# Combinatorial Optimization and Image Analysis: a literature survey.*

Alexander Toet

December 11, 1995

*CALMA Report CALMA.TNO.WP31.AT.95a

CONTENTS Page

**Combinatorial Optimization and Image Analysis: a literature survey**

**A. Toet**

## SUMMARY

The EUCLID (EUropean Cooperation for the Long term In Defence) CALMA (Combinatorial Algorithms for Military Applications) RTP (Research and Technology Project) 6.4 project has the main objective to investigate the relevance of various existing algorithmic techniques to the actual solution of complex combinatorial problems arising in military applications.

Various approaches to combinatorial problem solving have emerged over the last three decades (e.g. simulated annealing, mean field annealing, neural nets, genetic algorithms, tabu search). Most of these have only been tested on a few simplified problems (e.g. the "Traveling Salesman Problem"). It is presently not known to what extent they can solve complex realistic problems as those arising in military appplications, and what the required computing resources are.

This report presents a selection of combinatorial problems that arise at different stages in the analysis of realistic (and military relevant) imagery. They may serve as a testbed to compare different optimization techniques. This report is intended for researchers specialized in combinatorial optimization. The combinatorial optimization techniques are therefore only briefly introduced. A brief introduction to some image processing terminology is also provided.

Only few comparative studies on the effectiveness of combinatorial optimization techniques to solve combinatorial problems in image analysis are currently available. This report gives a survey of the results of these studies.

**Combinatorische Optimalizatie en Beeldanalyse:**
**een literatuuroverzicht**

**A. Toet**

## SAMENVATTING

Het EUCLID (EUropean Cooperation for the Long term In Defence) CALMA (Combinatorial Algorithms for Military Applications) RTP (Research and Technology Project) 6.4 project heeft als voornaamste doel het onderzoeken van de bruikbaarheid van verschillende bestaande combinatorische optimalizatie technieken voor het oplossen van complexe combinatorische problemen

Er zijn de laatste dertig jaar verscheidene combinatorische optimalizatietechnieken ontwikkeld (bijv. simulated annealing, mean field annealing, neurale netwerken, genetische algoritmen, tabu search). De meeste daarvan zijn alleen getest op sterk vereenvoudigde problemen (bijv. het "handelsreizigers probleem"). Het is momenteel dan ook niet bekend in hoeverre deze methoden in staat zijn om complexe realistische problemen (zoals die in militaire omgevingen voorkomen) op te lossen, en wat de benodigde rekencapaciteit is. zoals die voorkomen in militaire omgevingen.

Verschillende stadia van de vereisen het oplossen van grote combinatorische problemen. Dit rapport beschrijft een aantal geselekteerde problemen die optreden bij de analyse van realistische (militair relevante) beelden. Deze problemen kunnen dienen als een testomgeving waarin verschillende optimalisatietechnieken met elkaar kunnen worden vergeleken. Dit rapport is bedoeld voor onderzoekers op het gebied van combinatorische optimalisatietechnieken. Deze technieken worden derhalve slechts kort geïntroduceerd. Tevens wordt er een beknopte introductie gegeven van enkele gangbare termen uit de beeldanalyse.

Er zijn tot op heden slechts enkele studies verschenen waarin wordt onderzocht wat de effectiviteit is van verschillende optimalizatietechnieken voor het oplossen van combinatorische problemen in de beeldanalyse. Dit rapport geeft ook een overzicht van de resultaten van deze studies.

# 1 INTRODUCTION

In military operations, problems in planning and scheduling often require feasible and close to optimal solutions with limited computing resources and within very short time periods. Well known examples are the weapon-target assignment problem, cross-country movement planning, multiple-target tracking, the radio link frequency assignment problem, and optimal loading problems. Various approaches to combinatorial problem solving have emerged over the last three decades. Most of these have only been tested on a few simplified problems (e.g. the "Traveling Salesman Problem"). At the moment it is not known to what extent they can solve complex realistic problems as those arising in military appplications, and what the required computing resources are.

The EUCLID (EUropean Cooperation for the Long term In Defence) CALMA (Combinatorial Algorithms for Military Applications) RTP (Research and Technology Project) 6.4 project has the main objective to investigate the relevance of various existing algorithmic techniques to the actual solution of complex combinatorial problems arising in military applications.

