractor that can be represented as a binary image. They also define a unique invariant measure that can be represented as a greyvalue or color image. IFS are well suited for image coding, because even IFS with few functions are able to generate realistic images of natural objects. Firstly the mathematical basics of IFS will be presented in this work. Then the properties of affine transformation on euclidean spaces will be investigated. An convenient representation of affine transformations on the euclidian plane is developed. A discretization of the transformation is done next, this allows to present two efficient algorithms that compute a discrete attractor. By using image pyramids one of this algorithms is extended and enhanced. An efficient method for calculation of a discrete invariant measure is given. The discrete transformations introduce an error that results in a difference between the attractor, the invariant measure and their discrete counterparts. An error bound for this difference is shown. It follows from this error bound that it is always possible to achieve an accuracy that is sufficient for image coding purposes.

[PRIP-TR-018] Johann Wieser. Layoutanalysis, Finding Text, Titles, and Photographs in Digital Images of Newspaper Pages. February 1993.

**Abstract:** An important step in the analysis of printed documents is the segmentation and classification of blocks into categories such as photographs, titles, paragraphs, etc. This report presents an approach to enhance and combine two commonly used methods, a merging bottom up approach and a cutting top down approach, to segment pages of a newspaper. The planned procedure for an implementation of a layoutanalysis system as a preprocessing module for a commercial product is described.

[PRIP-TR-019] Cornelia Fermueller and Yiannis Aloimonos. The Role of Fixation in Visual Motion Analysis. 1993.

**Abstract:** The human eye is different from existing electronic cameras because it is not equipped with a uniform resolution over the whole visual field. With a small fovea in a large visual field it is not surprising that the human visual system has developed mechanisms, usually called saccades or pursuits, for moving the fovea rapidly. How does this particular ability of humans and primates to fixate on environmental points in the presence of relative motion help their visual systems in performing various tasks? To state the question in a more formal setting, we investigate in this paper the following problem: Suppose that we have an anthropomorphic active vision system, that is, a pair of cameras resting on a platform and controlled through motors by a computer that has access to the images sensed by the cameras in real time. If this machine can fixate on targets that are in motion relative to it, can it perform visual tasks in an efficient and robust manner? By restricting our attention to a set of navigational tasks, we find that such an active observer can solve the problems of 3-D motion estimation, egomotion recovery and estimation of time to contact in a very efficient manner. The possibility that a machine possessing gaze control capabilities can successfully address other problems, such as figure-ground segmentation, stereo-fusion, visual servoing for manipulatory tasks and relative depth.

[PRIP-TR-020] Cornelia Fermueller and Yiannis Aloimonos. Qualitative Egomotion. 1993.

**Abstract:** Due to the aperture problem, the only general unambiguous motion measurement in images is normal flow—the projection of image motion on the gradient direction. In this paper we show how a monocular observer can estimate its 3D motion relative to the scene by using normal flow measurements in a global and mostly qualitative way. The problem is addressed through a search technique. By checking constraints imposed by 3D motion parameters on the normal flow field the possible space of solutions is gradually reduced. In the four modules that comprise the solution, constraints of increasing restriction are considered, culminating in testing every single normal flow value for its consistency with a set of motion parameters. The fact that motion is rigid defines geometric relations between certain values of the normal flow field. The selected values form patterns in the image plane that are dependent on only some of the motion parameters. These patterns, which are determined by the signs of the normal flow values, are searched for in order to find the axes of translation and rotation. The third rotational component is computed from normal flow vectors that are only due to rotational motion. Finally, by looking at the complete data set, all solutions that cannot give rise to the given normal flow field are discarded from the solution space.

[PRIP-TR-021] Axel Pinz, Marek B. Zaremba, Horst Bischof, Francois A. Gougeon, and Michael Locas. Neuromorphic Methods for Recognition of Compact Image Objects. 1993.

**Abstract:** The issue of the recognition of tree species from high resolution aerial images is addressed in this paper. An approach based on the use of neural networks is presented and discussed in more detail. The networks perform classification and recognition operations on compact image objects, obtained by applying different tree isolation procedures. The recognition capabilities of two classes of networks, multilayer feedforward networks and holographic networks, are compared and some results of the research

carried out in Austria and Canada, using aerial photographs and multispectral scanner images, are given.

[PRIP-TR-022] Walter G. Kropatsch and Dieter Willersinn. Parallel Line Grouping and Irregular Curve Pyramids. 1994.

**Abstract:** Parallel lines are important features for object recognition by grouping. Regular 2x2/2 curve pyramids are hierarchical symbolic representations of curves that can be constructed and processed in logarithmic time. The rigidity of the regular structure causes an unstable, shift variant representation of parallel lines. In order to usefully apply the concept of the curve pyramid on grouping problems, the shift variance problem had to be overcome by extending the concept to irregular pyramids. These have a structure that adapts to the image data by deriving control information from curve relations. The algorithm that builds the irregular curve pyramid by deriving higher levels of abstraction from a set of relations goes far beyond merely solving the shift variance problem. It can reduce the computational complexity in comparable applications where all possible combinations of parts have to be checked in order to reassemble complex objects.

[PRIP-TR-023] Irene J. Leitgeb. Iterierte Funktionensysteme - Das eindimensionale inverse Problem. 1993.

**Abstract:** Since images of natural objects often have fractal features, Iterated Function Systems (IFS) seem quite suited for their representation in computers. A further reason to use IFS is their low requirements of memory. An IFS is a set of contractive functions that defines a fractal attractor; this attractor can be interpreted as a binary image. First of all, an explanation of IFS and the inverse problem will be given in this work, followed by an outline of existing approaches to the solution of the inverse problem. A brief introduction to the mathematical basis of IFS will then be given. Then an algorithm is presented that is based upon existing discretisation of transformations of an IFS and algorithms that find a discrete attractor of an IFS. The algorithm solves the inverse problem in the 1D case. It uses the ratio of the length of the black and the white connected components of an attractor which is invariant under the transformations of an IFS. In the discrete case an interval can be defined for this ratio. This information is used to find possible transformations. In consideration of sampling effects, the algorithm tries then to find the parameters for each possible transformation if they exist.

[PRIP-TR-024] Etienne Bertin and Horst Bischof. Voronoi Pyramids controlled by Hopfield Neural Networks. 1993.

**Abstract:** We present an algorithm for image segmentation with irregular pyramids. Instead of starting with the original pixel grid, we first apply some adaptive Voronoi tessellation to the image. This provides the advantage that the number of cells in the bottom level of the pyramid is already reduced as compared to the number of pixels of the original image. Furthermore the Voronoi diagram is a powerful tool for shape description and image compression. For the construction of the irregular pyramid we present a Hopfield neural network which controls the decimation process. In this paper we extend our previous results by proving a more general theorem. The contributions of this paper are the initialisation of the pyramid by a Delaunay graph and the extension of the results for Hopfield neural networks for decimation. The validity of our approach is demonstrated by several examples.

[PRIP-TR-025] Frank Davoine. Fractal Image Compression Based on Adaptive Tessellations. 1993.

**Abstract:** The principle of fractal image coding presented in this paper is based on the theory of L-IFS (Local Iterated Function Systems). The algorithm exploits the fact that a real-world image is formed approximately of transformed copies of parts of itself. Thus, the construction of fractal codes is directly made on partitions of the image support. It is based on piece-wise similarities between blocks of different sizes.

The paper starts with a regular block based approach first proposed by Jacquin. To improve the algorithm, adaptive partitions are proposed with, in particular, the Delaunay triangulation. The results show an improvement in computing times, compression ratios and visual quality of reconstructed images.

[PRIP-TR-026] Stefan Barth. LateX - Bilderkennung, Diplomarbeit. 1993.

**Abstract:** This Technical Report describes the goals of the Master thesis ‘ $\LaTeX$ -picture-recognition’. The main goal of the thesis, which is viewed as a contribution to the research topic ‘recognition of technical drawings’, is the automatic conversion of  $\LaTeX$ -drawings into  $\LaTeX$ -documents using the  $\LaTeX$  picture environment. Thus the process inverse to  $\LaTeX$  will be implemented.

Moreover, the basic idea for solving this problem and some methods capable of checking the correctness of the picture recognition are stated.

The report is given in German.

[PRIP-TR-027] Dieter Willersinn, Etienne Bertin, and Walter G. Kropatsch. Dual Irregular Voronoi Pyramids and Segmentation. 1994.

**Abstract:** We continue previous work about the combination of top-down and bottom-up adaptive segmentation techniques, Voronoi diagrams and irregular pyramids. We extend our considerations to the dual irregular pyramid to overcome the problem of increasing degree inherent to the “classical” irregular pyramid. Experimental results are presented, the analysis of which reveals inconsistencies in the theory of dual irregular pyramids. The conclusion of the report outlines two strategies for research in view of a solution.

[PRIP-TR-028] Dieter Willersinn. Parallel Graph Contraction for Dual Irregular Pyramids. 1994.

**Abstract:** Hierarchical representation of images is a crucial building principle of a system architecture that can cope with the complexity of visual perception. This paper presents an algorithm that builds a hierarchy which is flexible enough to provide a shift and scale invariant abstract representation of image content. The hierarchy is irregular in the sense that an element of the representation may have an arbitrary number of neighbors. At the same time, the representation can be described by a bounded data structure. Main properties of the algorithm with respect to visual perception are proved.

[PRIP-TR-029] Georg Duftschnid. Qualitaetskontrolle von Holzwoollplatten. 1994.

**Abstract:** Wooden fiber-plates, that are used in appearance-sensitive applications, have to be checked by aesthetical criterias. It will be discussed in this report, how defects on the surface of the plates, that disturb their optical appearance, can be found



by means of image processing. Three typical defect-types are examined. The resulting detection-process can be used as a step towards an automatic inspection system.

[PRIP-TR-030] Karin Hruby. Identifizieren von Gesichtern durch Steuerung der visuellen Aufmerksamkeit, Diplomarbeit. 1994.

**Abstract:** The goal of the diploma thesis is to develop a system that combines the eigenfaces introduced by Turk and Pentland and the template matching strategy of Brunelli and Poggio by a mechanism for focusing attention on different regions of the face. First the face is roughly classified at a coarse resolution, in a second step more information about the face is gained by focussing at different regions of the face at a higher resolution. The new information has to be included to get a new classification. After a general introduction into the field of face recognition the modules of the system and their interactions are explained.

[PRIP-TR-031] Elmar Thurner. Modulare Neurale Systeme, Aufgabenstellung der Diplomarbeit. 1994.

**Abstract:** During the last years neural networks have been applied to a wide range of pattern recognition applications. The reasons for this rapid growth of interest for neural networks are the innovative capabilities of neural networks: Neural networks are universal and adaptive function approximators. They provide self organization as well as distributed representation of knowledge. Despite of these advantages, the solutions found, suffered by new problems like the enormous effort in computational power to train the networks, bad scaling properties and insufficient extendability. In this report it is stated, that the reason for these disadvantages can be found in the architecture of the used networks. It will be shown, that a single general purpose network will not fit optimally to any given problem. The report gives an introduction to a new approach to overcome the mentioned problems: Modularity. Starting with basic definitions like modular and hierarchical system, a detailed analysis of the advantages of modular architectures in comparison to large, single networks is presented. Using these definitions the principal structures, found in modular neural architectures, like parallelism (integrative and competitive), cascades and supervisor actor structures are described. Finally an outlook to the goals of the diploma theses and a definition of performance criteria for neural networks for optical character recognition are given.

[PRIP-TR-034] Stephan Barth. Latex Bildererkennung. 1993.

**Abstract:** Within the scope of this diploma thesis a new algorithm for the recognition of LATEX-pictures, e.g. line drawings, was developed and implemented. The basic idea of this approach is the use of a regular  $2 \times 2/4$  pyramid for building up lines and arcs out of the huge mass of single edge points. In this manner the amount of data is reduced and the first step towards recognition of the image is done. The approximation of the edges with lines and arcs is accomplished by statistical methods. In addition to a detailed description of this algorithm, this diploma thesis gives a survey over the state-of-the-art in recognition of line drawings and a report on some experimental results.

