Theoretical Foundations of Computer Vision

Ruzena Bajcsy, Reinhard Klette, Walter G. Kropatsch, Franc Solina

(editors)

Dagstuhl-Seminar-Report

18.-22.3.1996 (Seminar 9612)

free copies:

Preface

This was the second TFCV meeting in Dagstuhl with the aim of bringing together scientists in Computer Vision from the West and from the former eastern block countries. The organizers believed that there was a mutual gain to all participants since we perceived a certain ignorance of each other's work. This fact explains that this seminar was less focused and covered a broad variety of computer vision topics. At the same time many participants hailed this fact, that it allowed them to educate and update themselves in such wide spectrum of computer vision topics.

The organizers assert that the goal of integrating the West and East communities at least in "small" has been achieved and no further meetings are necessary of this kind. Personal contacts have been established which now are leading to more direct exchanges between students and researchers of respective laboratories.

38 talks were given by 40 participants from 17 countries. 11 participants came from Germany, 5 from France, 3 from the Netherlands, 2 from Austria, Canada, Czech Republic, Italy, Poland, Slovenia and Ukraine, and 1 from Australia, Hungary, Japan, Russia, USA and Yugoslavia.

We are pleased to note that at this meeting by and large the presentations were highly professional, well structured, to the point and reporting most recent results as it is obvious from the bellow abstracts. This report contains the abstracts in alphabetical order sorted by the names of speakers.

We are grateful to the administration of the Dagstuhl enterprise for creating a conducive environment that enabled us this successful interaction and accomplishment of our primary goals.

For the future, we recommend that TFCV seminar be more topical and focused. "Evaluation and Validation of Computer Vision Algorithms" was the suggested topic of the next TFCV seminar in the final discussion.

Ruzena Bajcsy, Reinhard Klette, Walter G. Kropatsch, Franc Solina

List of Speakers

Monday

G. Sommer, R. Bajcsy, B. Krebs, R. Kozera, V. Kovalevsky, K. Voss, K. Daniilidis, G. Hartmann, Ch. Duperthuy

Tuesday

N. Petkov, J. Barron, St. Beauchemin, A. Gagalowicz, H. Jahn, S. Ben Yacoub, V. Di Gesu, V. Hlavac, W. Skarbek, F. Solina, A. Leonardis

Wednesday

J.-M. Jolion, A Koschan, R. Klette, M. I. Schlesinger, H. Süsse

Thursday

U. Eckhardt, J. B. T. B. Roerdink, A. Imiya, J. Zunic, F. Sloboda, W. G. Kropatsch, J.-O. Lachaud, D. Chetverikov, G. L. Gimelfarb, J. Weickert

Friday

J. Flusser, T. Belikova, R. Bajcsy

Signal to Symbol to Signal Transformation: A Fundamental Problem of Representation

by Ruzena Bajcsy (joint work with T. Kosecka, H.-H. Nagel, H. Christiansen)

Research Objectives: To find the formal framework for the Signal to Symbol and Symbol to Signal Transformation. We view this as a fundamental problem of representation. We begin with asking the following question: If the world is in some sense continuous (perceived and acted via signal) then why do we need symbols (or discrete representation)? Furthermore if we need symbols/discrete states/quantized information then the next questions are:
1. What is the alphabet into which signal(s) should be partitioned/segmented?
2. Is there a natural way, predicated on the physics (nature) of the problem that can determine this partitioning?
3. How task context and their complexity determines/constraints this partitioning?
Technical Approach: We have investigated this problem in the context and the task of visually guided navigation. We assert that:
1. Using Lagrange formulation of the dynamic interaction of the agent and its environment performing the task, the generalized coordinates (that are part of the above formulation) determine the degrees of freedom of this system, which in turn determines the number of discrete states of the system.
2. We furthermore partition the world into DISCRETE PLACES, which correspond to equivalence classes of visibility.
3. We model both the agent/environment interaction and the places which are aimed with discrete perceptual strategies using FINITE STATE MACHINES. Each state in the Finite State Machine is associated with a CONTINUOUS (modeled by an ordinary differential equation) control strategy to carry out either PERCEPTUAL or ACTION strategy.
This is for the FIRST time to our knowledge that such a UNIFIED framework has been achieved for both the perception and action strategies. We show complete physical experiments that TEST the above ideas and framework.

Computation of Time-Varying Motion and Structure Parameters from Real Image Sequences

by John L. Barron (joint work with R. Eagleson)

We outline a solution for recovering camera motion and scene structure parameters from a stream of optical flow fields computed from long image sequences. The solution involves solving simple 2 x 2 and 3 x 3 systems of linear equations to recover the camera direction of translation and the camera's rotation. Camera rotational acceleration can also be recovered and individual surface relative depth values by solving simple linear equations. The motion parameters are integrated over time using a Kalman filter, yielding at each time, a best estimate of the motion parameters. We conclude by analyzing the algorithm's performance for two real image sequences. The correct motion and structure parameters are known for these sequences, allowing a quantitative error analysis of the motion parameters. We also show that accuracy of the relative depth depends on the accuracy of the individual image velocities and depends on the presence of significant translational flow.