Image analysis has many military applications. For instance, techniques like contrast stretching, noise filtering, edge enhancement, and segmentation can be applied to improve the detectability of targets in images. They can be used to aid the human operator in surveillance tasks involving electro-optical imaging devices. The output of edge detection, segmentation and noise filtering routines also serves as input for higher level image analysis routines like automatic target recognition (ATR) algorithms.

This report presents some selected combinatorial problems from image analyis that may serve as a testbed to compare different optimization techniques. The presentation is aimed at researchers specialized in combinatorial optimization. The combinatorial optimization techniques are therefore only briefly introduced. A brief introduction to some image processing terminology is also provided.

Only few comparative studies on the effectiveness of combinatorial optimization techniques to solve combinatorial problems in image analysis are currently available. This report also gives a survey of the results of these studies.

The organization of the rest of this report is as follows. First some combinatorial optimization schemes are briefly introduced. Then some essential terminology from image analysis is presented. The following sections each present $(i)$ an introduction to some selected problems from image analysis, $(ii)$ a brief literature survey of attempts to solve these problems by combinatorial optimization techniques, and $(iii)$ the results of some comparative studies that evaluated the relative performance of different optimization techniques on the same problem. Finally, some concluding remarks and suggestions for further research are presented.

# 2 COMBINATORIAL OPTIMIZATION TECHNIQUES

This section presents a brief introduction to some optimization techniques that are frequently used in image analysis. The application of these techniques requires the definition of an objective or energy function. The energy function encodes the interaction between the variables and the system constraints. The quality of a solution is directly related to the corresponding value of the energy function, such that lower values of the energy function correspond to solutions of higher quality.

## 2.1 Local Search

The idea of a local search (LS) algorithm is to start with an initial solution and to move to nearby solutions as long as the energy function decreases. This approach is the basic principle underlying many classical optimization methods, like the gradient method for continuous nonlinear optimization or the the simplex method for linear programming (Crama *et al.*, 1992). It also explains the dynamics of many classes of neural networks, like e.g. the sequential iterations of Hopfield nets. In this latter case, the objective function corresponds to the energy (Lyaponov) function of the network, the feasible solutions are the different configurations, and two configurations are neighbours if they differ in the state of exactly one neuron (i.e. if the neuron is excited in one of the configurations and inhibited in the other). Vaessens *et al.*(1992) showed that the majority of optimization techniques that are currently around, such as iterative improvement, simulated annealing, threshold accepting, tabu search, and genetic algorithms, are special instances of a local search template. They differ in the remedies they use to avoid getting stuck in poor local optima.

## 2.2 Simulated Annealing

Annealing is a term from metallurgy. When the atoms in a piece of metal are aligned randomly, the metal is brittle and fractures easily. In the process of annealing, the metal is heated to a high temperature, causing the atoms to shake violently. If it were cooled suddenly, the the microstructure would be locked into a random unstable state. Instead, it is cooled very slowly. As the temperature drops, the atoms tend to fall into patterns that are relatively stable for that temperature. Providing that the temperature drop is slow enough, the metal will eventually stabilize into an orderly structure.

Simulated annealing (SA) is a stochastic optimization technique for solving large, multivariate, multi-modal problems. It is based on statistical mechanical techniques for calculating the equilibrium energy of molecules in solids at a constant temperature (Kirkpatrick *et al.*, 1983). It models the degrees of freedom in a problem as if they were a collection of atoms slowly being cooled into a ground state corresponding to the optimal solution to the problem. It combines gradient descent with a random process. It has the advantage over

other gradient descent techniques in that the random process or "thermal noise" can perturb the evolution of the solution to prevent it from becoming stuck in local minima.

The recently introduced mean field annealing (MFA) algorithm combines the characteristics of SA with those of the Hopfield neural network (Bout and Miller, 1990). It is equivalent to a neural network whose neural gains are gradually increased or annealed. Mean field annealing is a deterministic approximation of the average behaviour of the system's variables in simulated annealing.