[PRIP-TR-035] Walter G. Kropatsch. Building Irregular Pyramids by Dual Graph Contraction. 1994.

**Abstract:** Many image analysis tasks lead to or make use of graph structures that are related through the analysis process with the planar layout of a digital image. This

paper presents a theory that allows to build different types of hierarchies on top of such image graphs. The theory is based on the properties of a pair of dual image graphs that the reduction process should preserve, e.g. the structure of a particular input graph. The reduction process is controlled by decimation parameters, i.e. a selected subset of vertices, called survivors, and a selected subset of the graph's edges, the parent-child connections. It is formally shown that two phases of contractions transform a dual image graph to a dual image graph built by the surviving vertices. Phase one operates on the original (neighborhood) graph and eliminates all non-surviving vertices. Phase two operates on the dual (face) graph and eliminates all degenerated faces that have been created in phase one. The resulting graph preserves the structure of the survivors, it is minimal and unique with respect to the selected decimation parameters. The result is compared with two modified specifications, the one already in use for building stochastic and adaptive irregular pyramids.

[PRIP-TR-036] Martin Eder. Farbanalyse und Farbvergleich von Portaitminituren. 1994.

**Abstract:** The subject of this project is the development of a computer system to record and to classify miniportraits. A computer system should be able to register these works of art in order to analyse them objectively. Art historians should receive detailed support in making decisions about the classification of these miniportraits. This paper is about methods of modelling 3D-surfaces and about the analysis of colours concerning strokes in portraits. Each artist paints in its own characteristic way. These typical marks should be detected so that the miniportraits can be compared with each other.

[PRIP-TR-037] Paul Kammerer. Computer gestuetzte Klassifizierung von Portaitminituren. 1994.

**Abstract:** A computer supported classification of portrait miniatures is based on the detection of its primitives that are stroke and spot of a brush. A model of a stroke of a brush is defined. According to the model a method for the detection of primitives of portrait- miniatures is developed. The second part of the planned work deals with the determination of the primitives structure which is typical for an artist.

[PRIP-TR-038] Elmar Thurner. Modulare neuronale Systeme. 1995.

**Abstract:** This thesis gives an introduction and an overview to a new approach in the field of neural networks: Modularity. It will be shown, that a single general purpose network will not fit optimally to any given problem. Contrary, modular networks learn faster, have better generalization abilities and higher robustness and are easier to extend. To investigate the relationship between architecture and function the basic structures, found in modular neural architectures, like parallelism (integrative and competitive), cascades and supervisor actor structures are explored. Furthermore algorithms are presented to estimate the quality of results of networks, to combine the results of networks, to automatically decompose tasks and to combine different architectures. The theoretical analysed advantages of modular neural networks are demonstrated by experiments in the field of optical character recognition, OCR. For the recognition of printed characters in 20 different fonts by using modularity the training time is reduced to less than a quarter and the miss classification rate by one third in comparison to the single network solution. For the classification of hand written digits using modular neural preprocessing, an improvement of the miss classification rate of 7% is achieved compared to the non modular network.

[PRIP-TR-039] Thomas Nemec. Gesichtserkennungs-System zur Raumueberwachung. 1995.

**Abstract:** The goal of this diploma thesis is to develop a face recognition system that uses a CCD camera fixed on robot arm. In order to extend the “visual field” of the system the robot moves the camera in several directions. The task of the system is to identify people moving inside the observation room. After a generally introduction some methods of face recognition are explained. The structure of the system and its tasks are presented.

[PRIP-TR-040] Horst Bischof and Valentina Filova (Eds.). Theory and Applications of Digital Image Processing and Pattern Recognition. 1995.

**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.

[PRIP-TR-041] Thomas Nemec. Gesichtserkennungs-System zur Raumueberwachung. 1995.

**Abstract:** The goal of this diploma thesis is to develop a face recognition system that uses CCD camera fixed on a robot arm. In order to extend the “visual field” of the system the robot moves with the camera. The task of the system is to find people and to identify them. After a general introduction some methods of face recognition and detection are explained. The structure of the system and its tasks are presented. To evaluate the whole system a database consisting of people from the staff of the department ist created. The goal is then to check if a person entering the room belongs to the staff or not. In case the person belongs to the staff its name should be given.

The system consists of the following major modules: Motion detection, Face detection, and Face recognition. In particular, the following methods are used within the modules: Motion Energy detection for Motion detection, Multilayer Perceptrons for Face detection, and the Eigenface approach of Turk & Pentland for Face recognition.

After evaluating the individual modules we tested the face recognition system on different people entering the observation room. Provided one person stays app. 10 seconds in the observation room, the overall rate is 82,4 % which is a good result considering the speed of one frame per second on fully automatic face recognition.

[PRIP-TR-042] Walter G. Kropatsch. Equivalent Contraction Kernels and The Domain of Dual Irregular Pyramids. 1995.

**Abstract:** Dual graph contraction reduces the number of vertices and of edges of a pair of dual image graphs while, at the same time, the topological relations among the 'surviving' components are preserved. Repeated application produces a stack of successively smaller graphs: a pair of dual irregular pyramids. The process is controlled by selected decimation parameters which consist of a subset of surviving vertices and

associated contraction kernels. Equivalent contraction kernels (ECKs) combine two or more contraction kernels into one single contraction kernel which generates the same result in one single dual contraction. Decimation parameters of any individual pyramid level can be reconstructed from the ECK of the pyramid's apex if both vertices and edges of this ECK receive labels indicating their annihilation level in the pyramid. This is a labeled spanning tree (LST) of the base graph which allows efficient design and control of different types of dual irregular pyramids. Since the LST determines the pyramid, primitive modifications of the LST transform also pyramids into other pyramids on the same base graph. They open a large variety of possibilities to explore the domain of 'all' pyramids.

[PRIP-TR-043] Norbert Braendle. Rektifizierung oder direkte Tiefenberechnung? Untersuchungen zur digitalen Objektvermessung mit dem Stereoverfahren. 1996.

**Abstract:** Das Stereoverfahren ist als eine Methode zur Vermessung von dreidimensionalen Strukturen aus zweidimensionalen Bildern schon lange bekannt. Da die automatische, vom Rechner durchgefuehrte Stereoauswertung immer noch zahlreiche Probleme bereitet, insbesondere wenn es um die Bewaeltigung allgemeiner Szenen im dreidimensionalen Raum geht, sind Stereoalgorithmen weiterhin Gegenstand intensiver Forschungen. In den einschlaegigen Publikationen wird sehr oft eine einfache Kameraanordnung, der sogenannte Normalfall, angenommen. Da die geometrischen Bedingungen des Normalfalls in realen Aufnahmesystemen nicht zutreffen, werden durch perspektivische Rektifizierung Bilder erzeugt, die diese Bedingungen erfuellen. Eine Alternative ist der Verzicht auf die Vorteile der einfachen Geometrie und die Stereoauswertung auf den Originalbildern unter Beruecksichtigung der epipolaren Geometrie, was einen erhoeheten Rechenaufwand bedeutet und daher oft nicht in Erwaegung gezogen wird. Dieser Bericht behandelt das geometrische Kameramodell, die epipolare Geometrie einer Stereoanordnung, das Korrespondenzproblem, den Normalfall der Stereoauswertung und die perspektivische Rektifizierung von Stereobildern. Abschliessend werden die Ziele der Diplomarbeit vorgestellt.

[PRIP-TR-044] Roland Ebensberger. Robuste Eigenbildfunktionen. 1996.

**Abstract:** Die grundlegenden Einschränkungen der gegenwaertigen Eigenspace Methoden sind ihre Probleme mit gestoerten Daten und ueberdeckten Objekten. Am Institut fuer Mustererkennung wurde eine neue Methode entwickelt, die diese Probleme loest. Im Rahmen der Diplomarbeit soll diese Methode in ein bestehendes Softwarepaket fuer appearance-based matching (SLAM) integriert werden. Weiters soll die Koeffizientenberechnung in einem diskreten Raum ausgefuehrt werden. Die Methode soll auf ihre Robustheit und Berechnungskomplexitaet hin untersucht und mit der Standardmethode verglichen werden.

[PRIP-TR-045] Christian Menard. Robust Stereo using Correlation Scale Space. 1996.

**Abstract:** The stereo analysis method is similar to the human visual system. Due to the way our eyes are positioned and controlled, our brains usually receive similar images of a scene taken from nearby points of the same horizontal level. Therefore the relative position of the images of an object will differ in the two eyes. Our brains are capable of measuring this difference and thus estimating the depth. Stereo analysis tries to imitate this principle. This work contains two complementary and original contributions, one combines stereo techniques with robust statistics and the other solves the correspon-

dence problem in a multi-scale approach using correlation scale-space. Most standard algorithms make unrealistic assumptions about noise distributions, which leads to erroneous results that cannot be corrected in subsequent processing stages. In this work the standard area-based correlation approach is modified so that it can tolerate a significant number of outliers. The approach exhibits a robust behavior not only in the presence of mismatches but also in the case of depth discontinuities. Another central problem in stereo matching using correlation techniques lies in selecting the size of the search window. Small windows contain only a small number of data points, and thus are very sensitive to noise and therefore result in false matches. Whereas large search windows contain data from two or more different objects or surfaces, thus the estimated disparity is not accurate due to different projective distortions in the left and the right image. In this work a new method is proposed providing a continuous scale for the matching process, so that for each region in the stereo pair depending on the local information an optimal scale can be estimated. Results are given on synthetic images for the robust correlation technique. The adaptive matching method using correlation scale-space is tested on both synthetic and real images.

[PRIP-TR-046] Robert Sablatnig. A highly Adaptable Concept for Visual Inspection. 1997.

**Abstract:** To be acceptable in industry, vision systems must be inexpensive, within the speed of the production-line flow, and very accurate. While visual inspection is high in potential, at present the design and implementation of automatic visual inspection systems is labor-intensive. In addition, most of the visual inspection systems are developed in isolation with no systematic approach. Increasing flexibility to allow the inspection of parts whose positions are less constrained is desirable.

This work aims to show a systematic automated visual inspection concept that separates the detection of primitives from the model-based analysis process. This separation is obtained by defining a general analysis graph for inspection, containing detail relations that represent detection algorithms. Together with an object-specific description, defined in a so-called description language, the analysis graph is instantiated. Existing pattern recognition software is re-used in the detection stage and therefore the use of any detection algorithm is possible without changing the analysis. The concept can be seen as a “recipe” for solving industrial applications, stating which kind of decision have to be made at which stage.

An industrial application of the concept, for which industrial constraints have to be considered, is shown in the example of an automated visual inspection system for analogue watermeters used for calibration. Results concerning time, accuracy, and reliability of the specific inspection task are given. The flexibility of the concept is demonstrated by testing the analysis process with the description of other instruments (a hygrometer and a clock), which is performed by adapting the analysis graph but without changing the detection algorithms.

[PRIP-TR-047] Ales Leonardis, Horst Bischof, and Roland Ebensberger. Robust Recognition using Eigenimages. 1997.

**Abstract:** The basic limitations of the current appearance-based matching methods using eigenimages are non-robust estimation of coefficients and inability to cope with problems related to outliers, occlusions, and segmentation. In this paper we present a new approach which successfully solves these problems. The major novelty of our

approach lies in the way how the coefficients of the eigenimages are determined. Instead of computing the coefficients by a projection of the data onto the eigenimages, we extract them by a hypothesize-and-test paradigm using subsets of image points. Competing hypotheses are then subject to a selection procedure based on the Minimum Description Length principle. The approach enables us not only to reject outliers and to deal with occlusions but also to simultaneously use multiple classes of eigenimages.

[PRIP-TR-048] Thomas Melzer. Adaptive Robotersteuerung mittels visueller rueckkopplung. 1997.