A Theory of Occlusion in the Context of Optical Flow

by Steven S. Beauchemin (joint work with J. L. Barron)

A fundamental problem in processing sequences of images is the computation of optic flow, an approximation of the motion field defined as the projection of velocities of 3-d surface points onto the imaging plane of a (biological or mechanical) visual sensor. A wide variety of algorithms for computing optical flow has been reported in the literature. However, it is not until recently that multiple motion paradigms were seriously studied. Such approaches are necessary to encompass image events such as occlusion and translucency which violate the classical single motion hypothesis. While current approaches to this problem constitute valuable contributions, they often lack the ability to determine the nature of the image events giving rise to multiple motions. In these schemes, occlusion may not be differentiated from translucency and it remains difficult to explicitly identify the motions associated with both the occluding and occluded surfaces. In this analysis we demonstrate,! under a minimal set of assumptio

Image Processing and Syndrome Analysis for Decision-making under Conditions of Uncertainty

by Tatjana Belikova (joint work with I. Stenina, N. Yashunskaya

We consider an approach to the solution of complex tasks of image analysis and interpretation when uncertain data and informal knowledge about the task solution are in one's disposal. Such a case is very usual for human processing of visual information in the case of uncertainty. CV methods could be useful to support decision-making of the viewer reducing the level of uncertainty and number of ambiguous, wrong and subjective decisions-makings at image analysis and interpretation. The goal of the research was to develop methods and methodology for more relieve representation of visual information to the viewer; description of the image with the use of expert's knowledge; identification and recovery of number of features in terms of expert's description that are specific for image classification; analysis and interpretation of image descriptions in the case of uncertainty; recovering the goal and the strategy for image information computing at visual tasks solution.
The approach is proposed to develop image processing methods that give more relieve presentation of visual information for the viewer. Description of the images on the base of expert's knowledge and data analysis in the database of image descriptions helped to select discriminative feature and reduce the total number of features that should be used for image representation. The formal decision rules where developed that model the decision-making process of the expert at uncertainty. Developed formal decision rules can be used to recover the goal of computation of visual information and the strategy by which it can be carried out. Developed set of methods can support image interpretation by expert in the case of uncertainty and can be a navigator for visual data representation and computation with the methods of CV. Some experiments with analysis and interpretations of medical image information are presented.

Detection of Subjective Contours - By Means of the Hierarchical Hough Transform

by Souheil Ben Yacoub

The work deals with the extension of contour notions. Classically an edge or contour is defined us a variation of luminance variation. Subjective contours are perceived in areas where there is no luminance variation. So a contour could be seen us a discontinuity of continuous spatial property. Subjective contours (like Kumiszu Triangle) are suggested by a principle of "evidence accumulation". This principle is naturally present in the Hough transform. We propose to use the hierarchical version of the Hough transform (HHT) to design a 2 step algorithm for detecting subjective contours. A first map is created by detecting textures (end of lines, crossings, curvature points). The edges are then detected by accumulating the "texture map" with the HHT. This algorithm detects in a very easy way the "subjective contours". The next work on this topic should be an extension to more complex "subjective contours" which are created by many different phenomena (frequency, phase changes e! tc.).

Texture Feature Based Interaction Maps: Potential and Limits

by Dmitry Chetverikov (joint work with R. M. Haralick)

A new tool for texture analysis called feature based interaction map (FBIM) has been recently introduced by the authors. It has been shown that the FBIM can be efficiently used to assess fundamental structural properties of textures such as anisotropy, symmetry, orientation, and regularity. We have also demonstrated that the FBIM is suitable for rotation-invariant texture classification of patterns with regular, weak regular, or linear structure. In this talk, we (1) analyze relations between the fundamental structural properties; (2) propose a fast running implementation of the FBIM algorithm at such tasks as segmentation and detection; (3) present pilot experimental results demonstrating the potential of the FBIM approach in diverse tasks and applications; (4) discuss the limits of the FBIM method. A related Markov-Gibbs model based approach is also discussed in connection with the feature based approach presented. The first self-contained description of the FBIM method is! given.

Modeling 3D-Transformations Using Dual Quaternions: The Case of Hand-Eye Calibration

by Kostas Daniilidis

In order to relate measurements made by a sensor mounted on a mechanical link to the robot's coordinate frame we must first estimate the transformation between these two frames. Many algorithms have been proposed for this so called hand-eye calibration but they do not treat the relative position and orientation in a unified way. In this paper we introduce the use of dual quaternions which are the algebraic counterpart of screws. Then it is shown how a line transformation can be written with the dual quaternion product. We algebraically prove that if we consider the camera and motor transformations as screws then only the line coefficients of the screw axes are relevant regarding the hand-eye calibration. This new parametrization enables us to simultaneously solve for the hand-eye rotation and translation using the Singular Value Decomposition. The behavior of our new algorithm is experimentally analyzed in comparison to algorithms using the classical motion parametrization