### 2.3 Genetic Algorithms

Genetic algorithms (GA) are nonlinear optimization techniques (Goldberg, 1989; Holland, 1992a, Michalewicz, 1992). They employ mechanisms analogous to those involved in natural selection processes to conduct a search through a given parameter space for the global optimum of some objective function. Genetic algorithms maintain a population of candidate solutions, known as individuals. At each iteration, known as a generation, each individual is evaluated and recombined with others on the basis of its overall quality or fitness in solving the problem. The expected number of times an individual is selected for recombination is proportional to its fitness relative to the rest of the population. In other words, individuals with high fitness values have the largest chance to produce offspring (i.e. to copy their genetic code) in the next generation. This ensures the proliferation (by means of crossover) of optimal partial solutions throughout a population, and increases the population's average fitness. New individuals are created using two main genetic recombination operators known as crossover and mutation. Crossover operates by selecting a random location in the genetic string of the parents and concatenating the initial segments of one parent with the final segment of the other parent to create a new child. A second child is simultaneously generated using the remaining segments of the two parents. In practice, an arbitrary number of crossover sites and any number parents may be used. Mutation provides for occasional disturbances in crossover operation by introducing minor variations into one or more genetic elements during reproduction. Mutation serves $(i)$ to (re)introduce genetic strings and $(ii)$ to approach local extrema through hillclimbing.

Genetic algorithms are especially suited for problems involving a large number of parameters, with complex objective functions that have many local extrema. Conventional techniques like hill climbing are not likely to find a global extremum in these circumstances. By maintaining a population of solutions, genetic algorithms sample the objective function in many regions of the parameter space simultaneously. Notably, the density of sample points in a given region of the parameter space is proportional to the average value of the objective function in this region (i.e. to the fitness or quality of the solution in that vicinity). Through their ability to (re-) combine parameter strings representing partial solutions they are able to focus the search process on those parts of the solution space that are most likely to contain a global extremum of the objective function. They have been succesfully used in a wide range of applications to find global optima in very large multidimensional search spaces (Goldberg, 1989; Holland, 1992b). More recently, they have also been applied to

image analysis (Andrey and Tarroux, 1994; Ankenbrandt *et al.*, 1990; Bala and Wechsler, 1993; Hill and Taylor, 1992; Hill *et al.*, 1992).

### 2.4 Neural Nets

An artificial neural or cellular network consists of a set of interconnected simple nonlinear computational elements or formal neurons (for an introduction to neural networks we refer to Vemuri, 1988). These processing units generally work in parallel. The weights of their interconnections (synapses) are typically adapted according to learning rules. These rules specify an initial set of weights and indicate how these should be adapted during use to improve performance. Neural net models are therefore specified by the net topology, node characteristics, and training or learning rules.

Neural networks are unsupervised adaptive information processing systems. They can store information in their interconnection structure. Information retrieval and, most importantly, generalization is done by the (unsupervised) development of associations (mapppings or transformations) between stored patterns and incoming signals.

The potential benefits of neural nets extend beyond the high computation rates provided by massive parallelism. Neural nets typically provide a high degree of robustness or fault-tolerance. Damage to a few nodes or links need not impair their overall performance significantly because they generally have a large number of processing nodes, each with primarily local connections. Minor variabilities in the characteristics of their processing elements are compensated by their ability to adapt and continue learning. There has been a recent resurgence in the field of artificial neural nets caused by new net topologies, analog VLSI implementation techniques, and the need for high-speed processing in vision and speech recognition systems.

Neural architectures seem to be an efficient approach to solve the problems of scene partitioning and invariant target recognition. Neural nets have succesfully been applied to translation (e.g. Fukushima, 1980; Rosenblatt, 1961), orientation and scale (Reitboeck and Altmann, 1984; Carpenter *et al.*, 1989; Gupta and Sayeh, 1988) invariant target recognition (seee also Hecht-Nielsen, 1988).

## 3 DIGITAL IMAGE REPRESENTATIONS

A two dimensional *digital image* is represented by a function $f(x, y)$ of two variables $x$ and $y$ that represent the spatial coordinates in the image plane. The value of $f$ represents the local brightness (intensity), or $k$-tuples of brightness values in case of several spectral bands. In the black-and-white case, the values will be called *graylevels*. The range of $f$ is discrete (the brightness values are quantized), positive (brightness cannot be negative), and bounded (brightness cannot be arbitrarily great). The domain of $f$ is discrete (the brightness function is sampled), and finite, since an optical system has a limited field of view, so that the resulting image is of finite size.