**Abstract:** An adaptive system for kinematic robot control based on visual feedback is presented. The system is capable of moving the effector of an industrial robot to the position of an target object, whose coordinates are extracted from a pair of stationary mounted CCD cameras. The adaptive component consists of an extended neural gas network, which will - without any prior knowledge about camera orientation or robot arm architecture - eventually learn the mapping from stereo image coordinates to associated robot joint angles, the so called hand eye transform. Following the discussion of self organizing systems and the description of the system components, the results of four software experiments are presented, which shall illustrate the impressive performance, but also some weaknesses of the extended neural gas model. Finally, the results of an experiment conducted in a real hardware environment are presented.

[PRIP-TR-049] Thomas Melzer. TROL TU Vienna Robot Control library. 1998.

**Abstract:** TROL is a C-function library which gives the programmer access to virtually all kinematic functions of a CRS A465 industrial robot by providing him with a comfortable, high level Applications Programmers Interface. This document describes how to install, use and extend the library. Furthermore, it gives detailed insight in the design and implementation of TROL.

[PRIP-TR-050] Michael Dangl. Three-dimensional object reconstruction form layered spatial data. 1998.

**Abstract:** In this work, an approach for reconstructing 3D objects from layered spatial data points is presented. These data points are sampled at the surface of a 3D object using a 3D scanning system. In the first part, two algorithms for reconstructing triangle meshes from scanned objects are compared against each other showing different behavior concerning the quality and the number of reproduced triangles of the approximated model. Second, in order to allow fast visualization of the geometry, the mesh is optimized in respect of its number of triangles in a way that preserves the contour of the object.

[PRIP-TR-051] Christian Liska. Robuste und Adaptive Bildgewinnung mittels Laserlicht. 1998.

**Abstract:** Dieser Report spezifiziert die gestellte Problemstellung der Diplomarbeit. Es wird speziell die Anordnung der Hardware zur dreidimensionalen Oberflhenrekonstruktion mittels strukturiertem Licht, sowie die Kalibrierung dieser Hardware beschrieben. Weiters wird eine Einfhruung in Verfahren zur Planung und Berechnung der nhesten Kameraaufnahme-position zum Objekt, sogenannte Next-View-Planning-Techniken, gegeben. Abschliend werden die geplanten Ziele der Diplomarbeit definiert und ein grober Zeitplan zur Realisierung dieser Diplomarbeit angegeben.

[PRIP-TR-052] Martin Kampel. Tiefendatenregistrierung von rotationssymmetrischen Objekten. 1998.

**Abstract:** The aim of the master thesis is to develop an algorithm for registration of two different views of rotational symmetric objects. The method proposed uses the axis of rotation of the viewed objects to bring multiple range images into alignment. The task of the system is the construction of 3d models of archaeological finds. After a general introduction including the 3d acquisition technique different methods of registration are explained. The structure of the system and its tasks are presented.

[PRIP-TR-053] Christian Liska. Profilschnittermittlung und bildhafte Erfassung archaeologischer Fundscherben mittels Laserlicht. 1998.

**Abstract:** Profilschnitte dienen Archaeologen als wichtige Entscheidungsgrundlage zur Klassifikation von Fundstuecken. In dieser Arbeit wird ein System vorgestellt, das dem Archaeologen die Arbeit der manuellen Erfassung dieser Funde abnimmt. Neben der Vermittlung der Grundlagen zur Bildgewinnung wird die Erfassung der Oberflaechencharakteristik mittels strukturiertem Licht, speziell der Lichtstreifenprojektion beschrieben und an die Aufgabenstellung der Profilschnittermittlung angepasst. Das Verfahren wird weiters um eine Komponente zur Erfassung der sichtbaren Objektoberflaeche erweitert.

[PRIP-TR-054] Luc Brun and Walter Kropatsch. Dual Contraction of Combinatorial Maps. 1999.

**Abstract:** This paper presents a new formalism for irregular pyramids based on combinatorial maps. The combinatorial map formalism allows us to encode a planar graph thanks to two permutations encoding the edges and the vertices of the graph. The combinatorial map formalism encode explicitly the orientation of the planar graph. This last property is useful to describe the partitions of an image which may be considered as a subset of the oriented plane  $R^2$ . This new constraint allows us to design interesting properties for irregular pyramids. Finally the combinatorial formalism allows us to encode efficiently the graph transformations used in irregular pyramids.

[PRIP-TR-055] Christian Liska. Das adaptive Lichtschnittverfahren zur Oberflaechenrekonstruktion mittels laserlicht. 1999.

**Abstract:** This thesis describes a system, which allows a dynamic acquisition of the viewable surface of an object. Starting from the mathematic and geometric modeling of the acquisition process, an acquisition system is developed and the calibration process of the equipment used, consisting a turntable, a CCD camera and two laserdiodes, is described. After this, the surface reconstruction and the determination of the real 3d object coordinates out of their 2d projection in the camera plane is derived. Based on this fundamentals, a Next View Planning Technique is motivated and an adaptive algorithm is specified. The thesis concludes with expermental results and an outlook on future work.

[PRIP-TR-056] Martin Kampel. Tiefendatenregistrierung von rotationssymmetrischen Objekten. 1999.

**Abstract:** The aim of the master thesis is to develop an algorithm for registration of two different views of rotationally symmetric objects. The method proposed uses the axis of rotation of the viewed objects to bring multiple range images into alignment.

The task of the system is the construction of 3d models of archaeological finds. After a general introduction the 3d acquisition technique used and the calibration of the acquisition system is explained. Then the estimation of the axis of rotation is presented. After an overview about existing registration techniques the registration method used is illustrated. Finally results of four different experiments show the accuracy of the demonstrated registration algorithm.

[PRIP-TR-057] Luc Brun and Walter Kropatsch. Pyramids with Combinatorial Maps. 1999.  
**Abstract:** This paper presents a new formalism for irregular pyramids based on combinatorial maps. This technical report continues the work begun with the TR-54 report. Definition and properties of Contraction kernels are generalized and completed. The definition and properties of Equivalent contraction kernels are also given.

[PRIP-TR-058] Norbert Braendle, Hilmar Lapp, and Horst Bischof. Fully automatic grid fitting for genetic spot array images containing guide spots. 2000.  
**Abstract:** In the domain of biotechnology array-based methods are used to gain rapid access to genetic information based on the signals of the individual array elements (spots). For an automated analysis of the spots it is necessary to fit a grid to the spots in the digital image in order to represent the array distortions that may occur in the course of the experiment. In order to make the grid fitting problem tractable in a certain class of experiments spot arrays contain a sub-grid of guide spots with a known signal characteristic. We present an automatic grid fitting method for spot array images containing guide spots. Our approach uses simple image processing methods and takes into account prior knowledge inherent in the imaging process.

[PRIP-TR-059] Srdan Tosovic. Lineare Hough-Transformation und Drehtellerkalibrierung. 1999.  
**Abstract:** In dieser Arbeit wird ein Verfahren zur Kalibrierung eines Aufnahmesystems vorgestellt, das als Ergebnis die Transformationsmatrix liefert, die die Bildkoordinaten des Aufnahmeobjekts in die Objektkoordinaten umwandelt. Dabei besteht das Aufnahmesystem aus einem Drehteller und einer Kamera, deren optische Achse in der Drehtellerebene liegt. Da das Verfahren sehr stark auf der linearen Hough-Transformation beruht, wird die lineare Hough-Transformation im ersten Teil dieser Arbeit beschrieben. Danach erfolgt eine Beschreibung des Kalibrierungsverfahrens.

[PRIP-TR-060] Rémi Mégret and Caterina Saraceno. Building the background mosaic of an image sequence. 1999.  
**Abstract:** Images sequences present a high degree of redundancy because objects are repeated over the successive images. When their apparent displacements are well approximated by a simple parametric model, the whole sequence can be summed up by pasting together all the images onto a so called mosaic image. Then each image of the original sequence can be considered as a part of the final mosaic. When the displacements of distinct objects differ, we must choose which objects have to be represented and how.

In this report a framework is presented that produces the mosaic image corresponding to the background object of an image sequence. It is based on the dominant motion assumption, that is the background motion is parametric and the background occupies



the main part of the images. The foreground objects are localised by their different motion. This localisation is computed together with the background motion in an iterative method. The regions corresponding to the background are then pasted onto the mosaic image using classic methods adapted to our problem or a new pasting method based on the distance to the foreground objects that achieve clearer mosaics.

[PRIP-TR-061] Christian Wolf. Content based Image Retrieval using Interest Points and Texture Features. 2000.

**Abstract:** Content based image retrieval is the task of searching images from a database, which are visually similar to a given example image. Since there is no general definition for visual similarity, there are different possible ways to query for visual content. In this work we present methods for content based image retrieval based on texture similarity using interest points and Gabor features. Interest point detectors are used in computer vision to detect image points with special properties, which can be geometric (corners) or non-geometric (contrast etc.). Gabor functions and Gabor filters are regarded as excellent tools for texture feature extraction and texture segmentation. We present methods how to combine these methods for content based image retrieval and to generate a texture description of images. Special emphasis is devoted to distance measures for the texture descriptions. Experimental results of the query system on different test image databases are given.

[PRIP-TR-062] Alexander Selb. Modellselektion von Clusteringverfahren mittels minimaler Beschreibungslaenge. 2000.

**Abstract:** This diploma thesis presents the use of the *minimum description length* (*MDL*) principle for unsupervised clustering. Two frameworks are presented which determine the right size of neural networks used for *crisp/hard* and *fuzzy* clustering purposes. The optimization procedure starts with a network consisting of a large number of network nodes and iteratively reduces the net size until a complexity criterion is met. The *MDL* principle is used to measure the complexity of the network. Different instantiations of the *MDL* algorithm, depending on the coding procedure, are possible and derived in the thesis. In addition, we demonstrate how the *MDL* principle increases the robustness of the training algorithm by providing a criterion for *outlier* detection. Also a *growing* component based on *MDL* is presented to cope with the formation of new data clusters due to non-stationary data. In order to overcome the problem of initialization the *growing neural gas* (*GNG*) algorithm is used which thereby also increases the computational efficiency of the *MDL* algorithm. Furthermore, we present additional methods to increase the computational efficiency. Typical applications, like 2 dimensional clustering, data compression, *RGB* color image segmentation and learning the *hand-eye* coordination for robot control, are used to demonstrate the behaviour of the *MDL* algorithm.

[PRIP-TR-063] Luc Brun and Walter Kropatsch. The Construction of Pyramids with Combinatorial Maps. 2000.

**Abstract:** This paper presents a new formalism for irregular pyramids based on combinatorial maps. This technical report continues the work begun with the PRIP-TR-54 and PRIP-TR-57 reports. We provide in this technical report algorithms allowing efficient parallel or sequential implementation of combinatorial pyramids.

[PRIP-TR-064] Srdan Tosovic. Shape from Silhouette. 2000.

**Abstract:** An algorithm for automatic construction of a 3D model of an object is presented. The construction is based on a sequence of images of the object taken from different viewpoints. The object's silhouette is the only feature which is extracted from an input image. Images are acquired by rotating the object on a turntable in front of a stationary camera. The algorithm uses an octree representation of the model, and builds this model incrementally, by performing limited processing of all input images for each level of the octree. Beginning from the root node at the level 0 a rough model of the object is obtained quickly and is refined as the processed level of the octree increases. Results of the algorithm developed are presented for both synthetic and real input images.

[PRIP-TR-065] Thomas Melzer, Michael Reiter, and Horst Bischof. Kernel Canonical Correlation Analysis. 2001.

**Abstract:** This paper introduces a new non-linear feature extraction technique based on *Canonical Correlation Analysis* (CCA) with applications in regression and object recognition. The non-linear transformation of the input data is performed using kernel-methods. Although, in this respect, our approach is similar to other *generalized* linear methods like kernel-PCA, our method is especially well suited for relating two sets of measurements. The benefits of our method compared to standard feature extraction methods based on PCA will be illustrated with several experiments from the field of object recognition and pose estimation.

[PRIP-TR-066] Srdan Tosovic. Construction of 3D Models of Objects Using Combination of Shape from Silhouette and Shape from Structured Light. 2001.