Detection of Regions of Interest via the Pyramid Discrete Symmetry Transform

by Vito Di Gesu (joint work with C. Valenti)

A new approach to detect regions of interest in a digital scene is proposed. It is based on a pyramid implementation of the Discrete Symmetry Transform (DST) that allows to detect local symmetries (Di Gesu, Valenti 1995). The resulting hierarchies of symmetries are then used to find points of interest in the scene. Pyramid computation has been introduced in computer vision to design both top-down and bottom-up efficient vision algorithms. This paradigm of computation can be related to the work on visual perception made by Pomeranz and Soger (1975); the authors describe a visual process in which the attention goes from global to local information performing an holistic process or from local to global performing what they call an analytic process. The paradigm can be synthesized with the expression suggested by Navon (1977): Are we seeing the forest before trees or trees before the forest.

Pyramid computation has suggested both new data structure (quad-trees, multi-resolution), and new machine vision architecture. The concept of regular pyramid has been introduced by Kropatsch to handle connectivity problems that can arise mapping data through the pyramid layers. Yet, symmetry plays a remarkable role in perception problems. For example, peaks of brain activity are measured in correspondence with visual patterns showing symmetries. Relevance of symmetry in vision was already noted by psychologists (Koeler, 1929). Symmetry operators have been included in vision systems to perform different visual tasks. For example, a set of annular operators can be used to identify enclosed symmetry points, and then a grouping algorithm is applied to represent and to describe object-parts (Kelly 1994); the directional property of the symmetry axes has been applied to perform image segmentation (Gauch 1993).

The algorithm to compute the pyramid DST, has been implemented, in two versions: a) direct way, in this case the DST of the pyramid image is computed; b) indirect way, in this case the pyramid of the DST or the input image is computed. The near-commutativity of this two modalities is defined, and it is shown that it holds in the case of Gaussian pyramid. This result allows us to speed-up the computation and to retrieve areas of interest at low resolution pyramid layers. Preliminary experiments on real images confirm our claim, and are encouraging. Several problems are still under investigation, as: dependence from illumination, affine transformations invariance, presence of structured background.

Towards a Generalized Primal Sketch

by Christophe Duperthuy

The representation of visual data extracted from pictures is of importance according to the performances of the visual system to conceive. Such a representation should explicit all the real physical attributes, while being robust. Moreover, edge based approaches are not always suited to scene representation. In this way, we propose a new Primal Sketch based on a simplified pyramidal representation of the scene to deal with: it results in robustness and fast calculations (using a massive parallelism), especially with the use of 3x3 binary masks classifications (look-up tables for a real-time extraction!).

We obtain a set of attributes cards, such as features (noises, textures, shapes and blobs) on orientations (over 8 basic directions), for various resolution levels.

Topological Approach to Mathematical Morphology

by Ulrich Eckhardt (joint work with E. Hundt)

A very promising approach to coding of image sequences is object based coding. The first step of such a process is image segmentation. One class of segmentation methods is morphological segmentation, specifically watershed segmentation.

There are three structures imposed on an image in order to perform the method: A topological structure for defining connected components, a metric structure for geodesic distances and a morphological structure. Three questions raised during our experiments:

how should these three structures be chosen,

how should anomalies be treated,

how can the segments detected at low level be extended to spatial, topological and semantic relations on the set of segments as to handle them in the data stream.

It turns out that semi-topological spaces of character 1 (Latecki 1992) are the appropriate tool for answering these questions. Specifically, it is shown that such spaces coincide with morphological spaces.

Improving Medical Treatment of Gray Cloud by Advanced Image Processing Techniques

by Jan Flusser (joint with J. Novak, St. Saic and T. Suk)

The contribution shows a successful application of image processing in ophthalmology, particularly in the post-operation therapy of gray cloud. The intra-ocular lens implanted into the patient's eye is often attacked by the cells of immunity system (macrocytes or macrophages). The post-operation therapy consists on removing of the immunity

cells from the lens by means of appropriate medicine (eye drops). To determine an optimal medicine, a long-term image sequence of the patient's eye was registered by surface spline method with subpixel accuracy. This enables to investigate the movement of the cells and their behavior under various medical treatment.

Analysis/Synthesis of Monocular Image Sequences

by Andr   Gagalowicz
(abstract not available at time of editing this report)

Non-Markov Random Fields with Almost-Pairwise Pixel Interactions for Modeling Image Textures

by Georgy L. Gimelfarb

A novel class of non-Markov Gibbs random fields is proposed for modeling spatially uniform and piecewise-uniform grayscale image textures with due regard of possible changes of their gray level ranges. The model belongs to an exponential family of probability distributions. Gray level/region label histogram and gray level difference/region label coincidence histogram for a given learning sample which is normalized by reducing to a standard gray range from the sufficient statistics for describing this texture type. The model preserves, to a great extent, most attractive features of the original Markov/Gibbs model with multiple pairwise pixel interactions, in particular, analytical initial estimates of the parameters (Gibbs potentials), stochastic approximation refinement of these estimates by using a modified stochastic relaxation to generate field samples, etc. The model is close to the Markov/Gibbs one in computational complexity. Experiments in modeling, segmenting, and re! trieving natural textures show go