An element of a digital picture is called a picture element or *pixel*. A digital image is usually represented as a list of pixels. Each pixel can be represented by a record of integers, that represent its brightness value(s) and its spatial position in the sampling grid.

Digital images can be attributed a relational or graph structure by taking the topology (connectivity) of the support grid into account. This results in a neigbourhood graph representation, in which the vertices correspond to the pixels and the connecting arcs represent the spatial (neighbourhood) relations between the pixels. The vertices can be labelled with the brightness values(s) of the pixels. The arcs can be labelled with the (Euclidian) distance between the pixels.

Digital images defined on regular (e.g. rectangular or hexagonal) sampling grids can also be represented by matrices. The entries of the matrix represent the brightness value(s) of the pixels. The indices of an element in the matrix implicitly define the spatial coordinates of the pixel in the image plane.

For an introduction on digital image processing and computer vision see: Ballard and Brown, 1982; Hall, 1979; Pratt, 1991; Rosenfeld and Kak, 1982. For a glossary of terms used in computer vision see: Haralick and Shapiro, 1991.

# 4 COMBINATORIAL PROBLEMS IN IMAGE ANALYSIS

This part presents a brief overview of some complex combinatorial problems that arise in image analysis. The following sections each present $(i)$ an introduction to some selected problems from image analysis, $(ii)$ a brief literature survey of attempts to solve these problems by combinatorial optimization techniques, and $(iii)$ the results of some comparative studies that evaluated the relative performance of different optimization techniques on the same problem.

## 4.1 Edge Detection

*Problem description*

Local discontinuities in image luminance are called *luminance edges*. Edges can be detected by applying a local difference or gradient operator to the image luminance function and thresholding the output (Pratt, 1991; Rosenfeld and Kak, 1982). The result of this transformation is a cartoon-like image, representing the locations in the original image where the luminance gradient is strong.

Edge detection is an important task in computer vision. It is the front-end processing stage in target recognition and image understanding systems. The accuracy with which this task can be performed determines the overall system performance.

Despite the tremendous amount of reserach that has been done, the task of finding the edges in an image that correspond to true physical boundaries remains a difficult problem. Part

of the difficulty lies in the fact that the output of edge detectors is usually corrupted by the effects of noise, digitization, texture, lighting etc. As a result, many "false" edge points are detected which do not lie on border regions (do not correspond to physical edges), and at many border points no edges are detected. There are many techniques that exploit information about the local structure of the neighbourhood of an edge pixel to improve the output of edge detectors. This information can for instance be used to delete noise responses (thresholding techniques), to fill edge gaps (edge linking) and to enhance the edge articulation or pronouncedness (edge thinning, see: Pratt, 1991; Rosenfeld and Kak, 1982).

*Combinatorial optimization approach*

Edge detection has recently been formulated as an optimization problem (Bhandarkar *et al.*, 1994; Geman *et al.*, 1990; Jeong and Kim, 1992; Tan *et al.*, 1992). This was achieved by formulating a cost function that evaluates the quality of edge configurations. The function is a sum of weighted cost factors. The cost factors capture desirable characteristics of edges such as accuracy in localization, thinness, and continuity. Edges are detected by finding the edge configurations that minimize the cost function. This apporach is capable of detecting many different types of edges, including intensity and texture edges.

The search space for the minimum cost solution is extremely large as the number of possible solutions equals $2^P$, where $P$ is the number of pixels in the image (usually $256^2$ or $512^2$). Optimization techniques like SA (Bhandarkar *et al.*, 1994; Geman *et al.*, 1990; Tan *et al.*, 1992), MFA (Acton and Bovik, 1992), GA (Bhandarkar *et al.*, 1994), and LS (Bhandarkar *et al.*, 1994) were used to find a global minimum of the cost function.

*Comparative studies*

Bhandarkar *et al.*(1994) performed a qualitative and quantitative comparison of LS, SA and GA optimization techniques for edge detection on real and synthetic intensity images.

An "integrated genetic algorithm" (IGA) incorporating
- the elitism strategy (Davis, 1991),
- the Engineered Conditioning (EC) operator (Potter *et al.*, 1992), and
- a knowledge-augmented mutation operator, exploiting local edge structure and biased towards the construction of valid edge segments,

was found to perform very well in terms of robustness to noise, rate of convergence, and quality of the final edge image.