**Abstract:** This work proposes a method of three-dimensional reconstruction of objects using a combination of two existing methods, Shape from Silhouette and Shape from Structured Light. The purpose of combining these two methods is to overcome the disadvantages of one algorithm by applying the other and construct a 3D model of an object which will be more accurate and precise than a model acquired by applying one of the algorithms only.

[PRIP-TR-067] Georg Langs, Horst Bischof, and Walter G. Kropatsch. Irregular Eigenimage Pyramids and Robust Appearance- Based Object Recognition. 2002.

**Abstract:** In this report we deal with the possibilities to improve the performance of robust PCA algorithms by replacing uniform subsampling of images by an irregular image pyramid. The image pyramid is built based on knowledge gained from the training set of images. It represents different regions of the image with different level of detail, depending on their importance for the reconstruction process. As this is a technical report, not only the final and optimal results will be presented, but also some of the previous steps will be explained, even if they were abandoned later. The main reason is to support future work on the topic.

[PRIP-TR-068] Srdan Tosovic. Adaptive 3D Models of Objects by Combining Shape from Silhouette and Shape from Structued Light. 2002.

**Abstract:** This thesis proposes a method of three-dimensional reconstruction of objects using a combination of two different methods, Shape from Silhouette and Shape from Structured Light, focusing on reconstruction of archaeological vessels. Shape from

Silhouette is a method suitable for reconstruction of objects with handles, whereas it is unable to reconstruct concavities on an object's surface, such as inside of a bowl. Shape from Structured Light can reconstruct such concavities, but it often creates incomplete models because of camera and light occlusions. The purpose of combining these two methods is to overcome the weaknesses of one method through the strengths of the other, making it possible to construct complete models of arbitrarily shaped objects. The construction is based on multiple views of an object using a turntable in front of stationary cameras. The method is adaptive, because it automatically selects a subset of possible views, guided by the complexity of the object modeled. Results of the algorithm developed are presented for both synthetic and real objects.

[PRIP-TR-069] Maamar Saib, Yll Haxhimusa, and Roland Glantz. Building Irregular Graph Pyramid Using Dual Graph Contraction. 2002.

**Abstract:** In this technical report the new version of the software Dgc tool is presented. This tool allows us to build up irregular graph pyramids by dual graph contraction. The graph pyramid consists of a stack of levels (pair of graphs), each of which has a primal level and its dual. Every successive level is a reduced version of the level below. Primal level and its dual represent a primal graph and its dual, respectively. The primal graph base level of the pyramid may represent a two dimensional image.

[PRIP-TR-070] Norbert Braendle. Robust Analysis of Spot Array Images. 2002.

**Abstract:** Computer-aided image analysis deals with the automatic recovery of visual information from digital images. Recent image analysis research is trying to find more general and more efficient algorithms. Given a digital image with a certain contents or certain problem domain, it is still mandatory to manually evaluate different approaches and processing sequences to extract useful and plausible information from the image. Once a generic image analysis approach for the problem domain has been found, the method can often be applied independent of the operator's intuition and previous experience. Optical character recognition (OCR) is an example for a successful development from initial research to off-the-shelf products. This work provides a generic image analysis system for a class of images having a characteristic contents denoted as *spot arrays*. Spots are defined as simply connected, irregularly shaped regions lighter (or darker) than their background. Representatives of this image class include images of Braille paper for blind persons and DNA arrays - a tool of modern biotechnology. Analysis of spot array image has three main tasks. The first task is to detect the spots present in the image and therefore deals with the spatial localization process. The spots in the image are located on a grid which may be distorted in the course of the image production process. The second task therefore deals with the fitting of a grid to the detected spots, such that they can be correctly addressed. Once the spots are detected and addressed, they are characterized by their shape, intensity and local background. The automatic image analysis presented in this work is composed of a set of tools arranged in a general framework. This general framework enables to analyze spot arrays of high spot density with possibly multiple overlapping spots. Furthermore, the concept is robust in order to cope with outliers in the spot array and artifacts like image contaminations. These requirements can be fulfilled by robust statistical models: A key principle of grid fitting is to fit straight line models to the rows and columns of the grid. The input for the straight line models is given by a maximum search in matched filter response. Spot

characterization is performed by fitting a parametric spot model to the corresponding pixels with the help of a robust M-estimator. In a consecutive step, a semi-parametric fit is possible in order to cope with deviations from the spot model assumptions. Analysis of DNA array images serves as a demonstration of the presented general framework. Here, the intensity of a spot represents the amount of genetic material bound to the corresponding array element. The ultimate image analysis goal of this application is to quantify as exactly as possible the intensity of tens of thousands of possibly overlapping spots. The output of DNA array image analysis yields the raw data for the discovery of specific genes and the genetic control system of organisms. The results of DNA array images demonstrate the successful application of the framework presented in this work on thousands of images resulting from various biological experiments.

[PRIP-TR-071] Roman Pflugfelder. Visual Traffic Surveillance Using Real-time Tracking. 2002.

**Abstract:** This report presents an unusual event recognition approach in the field of traffic surveillance. Such events are unusual traffic behaviour like traffic jams, accidents or ghost drivers. An interest-point based tracking algorithm (KLT-tracker) is discussed which pursues features on vehicles through a static camera scene. Tracking data can be collected by observing normal traffic. Then, this data is used to learn a spatio-temporal model of normal traffic behaviour. Thereby, training samples are generated in a learning space by the tracking data. Thus, the spherical probability density function (p.d.f.) of the space can be estimated. We use a Growing Neural Gas in combination with a MDL-based pruning algorithm for unsupervised learning. The former method belongs to the class of soft-competitive algorithms which overcome the problems of stranded reference vectors. In contrast to other works, the number of reference vectors has not to be constant. The algorithm finds an optimal codebook according to the MDL-principle. As the p.d.f. only describes points and not trajectories of normal traffic behaviour, behaviour classes of normal traffic have to be learnt additionally. This work presents a novel approach by using the topology of the learning space which is created by Competitive Hebbian Learning. Beside the necessity of recognizing unusual events, it can also be used to analyze the behaviour of drivers at traffic sites like intersections or road works.

[PRIP-TR-072] Horst Wildenauer and Walter Kropatsch. Computer Vision Winter Workshop. 2002.

**Abstract:** This technical report collects the papers presented at the Computer Vision Winter Workshop 2002 (CVWW 02) held 4-7 February 2002, in Bad Aussee, Austria. The workshop was organized by the Pattern Recognition and Image Processing Group, Vienna University of Technology. Each paper submitted to this workshop was reviewed by two members of the Programm Committee, and I would like to take the opportunity to express my sincere appreciation for their services and timely efforts. I would also like to thank Beatrix Forsthuber and Christa Kropatsch for the excellent preparation of the workshop. Finally, my thanks go to all the authors for their contributions.

[PRIP-TR-073] Roman Pflugfelder. Visuelle Verkehrsueberwachung in intelligenten Transportsystemen. 2002.

**Abstract:** Dieser Bericht fasst den Stand visueller Sensoren zur Verkehrsueberwachung zusammen. Visuelle Sensoren sind moderne Bildverarbeitungssysteme, die Verkehrs-

formation aus Videobildern ermitteln koennen. Diese Information ist die Basis moderner intelligenter Transportsysteme (ITS). Der Bericht gibt eine Einfuehrung in ITS. Dann werden alternative Sensortechnologien diskutiert und ihre Staerken und Schwaechen analysiert. Vergleichend dazu werden visuelle Sensoren bezueglich Technologie, Vorteile bzw. Nachteile und moegliche Anwendungsgebiete hin besprochen. Der Bericht schliesst mit einer Vorstellung von kommerziellen Produkten ab.

[PRIP-TR-074] Yll Haxhimusa and Walter G. Kropatsch. Reduction Factors of Pyramids on Undirected and Directed Graphs. 2002.

**Abstract:** We present two new methods to determine contraction kernels for the construction of graph pyramids. The first method is restricted to undirected graphs and yields a reduction factor of at least. This means with our method the number of vertices in the subgraph induced by any set of contractible edges reduces to half or less by a single parallel contraction. Our second method also works for directed graphs. Our methods yield better reduction factors than Meer's stochastic decimation algorithm, in all tests. The lower bound of the reduction factor becomes crucial with large images. We also give a method to compare the structure of the image pyramid.

[PRIP-TR-075] Ahmed Nabil Belbachir and Horst Bischof. On board Data Compression: Distortion and Complexity Related Aspects. 2003.

**Abstract:** Digital signal Processing (DSP), and in particular data analysis and compression, has been studied for many years. However, only the recent advances in computing technology have made it possible to use DSP in day-to-day applications. Users expect the data to be transmitted in a minimum of time and to take up as little storage space as possible. These requirements call for efficient data compression algorithms in terms of results quality and algorithmic complexity. The users want a good data quality with very fast compression and decompression as not to have to wait for data to be usable. Therefore, an accurate investigation on compression algorithms performances has to be performed. The performance of an algorithm can be analyzed using two criteria: the result quality i.e. high Signal-to-Noise Ratio vs Compression ratio and the algorithmic complexity. This report addresses both aspects of compression. In the first part of the report, noise influence on the compression method is stressed. The notion of algorithmic complexity is formalized using both approaches qualitative and quantitative in the second part of the report. In the third part, a new concept 'Integrated data compression' is introduced. In this part, we provide an optimization of On-board compression algorithm in terms of algorithm complexity and results quality, introducing a distributed exploitation of the data in both sides remote and user side. Application results of the method for the case study 'HERSCHEL/PACS Infrared Camera and Spectrometer' are presented in another paper.

[PRIP-TR-076] Jocelyn Marchadier. Functional Graph Models. 2003.

**Abstract:** Functional models are frequently used in computer vision and photogrammetry, as they enable the mathematical formulation of several problems such as pose computation and more generally the parameter estimation problem. However, the structural properties of such models have only seldom been studied. This contribution is dedicated to the analysis of such properties. We propose a new formalism that enables the analysis and design of functional models.

[PRIP-TR-077] Allan Hanbury and Jean Serra. A 3D-polar Coordinate Colour Representation Suitable for Image Analysis. 2002.

**Abstract:** The processing and analysis of colour images has become an important area of study and application. The representation of the RGB colour space in 3D-polar coordinates (hue, saturation and brightness) can sometimes simplify this task by revealing characteristics not visible in the rectangular coordinate representation. The literature describes many such spaces (HLS, HSV, etc.), but many of them, having been developed for computer graphics applications, are unsuited to image processing and analysis tasks. We describe the flaws present in these colour spaces, and present three prerequisites for 3D-polar coordinate colour spaces well-suited to image processing and analysis. We then derive 3D-polar coordinate representations which satisfy the prerequisites, namely a space based on the L1-norm which has efficient linear transform functions to and from the RGB space; and an improved HLS (IHLS) space. The most important property of this latter space is a well-behaved saturation coordinate which, in contrast to commonly used ones, always has a small numerical value for near-achromatic colours, and is completely independent of the brightness function. Three applications taking advantage of the good properties of the IHLS space are described: the calculation of a saturation-weighted hue mean and of saturation-weighted hue histograms, and feature extraction using mathematical morphology.

[PRIP-TR-078] Yll Haxhimusa and Walter G. Kropatsch. Experimental Results of MIS, MIES, MIDES and D3P. 2003.

**Abstract:** In this technical report we present in detail the results of the first 100 experiments of stochastic irregular graph pyramid of 100 100 and 200 200 images i.e graphs using methods MIS, MIES, MIDES and D3P. For details about these methods and irregular images pyramid see Technical Report PRIP-TR-74 [HK02]. This report extends PRIP-TR-74.

[PRIP-TR-079] Georg Langs, Horst Bischof, and Philipp L. Peloschek. Automatic Quantification of Destructive Changes Caused by Rheumatoid arthritis. 2003.