The Influence of Neural Representations on the Reliability of Object Recognition

by Georg Hartmann

Holistic learning of object views is a very common approach to 2D and 3D recognition. At the level of a block diagram, the architectures are very similar and usually there are the following components:
- a log-pol representation for normalization of distance and orientation,
- a neural representation of contour based features,
- smoothing for tolerance against foveation errors and against minor variations of shape, and
- subsampling for data compression.
There are, however, essential differences between corresponding components of different systems, and in our opinion these differences are responsible for different performance. The original log-pol representation has two disadvantages:
- low quality of shape description due to underrepresentation in the periphery, and
- sensitivity to foveation errors due to overrepresentation in the center.
We could overcome this problem by multiple log-pol representation in a set of spatial frequency channels. Some authors are very careful with contour representations, and provide very elaborate coupling schemes between different types of neurons. In our system we successfully use very simple detectors only evaluating the pattern of positive and negative signs in the Laplacian image. Obviously, the results are not very sensitive to this component. While other authors provide spatial tolerance by (Gaussian) smoothing, we use a special type of complex model neurons. Our complex model neurons respond to "continuous sequences of simple cells within an enlarged receptive field, independent of the exact position of this sequence", and so they provide spatial tolerance without smoothing. A vector with a small number of analog components is provided by smoothing and subsampling in other systems. In our system the sheer activation pattern of our special complex model neurons is learnt. Though we learn long vectors with many binary components, there is no major difference in storage capacity. However, as generalization is due to the special representation with spatially tolerant complex model neurons, and not to learning in our case, we are able to learn very quickly by only one presentation of each view of the learning set. A moderate size of the prototype set, and very high recognition rates are obviously due to the special details of our representations.

Automatic Selection of Reference Views for Image-Based Scene Representation

by Vaclav Hlavac (joint work with T. Werner)

The work is related to displaying a real 3D-scene from any viewpoint. A sparse set of reference 2D-views is stored. The images that are in-between the reference views are obtained by interpolation of coordinates and brightness. This allows to represent and render complex scenes. The processing time is basically independent of the scene complexity. There is no attempt to understand the semantics of the image.
The novel approach to automatically determine a minimal set of reference views is presented. The method consists of two steps: viewing interval growing and selection. The first step searches for intervals (regions) of views from which scene can be rendered with appearance dissimilarity smaller than some predefined error.

Metric of Lines and Discrimination of Lines

by Atsushi Imiya

Duality of points and lines enables derivation of the randomized Hough transform which can detect any line on a plane by voting to a finite accumulator space embedded in a half-cube in three-dimensional Euclidean space. This paper also defines a measure for line segments on the plane by using one-to-one mapping between a line segment on the plane and a region between two large circles in the position unit semi-sphere. Furthermore, the metric of two line segments is defined as the Hamming distance between two regions corresponding to line segments. The arrangement of two line segments is classified into six categories. Thus, by using arrangement of line segments, an algorithm for the computation of the Hamming distance between two regions in the positive unit semi-sphere is also derived.

Image Segmentation with a Layered Graph Network

by Herbert Jahn

A new Layered Graph Network (LGN) for image segmentation is presented. In the LGN a graph representation of images is used. In such a Pixel Adjacency Graph (PAG) a segment is considered as a connected component. To define the PAG the layers of the network are divided into regions, and inside the regions the image is represented by sub-graphs consisting of sub-segments (nodes) which are connected by branches if they are adjacent. The connection of sub-segments is controlled by a special adjacency criterion which depends on the mean gray values of the sub-segments and their standard deviations. This way, the sub-segments of a layer l are merges of sub-segments of layer l-1 (the sub-segments of layer 0 are the pixels). The gray value averaging over the sub-segments is edge preserving and becomes more and more global with increasing number of the network layer. Bridge connections between the segments are prevented by the special regional structure of the network layers. The LGN ! can be understood as a special ne

Categorization through Temporal Analysis of Patterns

by Jean-Michael Jolion

We present some trends toward a new approach of artificial vision. We first introduce the main points related to already proposed methodologies based on the Marr's paradigm. We then show that the inside behavior of the observer is a source of information as well as the physical properties of visible surfaces. We argue that the Marr's paradigm mainly takes care of the "what" we are looking for. The behavior or goal based approaches like active perception, active vision and animate vision takes also care of "why" we are looking at the scene. We propose that we also take into account "How" we are looking at the scene. Examples are proposed in the particular case of connected components extraction in a binary image using a data flow pyramid computer showing that "How" we process information may be more important than "what" we actually process.