SA was found to converge slowly. It took about 40–60 iterations of the algorithm to get an edge image in which the image contents can be recognized. After about 100–150 iterations the edges are thin, continuous and correctly localized.

The edge image corresponding to the best individual in the population resulting after only a few generations produced by the integrated genetic algorithm, was found to be equivalent

to the edge image obtained after many iterations of the SA algorithm. The convergence of the IGA is therefore much faster than that of the SA algorithm. A simple GA was found to perform poorly in terms of convergence rate. This was attributed to the fact that the population of 512 individuals (images) represented only a small fraction of the search space of $2^{256\times256}$ possible edge configurations (the image size was $256 \times 256$, and each pixel is either an edge pixel or a non-edge pixel). Simple GA's rely on the diversity in the existing population (the richness in the gene pool) to search for better solutions. They perform poorly when the population size does not fully represent the diversity of the underlying search space.

Convergence of the GA improved considerably after the incorporation of a knowledge-based mutation operator and meta-level genetic operators. The computational complexity of each new generation of the GA was found to be roughly comparable to a single iteration of the SA algorithm with one transformations per pixel per iteration (i.e. for each temperature value).

The LS algorithm was found to have the fastest rate of convergence. Each iteration of this was found to be 15–18 times faster than a single iteration of the SA algorithm or the GA. After only 2 iterations a recognizable edge image was obtained. After about 10 iterations the edge image was more or less stable (the algorithm converged). However, the quality of the solutions obtained with the LS algorithm is considerably lower (the minimum of the cost function is higher) than than the quality of the solutions obtained with the SA algorithm and the GA. This was attributed to the tendency of the LS algorithm to get trapped in local minima.

In dealing with noisy images SA and the GA were found to perform considerably better than LS.

### 4.2 Clustering

*Cluster analysis* is an important technique for data analysis and is used in many scientific disciplines. Cluster analysis organizes a data set by abstracting its underlying structure either as a grouping of individuals or as a hierarchy of groups.

The "shape", "size" and number of clusters that can be identified in a given data set depend on the *resolution* with which the data are analyzed. A complete hierarchical structural description represents complex large clusters as a union of simple clusters of smaller size and can be obtained by analyzing a data set over a whole *range of scales*. A *clustering* is a partition of data or targets in a set of groups or clusters such that the similarity between any two elements in a cluster is larger than the similarity between an element of the cluster and any element of a different cluster. A *hierarchical clustering* is a sequence of partitions in which each partition is nested into the next partition in the sequence.

Clustering methods require that an *index of proximity*, or alikeness, or affinity, or association be established between pairs of targets. A proximity index is either a similarity or a

dissimilarity. The more two targets resemble each other, the larger their similarity index and the smaller their dissimilarity index.

Clustering can be regarded as finding the partitioning of a set into a given number of subsets that minimizes the overall intracluster variance.

*Graph theory* is a mathematical formalism that has a multitude of applications in cluster analysis (Harary, 1969; Jain and Dubes, 1988).

### *4.2.1 Image Segmentation*

*Problem description*

Image segmentation is a clustering procedure applied to a matrix representation of the image. It typically partitions the spatial domain of an image into mutually exclusive subsets, called regions. Each region is uniform and homogeneous with respect to some property such as tone or texture, and its property value differs in some significant way from the property value of each neigbouring region. The segmentation of an image into meaningful parts is a key step in nearly every image analysis problem. It is crucial to the successful identification of the depicted targets, and to the accuracy of target analysis such as shape and area.

*Combinatorial optimization approach*

Pattern recognition and image analysis are often defined in probabilistic terms. Hence, image segmentation can be regarded as an hypothesis testing process, which merges only segments belonging to the same region. In this approach, images are modelled as realizations of random fields (Geman and Geman, 1984). For segmentation, statistical optimal estimation techniques are used, such as minimum mean squared error, maximum likelihood, and maximum a posteriori probability. Recently Markov random field and Gibbs random field image models, in combination with SA, have been used to obtain optimal segmentations (or restorations) of noisy, textured and noisy textured images (Geman and Geman, 1984; Kim and Yang, 1994; Lakshmanan and Derin, 1989; Murray and Buxton, 1987).