**Abstract:** Rheumatoid arthritis is an incurable disease affecting predominantly peripheral joints of the appendicular skeleton. It can lead to severe disabling mutilations and even to the complete destruction of the joints. The accurate and reproducible quantification of the progression of the disease and of the destructive changes caused to the joints is a decisive factor during therapy and during clinical trials. The manual quantification methods are time consuming and lack accuracy as well as reproducibility. In this thesis an alternative approach to automatically quantify the destructive changes caused by rheumatoid arthritis is proposed. Based on a hand radiograph the positions of the bones and joints are determined by local linear mapping nets. They learn the visual appearance as well as the anatomical structure of the hand during a training phase. The ability to learn makes a straightforward transfer of the method to other anatomical structures possible. Based on the coarse position estimates of the bones, the contour is identified with active shape models and snakes. When these methods are combined in the ASM driven snakes algorithm, a control of a priori knowledge utilized during the search for the pathologically changed contour is possible. The resulting description and visual information of the bone contour and its surroundings are used for point wise classification of the contour with respect to the question whether or not individual points

are affected by rheumatoid arthritis. The automatic determination of the extent of erosions in the joint region allows for an accurate and operator independent quantification of the disease progression.

[PRIP-TR-080] Robert Sablatnig. Shape Based Machine Vision. 2003.

**Abstract:** The study of visual object recognition is often motivated by the problem of recognizing 3-d objects given that we receive 2-d patterns of light on our retinæ. Recent findings from human psychophysics, neurophysiology and computational vision provide converging evidence for a view-based recognition framework in which objects and scenes are represented as collections of viewpoint-specific local features rather than 2-d templates or 3-d models. Hence the recent decade saw a gradual shift away from the 3-d object reconstruction approach pioneered by Marr toward view-based approaches. This report summarizes our contributions to this problem where we focus on the shape as recognition feature and apply these findings in the area of Machine Vision. The first part presents an overview of the framework, motivates the view-based recognition strategy, and introduces the hierarchical matching concept. Next, a short summary of a collection of six representative publications of our work carried out in this field, and a discussion of how this fits into the framework is given. The second part consists of the six papers themselves, where we start with a paper on the general framework which is followed by three different applications of the framework in Visual Inspection, Archaeology and Art History. The remaining two papers describe recent work performed in 3-d vision as part of the object-based recognition concept. The first paper is on the registration of range data, in which we propose a novel technique for range image registration. The collection ends with a work on combining different 3-d acquisition techniques within the hierarchical framework.

[PRIP-TR-081] Yll Haxhimusa and Walter G. Kropatsch. Hierarchical Image Partitioning with Dual Graph Contraction. 2003.

**Abstract:** We present a hierarchical partitioning of images using a pairwise similarity function on a graph-based representation of an image. This function measures the difference along the boundary of two components relative to a measure of differences of the components internal differences. This definition tries to encapsulate the intuitive notion of contrast. Two components are merged if there is a low-cost connection between them. Each component's internal difference is represented by the maximum edge weight of its minimum spanning tree. External differences are the smallest weight of edges connecting components. We use this idea for building a minimum spanning tree to find region borders quickly and effortlessly in a bottom-up way, based on local differences in a specific feature.

[PRIP-TR-082] Luc Brun and Walter Kropatsch. Labeled Pyramids with Combinatorial Maps. 2003.

**Abstract:** Combinatorial Pyramids are defined as a stack of successively reduced combinatorial maps. The Pyramid construction plan defined in TR-63 allows to describe a pyramid by two functions level and state defined respectively on the set of darts of the initial combinatorial map and the set of levels of the pyramid. These two functions encode respectively the maximum level on which a dart survives and the type of each reduction operation. Based on these functions any combinatorial map of the pyramid may be built from the base by a one pass algorithm scanning all the darts of the ini-

tial combinatorial map. In this technical report we show that algorithms with a same sequential and parallel complexity may be designed in order to build all the reduced combinatorial maps of the Pyramid.

[PRIP-TR-083] Hubert Mara. Automated Profile Extraction of Archaeological Fragments. 2003.

**Abstract:** Thousands of fragments of ceramics are found at archaeological excavation sites. Till today archaeologists have drawn and classified them manually. This method is very time consuming and classification depends on the experiences of the archaeologists. Therefore we developed a system that speeds up this process by using a 3D-scanner for the acquisition and a software that generates a registered 3D-model of the sherd. The features for classification used by archaeologists and a 3D reconstruction of the unbroken vessel are estimated automatically. The registration of different views of the sherd are based on the estimation of the rotational axis by a Hough inspired method. The classification and reconstruction is done by extraction of the longest profile line, which is an intersection of the sherd along the rotational axis of the unbroken vessel. The extracted features for classification are diameters, heights and their relation to each other. These features are estimated by the use of extremal points of the profile line. Results of the system developed are presented for both synthetic and real input data.

[PRIP-TR-084] Physics based Segmentation of Colour Images in Spherical Coordinates. Allan hanbury. 2003.

**Abstract:** We present image segmentation and highlight detection algorithms based on the dichromatic reflection model. For image segmentation, we use the model prediction that objects of a certain colour produce lines (the matte lines) radiating away from the origin of the RGB colour space. These lines therefore show up as peaks in a 2-dimensional histogram of the angular coordinates of a spherical polar coordinate representation of the RGB space. An algorithm for automatically locating these peaks is suggested. When the matte line locations are known, one can define cylindrical polar coordinate systems having their  $z$ -axes centred on the matte lines. We suggest a Hough-based algorithm for the detection of highlight lines in 2D-histograms of the  $\rho$  and  $z$  coordinates of the cylindrical polar coordinate system. Examples of the results of applying these algorithms are given.

[PRIP-TR-085] Helmut Zollner. A Calibration Technique for CCD Cameras using Pose Estimation. 2003.

**Abstract:** In this report we present a new calibration technique for off-the-shelf CCD cameras and acquisition systems. The estimated camera model includes 6 extrinsic and up to 8 intrinsic model parameters including radial and tangential distortion. Also an inverse distortion function for correction of distortion effects can be estimated. Planar and non-planar objects can be used as previously measured calibration targets. The method is suitable for single image and multiple image input datasets. Tools for Mathematica<sup>TM</sup>, MATLAB<sup>TM</sup> and Khoros<sup>TM</sup> were developed and tested by the author at PRIP. In addition this report includes also a number of practical test results.

[PRIP-TR-086] Martin Kampel. 3D Mosaicing of Fractured Surfaces. 2003.

**Abstract:** A major obstacle to the wider use of 3D object reconstruction and modeling



is the extent of manual intervention needed. Such interventions are currently massive and exist throughout every phase of a 3D reconstruction project: collection of images, image management, establishment of sensor position and image orientation, extracting the geometric detail describing an object, merging geometric, texture and semantic data. This work aims to develop a solution for automated documentation of archaeological pottery, which also leads to a more complete 3D model out of multiple fragments. Generally the 3D reconstruction of arbitrary objects from their fragments can be regarded as a 3D puzzle. In order to solve it we identified the following main tasks: 3D data acquisition, orientation of the object, classification of the object and reconstruction. 3D acquisition with respect to archaeological requirements is described by four different methods, designed for the recording of fragments, complete vessels, profile sections and color. The range and pictorial information of the objects is the input for further classification and reconstruction. In the so-called documentation step the processing of the recorded data leads to orientation and the profile sections. The following classification step produces a systematic view and order of the material recorded and identifies possible candidates for subsequent fragment assembling. Reconstruction of pottery refers not only to the reconstruction of a pot from its fragments, but also to the reconstruction of a pot or fragment out of its profile section. This thesis describes a complete system for automated documentation and reconstruction of archaeological pottery. The main contributions are 3D scanning of pottery, pairwise registration of views, a scheme for automatic classification of pottery, and an approach for solving 3D jigsaw puzzles of fragmented surfaces. In order to evaluate the system, experiments and results are given on both synthetic and real data. The selected approaches are cross-checked with the associated archaeologists.

[PRIP-TR-087] Ahmed Nabil Belbachir, Ted Chilton, Marc Nunkesser, Sahbi Sidhom, and Georges Szajnowski. Image Compression using Hartley transform. 2003.

**Abstract:** The report presents a novel approach for image compression using the Hartley transform (HT). The Hartley transform has the advantage of solving the problem of phase wrapping from which the Fourier transform suffers. The magnitude and phase compression using this transformation (HT) have proved better performance than those of the Fourier Transform. Magnitude and phase were processed separately. The quantization of frequency samples in less bits has increased the compression ratio. Furthermore, the distributions used to generate the noise significantly influence the result. The lossy compression technique used seems not to degrade the image quality. A nonlinear filter for smoothing the resulting image would be suitable for image enhancement. In general, the overall compression ratio is acceptable it compresses to about 15–30 percent the size of the original image. A lossless compression technique could be performed additionally to increase the compression factor.

[PRIP-TR-088] Ernst Hirz, Ahmed Nabil Belbachir, and Robert Sablatnig. Software tool for herschel/pacs data decompression and analysis. 2004.

**Abstract:** This work is about the design and development of a standalone tool for the decompression of data from the Photodetector Array Camera and Spectrometer (PACS), one of the instruments housed inside the HERSCHEL Space Observatory (HSO). This is a part of an on-ground processing software package, with the purpose to provide an instrument for collecting, assembling, decompressing and later analyzing the

data fragments received from the telescope. The work is done in JAVA programming language for simplicity, portability, reliability and a distributed computing. Furthermore, JAVA is object-oriented. Object-oriented programming provides greater flexibility, modularity and reusability. Thus it is easy to maintain/extend the on-ground processing tool for different or further compression algorithms or general upgrades. Within this work, we developed a scientific software tool for the processing of the received data from the HSO/PACS instrument. This software tool is designed and tested for PACS data but not limited for astronomical applications. The on-ground processing concept can easily be adapted for other applications (Medical imaging, Telecommunication ...etc). Results from the evaluation of the software with real and virtual test data are given at the end of this report.

[PRIP-TR-089] Martin Lettner, Paul Kammerer, and Robert Sablatnig. Texture analysis of painted strokes. 2004.

**Abstract:** In this practical work texture analysis for painted strokes is reported. The work presents a study of stroke classification in which two classes of strokes are identified: fluid and dry strokes. The discrimination is done with a feature vector which is extracted from the stroke texture by the help of texture analysis methods. To find an adequate texture analysis method for this application, three different texture analysis methods are executed on test images from painted strokes. The methods applied are based on statistical features of first and second order and on the discrete wavelet transformation, whereas the statistical features of second order are extracted from the co-occurrence matrix. The results are compared and it turns out that the wavelet based texture analysis method yields the best discrimination rate for this application.

[PRIP-TR-090] Ahmed Nabil Belbachir and Florian Schmitzberger. On-board data compression of herschel-pacs control data. 2004.

**Abstract:** The aim of this work is to find a dedicated concept for efficient compression of Herschel-PACS control data. These data represent the setting parameters of the different instruments within PACS, which have to be transmitted with scientific data for diagnosis purpose. As these control data have a deterministic variability and because of the limited telemetry bandwidth, it was decided to lossless compress the control data for efficient exploitation of the downlink. In this report, an analysis of the control data has been made and a dedicated compression concept with different entropy coders has been developed and tested on simulated and real data. The obtained results have been evaluated and compared to a state-of-the-art codec.

[PRIP-TR-091] Michael Lienhardt. Morphological operations on surfaces. 2004.

**Abstract:** In this report, we propose an implementation of some basic morphological operations on triangulations of surfaces, in particular:

- dilation
- complement
- intersection, union, subtraction of two objects on the surface.

The definition of the distance and an object on the surface are the main contribution of this work. Some other functions for the visualisation of the scene are also described.

[PRIP-TR-092] Walter G. Kropatsch, Yll Haxhimusa, and Zygmunt Pizlo. Integral trees: Subtree depth and diameter. 2004.

**Abstract:** Regions in an image graph can be described by their spanning tree. A graph pyramid is a stack of image graphs at different granularities. Integral features capture important properties of these regions and the associated trees. We compute the depth of a rooted tree, its diameter and the center which becomes the root in the top-down decomposition of a region. The integral tree is an intermediate representation labeling each vertex of the tree with the integral feature(s) of the subtree. Parallel algorithms efficiently compute the integral trees for subtree depth and diameter enabling local decisions with global validity in subsequent top-down processes.