Height Data from Gradient Maps

by Reinhard Klette (joint work with K. Schlüns)

Several methods in the field of shape reconstruction (typically all reflectance based methods) lead to gradient data which still have to be transformed into (scaled) height or depth maps, or into surface data. Thus the reconstruction accuracy also depends upon the performance of such a transformation module. Surprisingly not much work was done so far in this area. There exist a few proposals of local integration methods (Coleman/Jain 1982, Healey/Jain 1984, Wu/Li 1988, Rodehorst 1993), and two proposals for global optimization (Horn/Brooks 1986 and Francot/Chellappa 1988). Several experimental validations of different methods for transforming gradient data into height data were performed, and the results depend upon the object shape. Polyhedral objects (e.g. very relevant to industrial applications) are the most critical shapes.

The talk reviews transformation algorithms for height from gradient data and specifies several methodologies for experimental validations of different transformation procedures. The proposed error measures have to be discussed in relation to applications of shape reconstruction. In the ideal case, certain "data sheets" about transformation modules should support the decision which one can be selected for a specific application. The studied (synthetic and real) object classes are spheres, polyhedral objects and curved objects. General qualitative evaluations of the compared transformation procedures are possible in relation to these object classes and in relation to different types of noise simulated for synthetic objects.

Dense Depth Maps by Active Color Illumination and Image Pyramids

by Andreas Koschan

Stereo is a well-known technique for obtaining depth information from digital images. Nevertheless, this technique still suffers from a lack in accuracy and/or long computation time needed to match stereo images. A new hierarchical algorithm using an image pyramid for obtaining dense depth maps from color stereo images is presented. We show that matching results of high quality are obtained when using the new hierarchical chromatic Block Matching algorithm. Most stereo matching algorithms can not compute correct dense depth maps in homogenous image regions. It is shown that using active color illumination may considerably improve the quality of the matching results. The benefit of this approach is shown for synthetic and for real images. In summary, we should like to emphasize that active color illumination always serves to improve stereo matching results. Therefore, we believe that more precise results can be efficiently obtained in dense stereo matching when combining hier! archical chromatic Block Matching

A New Approach to Shape from Shading

by Vladimir Kovalevsky

Recent works on Shape from Shading consider the problem as the restoration of the spatial shape of a smooth continuous surface when given a continuous function representing the brightness at each point of a planar projection of the surface. To solve the problem, properties of partial differential equations are investigated. The ultimate solution is then performed by some numerical methods. The author considers the great discrepancy between continuous models and the theory of differential equations on the one hand and the digital nature of the images as well as the use of numerical methods on the other hand as unsuitable. Thinking first about continuous functions and differential equations, and then coming back to some numerical representations and numerical solutions, is irrational. The author suggests therefore an essentially digital approach: the surface to be reconstructed is considered as a polyhedron. The image contains then a finite number of plane polygonal regions ea! ch having a constant gray value.

Recent Results in Shape from Shading

by Ryszard Kozera

In this paper (talk) we analyze the problem of representing solutions of a first-order partial differential equation in terms of complete integrals. We show that some of the existing results (see e.g. Snedton in "Elements of Partial Differential Equations") referring to the representability problem are incomplete and in fact incorrect. Additionally,

we analyze, in the context of complete integrals, the uniqueness problem for the shape recovery of a smooth Lambertian surface from an image obtained by illuminating this surface by an overhead distant light-source. Specifically, we revisit the uniqueness results already existing in the shape-from-shading literature (see e.g. Ph.D. by M. Brooks "Shape from Shading Discretely", Essex University, England) that concern eikonal equations corresponding to the images of a Lambertian hemisphere and a Lambertian plane. We show that the latter results are incorrect and indicate how to fill the gaps in the corresponding proofs.

Depth Analysis and Application to Model-Based Object Recognition

by Björn Krebs (joint work with F. M. Wahl)

Computer aided geometrical design provides very powerful methods to describe objects in terms of freeform surfaces. The generality and the useful properties had made this descriptions quite common in computer graphics, the car or airplane industry. The future goal will be to introduce those concepts to all components in computer aided manufacturing or similar applications.
We propose a CAD based 3D object recognition system. Thus, objects can be described by freeform surfaces like rational/nonrational Bezier patches, B-Spline-Patches or NURBS-Patches etc. in a common input format (DXF, IGES). Range Data is acquired by a sensor based on the coded light approach. In one recognition process multiple views are used. Thus, new sensor information is acquired incrementally until the location of all objects are computed. First estimation of the object's locations are generated by an efficient 3D segmentation into homologous regions. The initial locations are improved by a matching algorithm. Whereby, the system can handle even highly complex scene constellations containing overlapping objects by an accurate matching with a fuzzy iterative closest point algorithm. Furthermore, the best match is determined by a evidence of reference. Thus, we can provide a very powerful tool for further applications like grasp operations or assembly planning.

Equivalent Contraction Kernels to Build Dual Irregular Pyramids

by Walter G. Kropatsch

Dual graph contraction reduces the number of vertices and of edges of a pair of dual images 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 contraction kernels which consist of a subset of 'surviving' vertices and a subset of 'primary non-surviving' edges. Contraction kernels are required to form spanning forests. Connecting paths establish the relation between edges in the reduced graph and their connectivity in the level below. Each connecting path consists of a number of edges of the contraction kernels and one surviving edge which is called 'bridge'. 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. Contraction parameters of any individual pyramid l! evel can be reconstructed from th 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.
As a first application of ECK it is proved that dual irregular pyramids can represent any possible segmentation result in one single level. This overcomes the famous 'shift variance problem' of regular pyramids as stated by Bister, Cornelis and Rosenfeld.