A segmentation can also be regarded as the piecewise polynomial approximation of the image (e.g. Beaulieu and Goldberg, 1989; Blake, 1989). The approximation error can then be employed as a global criterion, and the image segmentation then consists of finding the partition that minimizes this criterion.

Images generally have a hierarchical structure (Koenderink, 1984). This property can be used to constrain the search space. The global optimization problem can then be replaced by a sequence of optimizations at succesively higher levels of resolution. The stepwize optimization criterion minimizes the probability of merging dissimilar segments. It can be derived from the global optimization criterion (Beaulieu and Goldberg, 1989).

*Comparative studies*

No comparative studies have been performed to evaluate the performance of different optimization techniques on the image segmentation problem.

### *4.2.2 Perceptual grouping*

*Problem description*

Perceptual grouping and figure-ground discrimination are augmented (rule based) forms of image segmentation. Their goal is to detect significant structures in image data without a priori knowledge of the contents of the scene. This process of structure detection uses general rules of perceptual organization, that appear to reflect general principles of invariance under the common imaging geometries (Lowe, 1985). Gestalt psychologists undertook the first detailed study of the grouping phenomenon in human vision in the first part of this century (Koffka, 1935; Kohler, 1970; Wertheimer, 1935). They observed that the human visual system uses rules such as proximity, similarity, continuity, closure, and symmetry in analyzing images. Many existing techniques to detect global structure in images do not take information about the local adjacency of image features into account. Consequently, the detected groupings often do not correspond to physical objects in the scene.

For perceptual grouping, the collection of all possible groupings in an image is the solution space. An energy function can be defined to evaluate the quality of groupings in terms of the abovementioned visual rules. The optimal solution is defined as the collection of groupings which minimizes the total sum of the energies of all groupings in the image (McCafferty, 1990; Kahn, 1990).

The most elementary information needed to construct a global perceptual grouping is the local adjacency relationship among the individual image samples (pixels). This requires the representation of the image by a neighbourhood graph (e.g. Tüceryan and Chorzempa, 1991). At the lowest level of abstraction, images can simply be attributed a neighbourhood graph structure by taking the topology (connectivity) of the support grid into account (e.g. Kong and Rosenfeld, 1989; Kovalevsky, 1989). At higher levels of abstraction, images can be represented by a neighbourhood graph in which the vertices correspond to image primitives and the connecting arcs represent the relations between these primitives. Primitives can either be pixels (simplest case) or they can be more complex image features (e.g. edges, corners, textons, shapes or targets; see: Nacken and Toet, 1994).

Local topological and metrical relationships restrict the features that can be grouped together. For example, edge segments can be grouped $(i)$ when their endpoints are within a given radius of one another, $(ii)$ when they are somewhat collinear, and $(iii)$ when they have a similar orientation (Nacken and Toet, 1994). Even when using extensive local or semi-local topoogical and metrical relationships to reduce the solution space of possible groupings, a great number of candidate perceptual grouping arrangements still remain. The

energy function measures the extent to which each grouping conforms to a given set of grouping principles (the constraints).

*Combinatorial optimization approach*

Because the perceptual grouping energy function is typically multimodal and often discontinuous, stochastic combinatorial optimization techniques like SA (Kahn *et al.*, 1990; McCafferty, 1990), MFA (Hérault and Horaud, 1992, 1993), MA (Hérault and Horaud, 1993), and NN (Hérault and Horaud, 1993) are commonly applied to compute perceptual groupings and to perform figure-ground discrimination.

*Comparative studies*

Hérault and Horaud (1993) used the figure-ground problem as a benchmark to compare the performance of SA, MFA and MA. They found that all these methods gave the same quality of the final solution: the energy minima obtained were all the same. However, for a sequential implementation of the algorithms, they found that MFA is from 32 to 47 times faster than SA and from 3.8 to 9 times faster than MCA. This is due to the small number of iterations needed in MFA as compared with the other methods. The number of iterations was the number of tested transitions in the case of SA and MCA, and the number of spin updates in the MFA. This number is 59 to 79 times smaller in MFA than in SA and 6.8 to 11.8 times smaller in MFA than in MCA. The number of iterations in MCA is smaller than in SA because it is easily determined when the equilibrium is reached in MCA. Summarizing, it was found that MFA is the most efficient method and SA the least efficient one.