[PRIP-TR-093] Christian Asinger. Classification of color pigments in hyperspectral images. 2004.

**Abstract:** Image spectroscopy in ultraviolet, visible and infrared regions provides a non-invasive method for analyzing paintings. It is used by restorers and art-historians to get valuable information about works of art without causing any damage to them. The purpose of this report is to analyze 15 different color pigments and three types of binder by means of near infrared hyperspectral images taken from 15 testpanels. After calibration three feature reduction methods, principal component analysis, block-based principal component analysis and linear discriminant analysis, are applied to the near-IR spectroscopic imaging data of the testpanels to reduce the originally 180 features to improve classification performance. Both, principal component analysis and linear discriminant analysis, are linear transformations (in contrast to e.g. Kernel PCA) that map the original data into a new vector space where certain constraints are met. On the basis of the outcomes of the feature reduction step k-nearest neighbor classification with leave-one-out cross validation is used to accomplish three different classification tasks. These three tasks contain the identification of the three binders, the identification of the color pigments for each binder separately and the identification of color pigments and binders together. The results of our work reveal that it is possible to distinguish the three types of binder as well as the color pigments for each binder separately with an accuracy of more than 90 percent. Classifying the 15 color pigments along with the three binders however results in a lower accuracy of at most 45 percent.

[PRIP-TR-094] Sebastian Zambanini. Segmentation and surveying of cutaneous hemangiomas. 2005.

**Abstract:** This report presents an automatic method for the surveying of cutaneous hemangiomas by means of a fully automatic hemangioma segmentation and a ruler visible in the images to be examined. The segmentation of regions belonging to the hemangioma is accomplished with a pre-segmentation based on a perceptron followed by a postprocessing step where regions with untypical properties are rejected. The spatial resolution of the images is determined by computing the maximum euclidean distance between two marks of the ruler. Various experiments with the images gathered are reported, showing not only the satisfactory results on most of the images but also the problems arising on particular images.

[PRIP-TR-095] Harald Entner. Real-time 3d reconstruction for autonomous football playing robots using a feature based stereo approach. 2005.

**Abstract:** This diploma thesis addresses the topics of stereo 3D reconstruction and

shows the results of a feature based real-time stereo algorithm. For understanding issues, it uses the pinhole camera to describe perspective projection. In reality, a lens is used to project the light rays onto the sensor chip, thus an extended model is presented in order to describe the projection using a lens and to introduce camera calibration. In a dynamic environment stereo from motion is not possible, as a result a multi-camera system is needed. Epipolar geometry describes this special geometry and can be useful to minimize the computational cost of searching corresponding points in two images. The feature extraction uses the Canny algorithm followed by a new iterative line detection. Straight lines are considered to be an appropriate representation of the football field and the robots. After features have been found the correspondences are searched and used to calculate the 3D position of the features. Chapter 2 covers the theoretical background, whereas Chapter 3 describes the implemented methods. In Chapter 4 every part of the algorithm is evaluated. It is followed by the conclusion and the future plans.

[PRIP-TR-096] Ahmed Nabil Belbachir, Roland Ottensam, and Angela Baier. Glitch detection algorithms analysis for herschel-pacs spectroscopy data. 2005.

**Abstract:** This document presents a summary of algorithms, susceptible for glitch detection of Herschel-PACS spectroscopy ramps. An analysis of these algorithms is performed for preliminary consideration on PACS spectroscopy ramps. Test data have been provided by an astronomer from KU Leuven. Some of the proposed algorithms have been evaluated on these data and the same astronomer analysed the test results. These are promising results as well as further improvement of these algorithm is possible by including additional detector know-how in the deglitching mainly for determining the threshold for the breakdown choice.

[PRIP-TR-097] RenDonner, Georg Langs, Michael Reiter, Horst Bischof, and Robert Sablatnig. Fast parameter prediction for active appearance models using canonical correlation analysis - an aam matlab implementation. 2005.

**Abstract:** Active Appearance Models (AAM) provide a compact statistical model of data encompassing both shape and texture variations. This report introduces a novel and fast search algorithm for AAMs based on canonical correlation analysis (CCA). In contrast to the standard AAM matching approach CCA exploits the correlation between texture residuals and model parameters more efficiently. In a set of experiments using face and medical images we show that CCA based search consistently outperforms the convergence speed of the standard method by a factor of four. The time needed for training is reduced by 80% standard approach and CCA are similar our results suggest that CCA can replace the standard AAM search.

[PRIP-TR-098] Adrian Ion, Yll Haxhimusa, and Walter G. Kropatsch. A graph-based representation for spatiotemporal information in cognitive vision. 2005.

**Abstract:** A concept relating story-board description of video sequences with spatio-temporal hierarchies build by local contraction processes of spatio-temporal relations is presented. Object trajectories are curves in which their ends and junctions are identified. Junction points happen when two (or more) trajectories touch or cross each other, which we interpret as the “interaction” of two objects. Trajectory connections are interpreted as the high level descriptions.

[PRIP-TR-099] Lech Szumilas and Allan Hanbury. Segment feature co-occurrence based texture detection. 2005.

**Abstract:** This work on texture detection was inspired by the general problem of object recognition in two dimensional still images. One of the crucial challenges associated with the object recognition is selecting and obtaining discriminative features. In analysis of real scenes, like nature or urban places, many objects contain textures, which can be considered as one of the object features. Texture detection may also significantly improve image segmentation, which is one of the tools used for object recognition. This report presents a novel method named Feature Co-occurrence Texture Detector (FCTD) which allows for fully automatic detection of textures common in real scenes, like water in lakes, tiger skin, fields of flowers or tree crowns. The method searches for an alternating color pattern, like for example black and orange stripes covering tiger skin, which is very often present in those textures. The final result is a hierarchy of textures (described by their boundaries and a set of features) detected at multiple precision levels, which can be used for further analysis or texture classification. It is achieved through hierarchical clustering of color pairs related to adjacent image segments, where each segment represents a low color gradient, simple shaped patch of pixels in the image. The results are presented on some images from the Berkeley database.

[PRIP-TR-100] Florian Seitner. Robust detection and tracking of objects. 2005.

**Abstract:** Due to the increasing availability of fast and cheap hardware in the past few years, today a wide range of complex visual tracking tasks is possible. Efficient mathematical methods can provide a high robustness which also makes visual tracking interesting for many industrial purposes. However, the high demands on quality and speed still provide a major challenge for each tracking application. In this thesis a tracking system is introduced, which tries to address both demands appropriately by using currently available algorithms to quickly track pedestrians in video streams. By combining these well-proved algorithms, a good solution regarding computational complexity, accuracy and stability is obtained. To achieve this task, a fast object detector similar to the approach of Viola et al. [Viola 2003] is used as one component in this tracking system. This detector uses Haar-like features which are very fast to compute and makes a quick pedestrian detection in a frame possible. Next to the detection system, an adaptive background model sub-divides each frame into foreground and background regions. As a compromise between complexity and robustness a single-mode parametric background model based on normal distributions and wrapped normal distributions is used. Both background model and detector are combined to provide the tracking system with locations of pedestrian-like regions and to sub-divide the body into three parts: head, upper body and lower body. After this segmentation into finer tracking units a set of colour and spatial features for further tracking is extracted from each part individually. Individual and spatially separated body parts also provide the possibility to use colour histograms in a spatial sense. Moreover, an appearance model provides accurate solutions and approximations when occlusions or missing detections occur.

[PRIP-TR-101] Stefan Kuthan. Extraction of attributes, nature and context of images. 2005.

**Abstract:** In this thesis a framework for deriving high-level scene attributes from low-level image features is developed. Examples of attributes derived are photo-painting,

indoor-outdoor, night-day and nature-city. The assignment of the attributes to images is done by a hierarchical classification of the low level features, which capture colour, texture and spatial information. A concise summary of current research and methods used in this field of research is given. Furthermore, a prototype for image classification is implemented, which aids in the evaluation of the different methods available. Training and test images are provided by the ImagEVAL project, a French computer vision evaluation project.

[PRIP-TR-102] Allan Hanbury. Review of image annotation for the evaluation of computer vision algorithms. 2006.

**Abstract:** In the field of computer vision, automated image annotation and object recognition are currently important research topics. It is hoped that these will lead to improved general image understanding which can be usefully applied in Content-based Image Retrieval. Three approaches to image annotation are reviewed: free text annotation, keyword annotation and annotation based on ontologies. An analysis of the keywords which have been used in automated image and video annotation research and evaluation campaigns is then presented. The outcome of this analysis is a list of 525 keywords divided into 15 categories. Given that this list is collected from existing image annotations, it could be used to check the applicability of ontologies describing entities which are portrayable in images.

[PRIP-TR-103] Hurbert Mara. Documentation of rotationally symmetric archaeological finds by 3d shape estimation. 2006.

**Abstract:** Tens of thousands of fragments of ceramics (called sherds for short) are found at every archaeological excavation site and have to be documented for further archaeological research. The traditional documentation is based on the profile line, which is the intersection of the sherd along the axis of symmetry in the direction of the rotational axis. Traditionally this is done by experts by manually drawing the profile line, using different tools like a Profilkamm (profile comb), flexible wires, circle-templates, etc. to estimate the axis of rotation and the profile line. The traditional drawing is error prone and time consuming, therefore a semiautomatic method using a Profilograph was introduced to increase accuracy. Since the measurement is still manually, the time for drawing was not decreased. We propose a fully automatic system for the estimation of the rotational axis and the profile line. For data-acquisition we are using acquisition methods based on the principle of structured light, which have also been compared in respect to accuracy and performance to traditional methods of documentation. Based on continuous experiments and comparisons we show a new method for estimation of the rotational axis and the profile line, which is inspired by traditional archaeological methods. The methods shown in this thesis were tested on synthetic and real data. The experiments with real data were done at the archaeological excavation in Tel Dor in Israel. The results for estimation of the profile line and the comparison between the manual drawings, the Profilograph and the 3D-acquisition by structured light are shown in this thesis. Furthermore methodological experiments of geometrical surface analysis are shown, which demonstrate the possibility of estimation of ancient manufacturing techniques of ceramics. Finally an outlook towards detection of lines and analysis of painted ceramics is given.

[PRIP-TR-104] Irfan Adilovic. An audio filter framework for jop. 2006.

**Abstract:** The aim of this project is to develop a conceptual implementation of a filter framework for the Java Optimized Processor (JOP). The target system is a board with (among other things) a Cyclone FPGA and an AC'97 audio codec. The filter framework and the target system are not intended for industrial use, but rather for educational purposes about programming embedded systems in pure Java without foregoing any cross-platform capabilities of Java. The filter framework needs to run without changes on desktop systems (in a free JVM, like SUN's or IBM's) as well as on the targeted embedded system.

[PRIP-TR-105] Stefan Fiel and Paul Guerrero. Motion tracking with normalized cut and minimum spanning tree. 2006.

**Abstract:** We propose using the Normalized Cut method for motion tracking (J. Shi and J. Malik . "Motion Segmentation and Tracking Using Normalized Cuts". In IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(8), pages 888 - 905, 2000) on regions in frames that were identified by a Minimum Spanning Tree (MST) method in a pre-processing step. The purpose of the pre-processing step is to reduce the spatial resolution without losing important image information. An energy function, based on some selected properties of regions, is then calculated for each pair of regions in a fixed number of consecutive frames. This energy function represents the similarity of two regions. Based on this similarity, the Normalized Cut method is used to identify salient groups of regions. Finally, corresponding salient groups of regions of neighbouring sets of consecutive frames are found. We show on different experiments, how pre-segmentation can help to reduce computation time by reducing the spatial resolution of the input frames. We tested different methods to reduce temporal resolution to gain an additional speedup.

[PRIP-TR-106] Thomas Illetschko, Adrian Ion, Yll Haxhimusa, and Walter G. Kropatsch. Effective programming of combinatorial maps using coma - a c++ framework for combinatorial maps. 2006.