Hierarchic and Dynamic Method for Modeling Surfaces: Application to Segmentation of 3D Data

by Jacques-Olivier Lachaud (joint work with A. Montavert)

Volume segmentation has become a major research topic in the last years. This is due to the appearance of three dimensional data in medical, geological or biological domains. This kind of data can come from either MR,

tomography or confocal microscopy. Whereas bidimensional segmentation tries to mimic the vision process of the human being, volume segmentation widens the detection of shapes to the reconstruction of complex volumes, which is a difficult task for our mind. Direct segmentations one often inadequate and the report to a deformable model is then essential.

We present a method for segmenting three dimensional images. It is based on a dynamic triangulated surface and on a pyramidal representation. The triangulated surface, which follows a physical modeling and which can as well modify its geometry as its topology, segments images into their components by altering its shape according to internal and external constraints.

In order to speed up the whole process, and algorithm for pyramid building, with any reduction ratios allows us to transform the image into a set of images with progressive resolutions. This organization into a hierarchy, combined with a model that can adapt its mesh refinement to the resolution of the workspace, authorizes a fast estimation of the general shapes included in the image. After that, the model searches for finer and finer details while relying successively on the different levels of the pyramid.

Modeling and Recovery: Analytic Parametric Models versus Eigenimages

by Ales Leonardis

Analytic parametric models are commonly used in computer vision to achieve "signal-to-symbol" transformation. Their main advantage is that there exist robust techniques for their recovery from image data. However, the analytic parametric models are often quite limited in their expressive power and flexibility. On the other hand, the appearance-based representations (eigenimages) have recently received a renewed attention primarily because they can be acquired through automatic learning. The basic limitations of the current recovery of appearance-based models are non-robust estimation of coefficients and inability to cope with problems related to occlusions and segmentation.

We show that there exist certain similarities between the analytic parametric models and the representations in terms of eigenimages, and that a robust and efficient methodology, which we developed for the recovery of analytic parametric models, can be utilized for the recovery of eigenimage-representations. The major novelty 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 MDL (minimum description length) principle. The approach not only enables the robust recovery of appearance-based representations in cases of outliers, cluttered background, and occlusions but also to simultaneously use multiple classes of eigenimages. More generally, it can also be seen as an extension of the robust recovery of analytic par! ametric models to learnable class

Visual Cortex Simulation and Image Pattern Classification

by Nikolay Petkov

A preprocessor based on a computational model of simple cells in the mammalian primary visual cortex is combined with a self-organizing artificial neural network classifier. After learning with a sequence of input images, the output units of the system turn out to correspond to classes of input images and this correspondence follows closely human perception. In particular, groups of output units which are selective for images of human faces emerge. In this respect the output units mimic the behavior of face selective cells that have been found in the inferior temporal cortex of primates. The system is capable of memorizing image patterns, building autonomously its own internal representations, and correctly classifying new patterns without using any a priory model of the visual world.

Morphological Segmentations by Watersheds: Algorithms and (Parallel) Implementation

by Jos Roerdink

The watershed transformation is used in morphological image processing for segmenting gray scale images. Two approaches for arriving at a parallel implementation of this algorithm are discussed, both using a graph formulation. The first approach starts from the recursive definition of watersheds by Vincent and Soille. First the image is transformed to a graph in which each vertex represents a connected component at a certain gray level L. Then we

compute the watershed of this graph as in the Vincent-Soille method, and finally transform the result back to the image domain. The computation of a skeleton of plateau regions is performed as a postprocessing step. The three stages of this algorithms are parallelized using a SPMD approach.

The second approach is based on Meyer's definition of watersheds in terms of shortest paths with respect to a topological distance function which takes gray level information into account. The shortest paths can be computed by Dykstra's algorithm, which can easily be adapted to arrive at a parallel implementation. This algorithm has been implemented on a Cray Jg32. Speedups for a set of test images are presented as a function of the number of processors.

Generalized Convolutions and their Application to Solution of the Certain Best Matching Problems

by M. I. Schlesinger

A problem of calculation of distance by Levenstein between string and languages of same types is investigated. It is shown, that for regular, context-free and two-dimensional context-free languages a calculation of the distance by Levenstein between string and a language has a computational complexity of the same order as a testing, whether the string belongs to language or not. It means that for regular, context-free and two-dimensional context-free languages the problem of the best matching is computationally equivalent to the exact matching problem. A proof of such type of equivalency is fulfilled with a help of generalized convolutions, and equivalent transformations of expressions composed of convolutions.

A proof is published in [Schlesinger, Systeme von Funktionsoperationen angewendet auf eine Aufgabe der besten &UUML;bereinstimmung. ISSN 0863-0798, Wissenschaftliche Beiträge zur Informatik - Fakultät Informatik der TU Dresden, 7(1994) Heft 3, S. 62-79].