### *4.2.3 Computing partial relational image descriptions (graph partitioning)*

*Problem description*

When complex large relational image descriptions are to be matched it is convenient and sometimes even necessary to perform a piecewise comparison to make the problem tractable and/or suitable for distributed processing. This requires the partitioning of large graphs into smaller subgraphs. The interconnection cost between the subgraphs, defined as the total weight of arcs connecting vertices in distinct subsets, must be minimized to prevent the disruption of highly significant local topological relations and to ensure that only a minimum number of arcs is cut. Constrained graph partitioning is a NP-complete problem with a lot of potential applications (Garey and Johnson, 1979).

The graph partitioning problem has typically been solved by using problem-specific heuristics, branch- and bound searching, or by relaxing one or more constraints and solving the simplified problem. Methods like exhaustive search and linear programming are not adapted to problems involving large size graphs: they are inextricable. Moreover, their parallel implementation seems to be very difficult (Hérault and Niez, 1989).

*Combinatorial optimization approach*

A significant advance in finding optimal solutions for the graph partitioning problem was made with the invention of new techniques like SA and NN. The main drawback of SA is the large computing time on a sequential computer (Hérault and Niez, 1989).

*Comparative studies*

The rate of convergence of a mean field annealing algorithm on graph partitioning problems is 10 to 100 times that of SA, with nearly equal quality of solutions (Bout and Miller, 1990; Hérault and Niez, 1989). This speedup is achieved due to the fact that, like neural networks, continuous *average values* of the discrete degrees of freedom are used in the computations, and these relax to their equilibrium values at a given temperature much faster than the corresponding Markov chain employed in SA. The temperature behaviour of mean field annealing during graph partitioning posseses a critical temperature at which most of the optimization occurs. This temperature is analogous to the gain of neurons in a neural network and may be used to tune such networks for better performance.

### 4.3 Contrast Enhancement

*Problem description*

One of the most common defects of photographic or electronic images is poor contrast resulting from a reduced, and perhaps nonlinear, image amplitude range. Image contrast can often be improved by amplitude rescaling of each pixel. The choice of the (nonlinear) remapping function is task-dependent and subjective. Measures like entropy, compactness, and area coverage can be used to objectivate the evaluation (e.g. Ballard and Brown, 1982; Hall, 1979; Pratt, 1991; Rosenfeld and Kak, 1982). Entropy of an image considers the global information content and provides a measure for the average amount of fuzziness in grayness, i.e. the degree of difficulty (ambiguity) in deciding whether a pixel should be considered bright or dark. Compactness on the other hand is a local information measure that reflects the amount of fuzziness in shape and geometry of an image. The local and global measures can be combined into a single overall evaluation function that measures image quality in both respects. Finding the optimal mapping is a large combinatorial problem because of the large number of parameters involved. Also, the search space has many local optima.

*Combinatorial optimization approach*

Pal *et al.*(1994) recently introduced a GA that estimates the optimal contrast remapping. The convergence of the algorithm was experimentally verified for the images of both compact and elongated objects. The algorithm on average only needs about 3000 evaluations (out of a search space of $2^{120}$ points) to converge to an acceptable solution.

*Comparative studies*

No comparative studies have been performed to evaluate the performance of different optimization techniques on the image contrast enhancement problem.

## 4.4 Target Recognition

Model-based target recognition involves the detection and classification of instantiations of predetermined patterns or target models in an image. This involves the estimation of the parameters of a model -to- image transformation that results in the observed image. For example, for a 2D image of a known rigid 3D target the goal is to find the (position- and viewpoint-) parameters of the model which best explain the target projection observed in the image. These parameters can be obtained by means of an iterative generate -and- test procedure which moves from an initial solution via transformations to one that optimally represents the data. The extent to which a model projection and the target contours in the image spatially coincide is called the *fit*. A *match* is a model projection whose fit has been optimized.

Images are frequently degraded by sensor noise, target occlusion and non-uniform illumination. Consequently, images do not necessarily give a correct or complete representation of the targets in the scene. As a result, perfect matches between target models and image representations seldom occur, and partial matches must be considered. Also, models are usually oversimplified and provide only a partial description of image shapes. As a result, different (noisy) image shapes may result in equally good fits. A *globally optimal match* is a model instantiation with the lowest match error of all possible model instantiations (i.e. which is optimal over the entire search space). A *locally optimal match* is a model instantiation with the lowest match error of all elements in a subset of all possible model instantiations (i.e. which is optimal in a restricted neighbourhood of the search space). Generally, a *best match* will refer to a locally optimal match that is likely, but not guaranteed, to be globally optimal.