**Abstract:** Combinatorial maps and pyramids have been studied in great detail in the past, and it has been shown that this concept is advantageous for many applications in the field of image processing and pattern recognition by providing means to store information of the topological relations of the represented data. In the course of these studies, the properties of combinatorial maps have been investigated using different sets of permutations, different operations and different algorithms. In each case new software had to be created in order to conduct experiments, as the existing programs were designed to work only for a specific model. Due to the complexity of combinatorial maps, the implementation of such a software is a time and resource intensive task. Thus these programming efforts were often responsible for delaying the presentation of new results in the past. This paper presents COMA - a C++ framework for combinatorial maps - that has been created during recent studies of combinatorial maps, motivated by this problem. Using an object oriented approach, COMA was specifically designed to allow an efficient and quick integration of changes to the model of combinatorial maps used, as well as the implementation of new algorithms. As a consequence COMA significantly reduces the amount of time needed to set up new experiments.

[PRIP-TR-107] Thomas Flanitzer. The eccentricity transform (computation). 2006.

**Abstract:** The eccentricity of a vertex is the longest shortest distance to any other vertex in a graph. We introduce the eccentricity transform which calculates the eccentricity for every point in a graph. Applied to digital images it offers some interesting properties including invariance to articulated motion and robustness with respect to salt & pepper noise. Applied to graphs with an embedding it can be used for boundary determination. Its characteristics make it a good candidate for supporting or even replacing the distance transform as a basic tool in many feature extraction tasks (e.g. shape description). This report focuses on the computation of the eccentricity transform and explains implementation approaches.

[PRIP-TR-108] Julian Stoettinger. High performance profile line generation and visualization using opengl. 2007.

**Abstract:** On archaeological sites, thousands of fragments of ceramics are found. The classification of these findings is very time consuming and still done manually until today. To speed up this process, a 3d scanner for the data acquisition and software for the profile line generation is used. This work provides an end-user friendly user interface for profile line generation of archaeological findings. For performance reasons, the visualization is done by the GPU using OpenGL. 3d scan data with several hundreds of thousands of vertices can be processed in real time. The profile line generation is implemented in three levels: First, profile line generation is simulated by using OpenGL clipping planes and thus, the GPU only. Then, the nearest vertices of the intersection plane are marked and the intersection vertices are estimated. Finally, the profile segments of this diminished 3d scan data are decided. Comparisons to a Matlab implementation and a commercial software have been carried out and these experiments with synthetic and real data are shown. Finally, an outlook towards future development is given.

[PRIP-TR-109] Axel Pinz, Horst Bischof, Walter Kropatsch, Gerald Schweighofer, Yll Haxhimusa, Andreas Opelt, and Adrian Ion. Representations for cognitive vision: A review of appearance-based, spatio-temporal, and graph-based approaches. 2006.

**Abstract:** The emerging discipline of cognitive vision requires a proper representation of visual information including spatial and temporal relationships, scenes, events, semantics and context. The goal of this review article is to summarize existing representational schemes which might be useful for cognitive vision, and to discuss promising future research directions. We structure the various approaches into appearance-based, spatio-temporal and graph-based representations for cognitive vision. The representation of objects has been covered extensively in computer vision research, both from a reconstruction as well as from a recognition point of view. Cognitive vision, however, will also require new ideas how to represent scenes. We introduce new concepts for scene representations and discuss how these might be efficiently implemented in future cognitive vision systems..

[PRIP-TR-110] Thomas Illtischko. Minimal combinatorial maps for analyzing 3d data. 2006.

**Abstract:** Combinatorial maps and irregular pyramids based on combinatorial maps for 2D data have been studied in great detail. It has been shown that this concept is advantageous for many applications in the field of image processing and pattern recognition by providing means to store information of the topological relation of the



represented data. While the concept of combinatorial maps has been defined for any dimension, most of the studies concentrated on the representation of two dimensional data and only few results exist regarding higher dimensions. This report studies the properties of combinatorial maps for 3D data. Especially collapsing an initial map of the volumetric data by applying contraction and removal operations to produce a minimal representation while preserving the topological relations is presented in this report. Formal conditions for applying these operations as well as the minimal configurations of the topological relations found in volumetric data are presented in this report and means for discriminating and identifying these minimal configurations using pseudo elements are introduced.

[PRIP-TR-111] Samuel Peltier, Adrian Ion, Yll Haxhimusa, and Walter G. Kropatsch. Computing homology group generators of images using irregular graph pyramids. 2006.

**Abstract:** We introduce a method for computing homology groups and their generators of a 2D image, using a hierarchical structure i.e. an irregular graph pyramid. Instead of computing homology generators in the base where the number of entities (cells) is large, we first reduce the number of cells by a graph pyramid. Then homology generators are computed efficiently on the top level of the pyramid, since the number of cells is small, and a top down process is then used to deduce homology generators in any level of the pyramid, including the base level i.e. the initial image. We show that the new method produces valid homology generators and present some experimental results. In this report we also show that the generators of the first homology groups of a 2D image, computed with this pyramid based method always fit on the borders of the regions.

[PRIP-TR-112] Martin Stubenschrott, Walter Kropatsch, and Yll Haxhimusa. Combining an optical flow feature detector with graph-based segmentation. 2007.

**Abstract:** Object tracking is the complex task to follow a given object in a video stream. This paper describes an algorithm which combines an optical flow based feature tracker with color segmentation. The aim is to build a feature model and reconstruct lost feature points when they are lost due to occlusion or tracking errors. These feature points are tracked from one frame to another with the Lucas & Kanade optical flow algorithm. Additionally, we segment each frame with the Felzenszwalb-Huttenlocher graph-based segmentation algorithm. Optical flow and segmentation are then combined to track an object in a video scene. By using this strategy, also occlusion and slight rotation or deformation can be handled. The tracker is then evaluated on an artificial video sequence with moving balls but also on realworld sequences of a moving person. For all video sequences, ground truth data is available and compared to our results.

[PRIP-TR-113] Andreas Hubmer, Adrian Ion, Walter G. Kropatsch, Yll Haxhimusa, and Hubert Haussegger. How humans describe short videos - details of an experiment. 2007.

**Abstract:** Human vision can be used as a model for computer vision. We have conducted an experiment to investigate several properties of human vision that can be applied to, and that can improve computer vision. This report describes in detail the description of videos done by human subjects. Human descriptions of videos show the importance of higher levels of abstraction and that features of an object related to a task can raise the object's relevance.

[PRIP-TR-114] Sebastian Zambanini. Automatic registration of cutaneous hemangiomas in digital image series. 2007.

**Abstract:** This report presents an automatic method for registering follow-up hemangioma images taken during clinical trials in specific time intervals. The method finds interest points in two images on the basis of edge points and matches corresponding interest points using SIFT features. Under the assumption that hemangioma regions are planar, these correspondences are used to determine a homography between the two images by means of RANSAC. Experimental results are reported for image pairs acquired at the same time and for image pairs acquired during follow-up showing hemangiomas at different times. Registration errors on the images are acceptable for subsequent processing, however, gross changes of hemangiomas deteriorate the performance and are subject to ongoing research.

[PRIP-TR-115] Sebastian Zambanini. Automatic assessment of lesion development in hemangioma follow-up images. 2007.

**Abstract:** This thesis presents an automatic method for the assessment of the development of cutaneous hemangiomas in digital images. The overall method provides two measurements on photographs taken during follow-up examinations: (1) the current skin area affected by the lesion and (2) the percentage/area of the hemangioma showing a regression, a so-called graying. For both analyses a pixel-wise classification scheme is applied to the images. The actual area measurement is accomplished through an image scale computation by means of a ruler attached to skin and visible in the images. Image registration is included in the assessment procedure to align follow-up images providing a direct comparison of color values necessary for a reliable detection of regressions. For image registration a robust feature-based method is presented that is able to deal with changing hemangioma appearances during follow-up. Experimental results are reported for the individual algorithms presented as well as for the whole procedure applied to follow-up images. In general, the proposed procedure working on follow-up images is expected to provide a more accurate and objective evaluation of the course of disease than the current clinical practice of manual measurement during an examination.

[PRIP-TR-116] Michael Rauter. Effizientes object tracking durch programmierung von mehrkernprozessoren und grafikarten. 2007.

**Abstract:** In dieser Masterarbeit wird ein Object Tracking System realisiert, das bei Video-Material mit Auflungen von bis zu 768 x 576 Bildpunkten bei 3 Farbkanen Echtzeitfigkeit erreichen soll (Def.: das Verarbeiten von 25 Bildern pro Sekunde - in 40 Millisekunden mssen alle Algorithmen fertig berechnet werden). Dabei werden state-of-the-art Algorithmen verwendet, die auf CPUs mit nur einem Rechenkern nur bei geringeren Video-Auflungen Echtzeitfigkeit erreichen. Ziel ist die Performance-Steigerung eines Object Tracking Systems. Die Motivation fr eine Performance-Steigerung rhr daher, dass sich, je weniger Zeit fr das Berechnen von Verarbeitungsschritten eines Object Tracking Systems benigt wird, desto here Video-Auflungen und -Frameraten handhaben lassen (resultiert in besseren Tracking-Ergebnissen), und es bleibt mehr Zeit fr weitere Verarbeitungsschritte (z.B. herentwickelte Data Association, Behaviour Detection). Die Berechnungsschritte, die im entwickelten System durchgefhr werden und als Gesamtpaket in Echtzeit laufen sollen, sind Datenerfassung, Bewegungserkennung samt Schatten- und Reflexionserkennung inklusive deren Entfernung, Connected

Components Analysis, Verwalten der zu trackenden Objekte (Initialisierung, Tracking, Data Association, Lchen) und die grafische Ausgabe. Die Echtzeitigkeit wird dadurch erreicht, dass einerseits Algorithmen bzw. Algorithmusschritte, die parallelisierbar sind, auf die Grafikkarte ausgelagert werden und andererseits durch Multi-Threading Mehrkernprozessoren ausgereizt werden, indem verschiedene Phasen der Verarbeitungskette des Systems auf alle vorhandenen CPU-Kerne aufgeteilt und so Berechnungen simultan auf verschiedenen Rechenkernen ausgeführt werden (Datenerfassung, Bewegungserkennung, Object Tracking, Visualisierung). Mit diesen Konzepten erreicht das implementierte Object Tracking System eine Performancesteigerung um mehr als den Faktor 9 im Vergleich zu einer optimierten Single-Core CPU-Variante.

[PRIP-TR-117] Adrian Ion Mabel Iglesias Ham and Walter G. Kropatsch. Documentation for the graph pyramid drawing application. 2008.

**Abstract:** The aim of this document is to support an easy orientation in the application developed so far for the pyramid drawing problem. This application is based on an algorithm that uses paths by means of the equivalent contraction kernels to draw the edges. The drawing shows a planar graph which preserves topology but also geometry of the original image. Also, it can deal properly with the presence of multiple edges and self loops which commonly appear in the top level of irregular pyramids. Using only straight lines, the self loops would disappear and multiple edges overlap. The functionality of detecting and drawing a set of generators in the top of the pyramid has been added, by means of computing a fake new level by a last contraction using a spanning tree, and finally reconstructing the remaining loops in the previous last level. For supporting the studies to measure new topological invariants the edges have been classified in contracted, removed and surviving edges using a code of colors. Details about the input text file, set of classes, and comments about future work have been included.

[PRIP-TR-118] Martin Reiterer, Hermann Czedik-Eysenberg, and Andreas Zweng. Bachelor thesis at prip - collection of summer term 2008. 2008.

**Abstract:** This is the first collection in the series of Bachelor Thesis at PRIP that will be published twice a year, namely the summer term collection will be published at the end of October, and the winter term collection at the end of March of each year. In this collection three bachelor thesis are selected. These works are done during the summer semester 2008. The first work deals with the new image distance transformation called eccentricity transform. The author of the work, Martin Reiterer, presents a novel approximation that uses thinning of a given image region to reduce the amount of data for accelerating the eccentricity transform. In the second thesis, Hermann Czedik-Eysenberg shows a system for identifying unknown coins by matching their shape with a database of known coin shapes. In the third work, Andreas Zweng in his thesis deals with the problem of automatic number plate recognition. We would like to thank all the authors and their supervisors for their support.