Associative Memory for Images by Recurrent Neural Subnetworks

by Wladyslaw Skarbek

Gray scale images are represented by recurrent neural subnetworks which together with a competition layer create an associative memory. The single recurrent subnetwork N_i implements a stochastic nonlinear operator F_i constructed for the given image. It is shown that under realistic assumptions F_i has a unique attractor which is located in the vicinity of the original image. The associative recall is implemented in two stages. Firstly, the competition layer finds the most invariant subnetwork for the given input distorted, incomplete or noisy image g. Next, the selected recurrent subnetwork reconstructs (in few iterations) the original image. The degree of invariance for the subnetwork N_i on the input g is measured by a norm $\|G-F_i(g)\|$. Comparing to Aman-Hopfield associative memory, this solution has no spurious states, is less sensitive to noise and data incompleteness, and its complexity is significantly lower. It can remove additive and multiplicative noise with s # 2 * "! pixel range", also impulsive nois

On Boundary Approximation

by Fridrich Sloboda

Piecewise linear approximation of planar Jordan curves and arcs by gridding techniques are described. The algorithms are based on the shortest path and geodesic diameter calculation in corresponding polygonally bounded compact sets. Applications to approximation of planar curves in implicit forms are investigated. The most important curves in implicit forms for segmentation purposes are presented.

Optimal Models for Visual Recognition

by Franc Solina

Various approaches have been proposed for building models of objects from sensor data. This paper discusses the

criterion for selecting an optimal model for given objects from a set of possible models. The application of MDL (Minimal Description Length) principle as a criterion of model selection is discussed in detail The relation among ML (maximum likelihood) principle, Occam's razor principle and MDL is presented. The MDL principle which is the most general states that the model which minimizes the sum of the encoding of objects given model and the encoding of the model should be selected. Since finding the optimal encoding of model is intractable we propose suboptimal encoding for non graph theoretic models for modeling objects from sensor data. Experiments of finding line models of binary edge images are shown. From a redundant set of line segments obtained by the Hough transform only those line segments giving the most compact model of edges in the MDL sense is selecte! d. The algorithm performs local o

Geometric Algebra - A Framework for the Perception - Action Cycle

by Gerald Sommer

The perception- action cycle (PAC) as the frame of autonomous behavior relates perception and action of an animal in a purposive manner. Behavior as the observable manifestation of competence shows such attractive properties as robustness and adaptivity. These properties favor behavior based design of artificial systems as alternative to knowledge based system design. Within PAC perception and action cannot be separated, neither in the stage of learning competence by experience nor in the matured stage. The design of artificial systems, which can handle the PAC, calls for a fusion of computer vision, robotics, and neural computation. Evaluating the actual situation in these fields on several levels of abstraction the great importance of representations will became obviously. Representations have to relate to the agent, the environment and to the relations of both in space-time.
There are some well identified dichotomies and representation gaps which hinder the requested fusion. These dichotomies have to been bridged by a unified language describing the use of several degrees of freedom in space-time. This language has to express a lot of group theoretically constrained geometric isomorphisms between the phenomena of the world and their representations. It has to be a rich algebraically structured geometry or a geometrically interpretable rich structured algebra. Several serious shortcomings of signal theory, computer vision, robotics, and neural computation are identified. Mostly they are based on the poverty of structures in usually used vector space algebra.
The Geometric Algebra (GA), a geometric interpreted version of Clifford Algebra has been formulated by D. Hestenes (1984). This algebra of directed numbers of any dimension will be proposed as the most attractive and flexible algebraic structure of a vector space with several important implications to signal theory, projective geometry, robotics, and neural computation. It offers the advantages of both representing geometric entities of any dimension within the GA of a vector space and representing structures (or patterns) and operations on these in the same language. Both perceived and by motoric activities generated patterns in space-time can be understood as orbits of a Lie Algebra of geometric transformations of interest which are embedded into that GA. Which GA represents the general frame for any actions (operations) or patterns, Lie Algebra is the frame of differential generation of such patterns.
The basic ideas of GA of vector spaces are introduced. These are geometric product, multivectors and blades, projection and rejection. GA of Euclidean space, and linear operators and outermorphism. A short sketch outlines the application to design a GA-structured linear signal theory.
The relation of GA and projective geometry is based on the duality principle as element of GA. Similar as a linear vector space Vn can be extended to a GA of that space, G(Vn), there exist an extension of projective space Pn to a GA of its metric pendant G(Vn+1). While the physical space E3 has Euclidean metric, the space of homogeneous coordinates V4 has Minkowski metric. Projective split represents a canonical relation between GA's of different dimension. It is of high importance to relate projective, affine, and metric geometry. Outermorphism enables to compute projective transforms. The algebra of incidence is constituted by the operations join and meet of blades and is deeply related to the geometric product.
Several results of computing invariants on n-linear constraints in projective geometry are sketched. Concerning robotic and neural computation some outlines of projects of the Kiel Cognitive Systems Group are presented.

Invariant Fitting of Planar Objects by Primitives Using Affine Normalization

by Herbert Süsse (joint work with K. Voss)

The determination of invariant characteristics is an important problem in pattern recognition. Many invariants are

known which have been obtained by the method of normalization. In this talk, we introduce a new approach of fitting planar objects by primitives using the method of normalization (for instance: fitting by lines, triangles, rectangles, circles, ellipses, super-quadrics etc., i.e. also a fitting by non-analytical primitives).

Objects and primitives are described by features, for example by moments. The main advantage is that the normalization process provides us with a canonical frame of the object and the primitive. Thereby caused by technical requirements, different normalization methods are necessary (for instance, the standardized objects - the "canonical frame" - are characterized by m10=m01=m11=m30+m12=0, m20=m02=1 or they are characterized by m10=m01=m30=m03=0, m20=m02=1 etc.). The basic idea is to determine the affine transformation T between object O and canonical frame F at one side and the affine transformation T' between primitive P and canonical frame at the other side. The transformed primitive T"T'P (with T" as inverse transformation of T) then yields an optimal fit of O by a primitive of the assumed equivalence class (triangles, rectangles, circles etc.). Therefore, the fit is invariant with respect to the used transformation. By this new method, an analytical fitting of non-anal! ytical objects can be done, for e

Affine Normalization and Projective 3D-Reconstruction of Planar Smooth Boundary Objects

by Klaus Voss (joint work with M. Schubert, H. Süsse)

In robot vision, the 3D-reconstruction of planar objects characterized by CAD-information is often necessary. This reconstruction must be taken into account the projective distortion of the corresponding image object. If the object is a polygon, the reconstruction can be done by point matching and geometrical/numerical calculations using the parameters of a calibrated camera. Especially for triangles, a polynom of order four can be derived which gives the solutions with accuracy depending of the measurement errors of corner point coordinates in the image and of the errors of the calibrated camera. This method shall be generalized to objects with smooth boundaries. Using the method of affine normalization, the CAD-object C is affinely transformed by N=TC to a "standard object" N characterized by moments (for example by m10=m01=m30=m03=0, m20=m02=1). The image object O is affinely transformed by N'=T'O to a very similar standard object N' also characterized by the above given ! moments. Therefore we can approxi Finally, we determine for three or more points of C the corresponding image points what means an "accurate" point matching. From this point matching, we can determine location and orientation of the smooth boundary object in the same manner as for polygons.

Semidiscrete and Discrete Scale-Space Theory for Nonlinear Diffusion Filtering

by Joachim Weickert

Although much effort has been spent in the recent decade to establish a theoretical foundation of certain PDEs as scale-spaces, the question of how to construct discrete scale-spaces with similar properties as their continuos counterparts has hardly been addressed. This problem, however, is of great practical relevance, since every digital image is sampled on a finite pixel grid. Thus, scale-spaces cannot perform better than their numerical discretizations. The goal of this communication is to present a discrete scale-space framework for nonlinear diffusion filtering. Two cases are distinguished: The semidiscrete case with a continuos time parameter and a spatial discretization on a fixed pixel grid, and the (fully) discrete case, where in addition a temporal discretization leads to a finite number of scales. In the first case we study nonlinear systems of coupled ordinary differential equations. Conditions are established under which one can prove existence of a unique solution which is stable against perturbations of the initial image and preserves the average gray level. An interpretation as a smoothing scale-space transformation is introduced which is based on an extremum principle and the existence of a large class of Lyapunov functionals comprising for instance p-norms, even central moments and the entropy. The guarantee that the process is not only simplifying and information-reducing, but also converges to a constant im! age as the scale parameter t tend In the fully discrete scale-space framework prerequisites are presented leading to the same results as in the semidiscrete case. It is possible to obtain some of these processes as explicit or a-semi-implicit time discretizations of the semidiscrete filter class. Of special practical importance is the semi-implicit case, since this framework is proved to satisfy all the all the above mentioned well-posedness and scale-space properties for arbitrary large time step sizes. This gives rise to absolutely stable algorithms which are far more efficient than the widely-used explicit schemes.

Least Squares Fit Representations of Digital Objects

by Jovisa Zunic

The idea for using the least squares fitting technique for representation of digital objects was introduced by Melter and Rosenfeld (1989). Namely, they introduced the concept of a "noisy straight line segment" based on least squares line fitting and defined in terms of bounds on correlation coefficients and show that it is a generalization of a digital straight line. The following question is natural: If a continuos line is digitized and the least squares method is applied to the points of the digital line segment, can the original line be recovered? The answer is positive. Moreover, it can be shown that there exists one-to-one correspondence between the digital polynomial curve segments of a given degree and the corresponded least squares polynomials fit. Such a correspondence preserves a constant space representation of digital objects which result from the digitalization of polynomial curves of an arbitrary degree. The above result can be extended to a constant space coding scheme for sets consisting of digital curves whose "original curves" are graphs of explicitly given continuous functions, having the bounded number of intersection points, pairwise, on the observed interval.

edited by R.Klette. April 9, 1996