[PRIP-TR-119] Adrian Ion Tamir Hassan and Walter G. Kropatsch. A demonstration for the smart room. 2008.

**Abstract:** This report describes the work carried out between January and April 2008 to create a demonstration for the Smart Room which involves tracking a blind (or blindfolded) person from several cameras in the room, which contains an obstacle.

Feedback is generated by means of audio tones from four loudspeakers positioned in the periphery. This report describes in detail how the set-up can be recreated, and suggestions are also included for continuation work.

[PRIP-TR-120] Eva Dittrich. Automatic model generation for sparse mrf appearance models using minimum description length. 2009.

**Abstract:** This report explores the combination of the Minimum Description Length (MDL) approach and Sparse Markov Random Fields Appearance Models (SAMs). SAMs are a method to locate a structure that is learnt from annotated training data in a new and unseen image. However, to achieve this result it is necessary to provide the SAMs with manual annotations of the images (landmarks) for a large set of training examples, which is a time consuming and error-prone requirement. The goal of this work is to become independent from manual annotations and to obtain the annotations automatically by using an MDL based approach. We report experimental results for different types of data (synthetic data, 2D X-rays and 3D CTs) and the method was modified to reach the best possible results for each of them. The resulting approach allows to construct SAMs in a fully automated fashion.

[PRIP-TR-121] Georg Zankl, Lukas Fischer, and Timo Kropp. Bachelor thesis at prip - collection of summer and winter term 2009. 2009.

**Abstract:** This is the second collection in the series of Bachelor Thesis at PRIP. We have collected students best works done in Summer and Winter Term 2009. First work in the collection is the bachelor thesis of Georg Zankl. He presents an algorithm for computing the eccentricity transform of a polygonal (continuous) shape which can have holes. The algorithm decomposes a polygonal shape, by gradually refining a partitioning. The partitioning is then used to efficiently compute the eccentricity for any point inside the shape. In the second work, Lukas Fischer, presents a survey on sequential Monte Carlo Methods, also known as Shape Particle Filters, used for the segmentation in medical imaging. This report surveys the robustness of these methods on different medical example images. Results on different data (synthetic rectangles, MRI slices and radiographs) are reported. The third work done by Timo Kropp, presents a software implementation that allows the user to visualize trajectories of moving object in the Microsoft Bing maps.

[PRIP-TR-122] Hubert Wagner, Chao Chen, and Erald Vucini Abstract. Efficient computation of persistent homology for cubical data. 2010.

**Abstract:** In this paper we present an efficient framework for computation of persistent homology of cubical data in arbitrary dimensions. An existing algorithm using simplicial complexes is adapted to the setting of cubical complexes. The proposed approach enables efficient application of persistent homology in domains where the data is naturally given in a cubical form. By avoiding triangulation of the data, we significantly reduce the size of the complex. We also present a data-structure designed to compactly store and quickly manipulate cubical complexes. By means of numerical experiments, we show high speed and memory efficiency of our approach. We compare our framework to other available implementations, showing its superiority. Finally, we report persistent homology results for selected 3D and 4D data sets.

[PRIP-TR-124] Tamir Hassan, Rainer Splechtna, Anton Fuhrmann, and Walter G. Kropatsch. Experiments for assistive living in the smart room. 2009.

**Abstract:** This report describes the experiments that were carried out in December 2009 at VRVis to demonstrate how blind (or blindfolded) persons can be assisted by computer-based tracking and feedback systems. The positions of the user are tracked precisely in six degrees of freedom (three dimensions position and rotation) using three infrared cameras. Feedback is generated by the use of a 3D headset. This report describes in how the set-up was created, and suggestions are also included for continuation work.

[PRIP-TR-125] Tamir Hassan. Towards a common evaluation strategy for table structure recognition algorithms. 2010.

**Abstract:** A number of methods for evaluating table structure recognition systems have been proposed in the literature, which have been used successfully for automatic and manual optimization of their respective algorithms. Unfortunately, the lack of standard, ground-truthed datasets coupled with the ambiguous nature of how humans interpret tabular data has made it difficult to compare the obtained results between different systems developed by different research groups. With reference to these approaches, we describe our experiences in comparing our algorithm for table detection and structure recognition with another recently published system using a freely available dataset of 75 PDF documents. Based on examples from this dataset, we define several classes of errors and propose how they can be treated consistently to eliminate ambiguities and ensure the repeatability of the results and their comparability between different systems from different research groups.

[PRIP-TR-126] Adrian Ion, Yll Haxhimusa, and Walter G. Kropatsch. Skewed coordinate system for dense point correspondences inside articulated shapes. 2011.

**Abstract:** This paper considers using a non-rigid coordinate system to find corresponding points in different poses of the same articulated 2D shape. The shape-centered coordinate system is mapped on top of the eccentricity transform of the shape, which uses maximal geodesic distances and is bounded under articulation. The isolines of the eccentricity transform are used as one of the coordinates, the radial-like, and the other one, the angular-like, is stretched to compensate for changes in the widths of parts. The polar-like coordinate system is first computed on inter-pixel isolines and then mapped to the pixels. The angular-like coordinates are aligned using the 1D signals of the eccentricity values along the boundaries of the two shapes. Correspondences between points are established by minimizing the difference of their coordinates. Detecting failed correspondences is done using an adaptive threshold which adjusts to the changing local variation of the coordinates. Experimental results are shown on a set of hand poses, ranging from minor movement to touching or missing fingers.

[PRIP-TR-127] Georg Zankl. Semi-automatic annotation on image segmentation hierarchies. 2012.

**Abstract:** In the field of object recognition in natural images, a variety of established tasks exist, which are focus of attention when it comes to comparing different methods, for example image segmentation, semantic image segmentation or object detection. Image segmentation is the task of grouping pixels in an image that belong to the same region or object. Semantic image segmentation is the task of assigning a semantic label

to each pixel of the image. The semantic labels can be objects: for example car, person, building; or classes of areas in an image: sky, floor, vertical surface. Object detection is the task of predicting occurrence and position in an image, for example by determining a bounding box of the object. Traditional object recognition challenges have limitations such as ambiguity in more general contexts. For example for a single natural image, there are often multiple image segmentations a human would consider to be correct, depending on the object that person is particularly interested in. We raise the question: 'Is there a different task, that overcomes these limitations?' As an example we propose the task of interactively assigning a semantic label to each segment of a segmentation hierarchy. The result can be represented as a stack of semantic segmentations, with an inclusion-relationship between segments of adjacent segmentations. The focus of this work is to provide a solution to this task and discuss advantages and problems that arise. The main disadvantage is that it is harder to obtain suitable ground-truth that consists of annotated segmentation hierarchies. Also the quality of underlying segmentation methods is, in general, sub-optimal for natural images. The main advantage is that the structure implied by the occurrence of labels in the ground-truth can be used to aid the user in labeling the segments of the hierarchy. We propose a framework that consists of a feedback loop, where a label prediction is provided by the framework and a human user may select one or more misclassified segments and assign the correct label. This process can be repeated until the user is satisfied. The prediction is done using a Conditional Random Field (CRF) that is modified so that we are able to condition the model on the segmentation hierarchy as well as the user input. The framework is evaluated on two distinct datasets by comparing its quality to a straight-forward baseline. The baseline consists of a single prediction step of the proposed framework followed by fully manual correction of the segments without new predictions. The results show a significant difference in quality, after several user interactions. For example after 20 user interactions the baseline adjusts 20 misclassified segments, while the CRF-based framework adjusts about 130 misclassified

ed segments for the two datasets. This experiment illustrates the potential of structured prediction for the proposed task.

[PRIP-TR-128] Philip Limbeck. Interactive tracking of markers for facial palsy analysis. 2012.

**Abstract:** The human face provides a rich source of information which can be exploited to diagnose facial impairments. Facial palsy is one of these impairments, and is caused by restrictions of the neural actuation of muscles responsible for facial expressions. The main symptoms of this condition are asymmetrical facial movement and partial facial paralysis. To measure its progress, physicians require clinical measures extracted from those locations of the face which provide most information about the facial expression. Small artificial markers indicate these locations by being placed on the patient's face before an evaluation session. A video of the patient is recorded which is used to localize these markers in every frame. This task is currently performed manually by an operator and can take up to 5 hours for a single video. Object tracking refers to the estimation of the position of objects from an image sequence. Illumination and occlusion are considered as the main problems when tracking artificial objects. Natural objects, such as the human face, have a high potential for deformation and are characterized by an irregular texture. As not only one, but multiple markers have to be tracked simultaneously,

additional difficulty is imposed by ensuring that markers can be uniquely identified. The thesis explores the possibility of tracking these markers semi-automatically by applying a sequential Bayes estimation technique, which assesses a set of hypothesis using their congruence with the target model. Hence, the location of each marker can be accurately estimated and occlusions handled efficiently. To improve the accuracy and to reset lost markers, the clinical operator can interact with the tracking system. The results showed that our chosen methods are superior when compared with traditional trackers which use only a single hypothesis concerning the marker locations, while at the same time being able to preserve an accuracy comparable to manual tracking.

[PRIP-TR-129] Michael Gerstmayer. Interactive hierarchical image segmentation on irregular pyramids. 2013.

**Abstract:** Image segmentation, in general, is the process of dividing a digital image into segments having a strong correlation with objects in it. Various techniques exist to locate objects of interest formed by visual cues. However, general purpose segmentation methods cannot produce a perfect final segmentation by using low-level cues only. A way round the problem is rather to create a stack of segmentations with different resolution levels. Higher level knowledge shall then be used to confirm or select regions for further processing. In automatic region-based segmentation, usually such a stack of segmentations is built in a bottom-up manner, guided by low-level image feature data and the defined homogeneity criteria. We should take into account as well that the accuracy of an image segmentation is measurable, but its quality and usability are highly subjective and depend also on the scope of the application. This thesis deals with modifications of such an irregular image segmentation pyramid and embedding additional knowledge about the problem domain such that the results of the image segmentation best suit the user. Based on an existing automatic segmentation framework where the minimum spanning tree based method tries to capture perceptually important groupings we bring the user into the loop and define interactive operations guiding the segmentation process. Semi-automatic approaches show multiple benefits (like flexibility and acceptance), but may sometimes be required also from juridical point of view. The interactive operations of merging and inhibition from merging require a representation that encodes the edge and the parent-child relationship of the merging tree. In this work each level of the irregular pyramid is represented by a combinatorial map, encoding both region and boundary information in a single combinatorial map structure. Using the connecting paths between the different levels of the pyramid, it is possible to set focus on regions from different granularities. In contrast to related approaches, this work is not limited to a single working level and pure sequential processing. Moreover, regions having different resolutions down to pixel level may be selected in parallel. This requires dedicated (pre-)processing and conflict resolution methods which guarantee consistency of applying the operations throughout the hierarchy. The output is a stack of segmentations with a final result that best suits the users' applications, in the topmost level of the hierarchy. We try to find out answers related to usability questions of the interactive segmentation tool developed and empirical values on the operations defined. As it turned out, the candidates (beginners) were able to produce results satisfying their expectations. The data recorded during the segmentation-sessions reveals different strategies and gives evidence on the usage of the interactive operations. This work can be used for problems where accuracy in image segmentation, annotating images or creating ground truth

among others is needed.

[PRIP-TR-130] Katrin Lasinger. Overview of existing software tools for graph matching. 2013.

**Abstract:** This report was created in the course of the seminar 'Selected Chapters in Image Processing' at the Vienna University of Technology under the supervision of Walter Kropatsch. The report aims to give an overview of publicly available software toolkits for graph matching. Existing graph datasets for graph matching benchmarking as well as common data structures to store graphs are presented. Libraries and toolkits for exact and inexact matching are covered in the report and their implemented algorithms are stated.