The model-based target recognition task can be formulated as a combinatorial optimization problem. An objective function can be defined which measures the evidential support for any particular projection of a parametrized target model onto the input image. A combinatorial optimization algorithm can then be employed to find a set of parameters which provide a good (though not necessarily optimal) interpretation of the image in terms of the model.

Bengtsson (1990) presented a matching scheme based on a stochastic optimization method that is related to SA. It was applied to match a template contour of an target to edges in a digital image, using Borgefor's (1984b) r.m.s. distance measure as an evaluation function. The method was limited to a three parameter space: two translation parameters and one rotation parameter. The size and shape of the template contour were fixed. In a pilot test, the stochastic method was succesful in 4 out of 8 trials, while a deterministic approach failed in all cases to find the correct match.

### *4.4.1 Matching relational image descriptions (graph matching)*

*Problem description*

Target recognition in graph representations generally involves the detection of local instatiations of the target's graph representation. This can be done by placing the (relatively small and simple) graph representation of the target at every vertex of the graph and monitoring the (partial) matches. Procedures like these are intricate and computationally expensive. There exists no known polynomial time algorithm for finding subgraph isomorphisms, despite the fact that the problem has not been shown to be NP-complete.

In military operations, the speed of the target recognition, interpretation and manipulation process is important. Hence, the data structures should $(i)$ enable the flexible encoding of a vast variety of graphs (e.g. non-planar graphs and oriented graphs), and $(ii)$ allow the efficient computation of graph transformations. Appropriate data structures can be derived from the adjacency matrix representation (Gondran and Minoux, 1984; Heijmans *et al.*, 1992).

The subgraph isomorphism problem needs exhaustive computation if exact algorithms are used.

*Combinatorial optimization approach*

Constrained continuous optimization techniques, such as relaxation labeling and neural network strategies, can solve graph matching and recognition problems within reasonable time, even for rather complex relational structures for which heuristics can easily fail (Dreyfus and Zippelius, 1989; Kree and Zippelius, 1988; Kuner and Ueberreiter, 1988; Mjolsness *et al.*, 1989; von der Malsburg, 1988). Hierarchical neural nets, in which there is a two-way passage of information between different levels of abstraction, are particularly efficient in solving the abovementioned problems (Mjolsness *et al.*, 1991). These methods are well suited for parallelism. Moreover, associative networks can also handle incomplete and distorted input graphs.

An ongoing research topic is the development of an appropriate neural architecture to represent graphs and to implement graph matching. It has been proposed to represent both the nodes and the arcs in a graph by neurons, such that there are node-neurons and edge-neurons (Hopfield and Tank, 1986; Kree and Zippelius, 1988; Philips *et al.*, 1988). Nodes represent structural primitives and may be labeled with the feature type. Arcs represent the topological relations between structural primitives. An alternative is to represent the arcs by dynamical connections between node-neurons and by the temporal correlations between their signals (Bienenstock, 1988; Bienenstock and von der Malsburg, 1987; von der Malsburg, 1985; von der Malsburg and Bienenstock, 1986, 1987). The advantage of the latter architecture is its parsimony in terms of network structure. It performs graph matching not because it has been specifically wired in, but because it is inherent in the network dynamics.

*Comparative studies*

No comparative studies have been performed to evaluate the performance of different optimization techniques on the graph matching problem.

### *4.4.2 Contour Matching*

*Problem description*

In stead of matching the full representation of a 2D model to the image data (as in template matching), it is also possible to match the model's outlines (contour) to the available image edges. This can be done by $(i)$ placing the model contour at every location in the image, and by $(ii)$ measuring the coincidence of model edges and image edges. The distance transformation is highly suitable for this purpose (Borgefors, 1988; Liu and Srinath, 1990). It converts a binary edge image into a distance image, which is a graylevel image in which all pixels have a value that corresponds to the distance to the nearest edge pixel.

An efficient method to measure the fit or distance between two sets of edges is to $(i)$ project the first set of edges onto the distance transform of the second set, and to $(ii)$ compute an average of the values of the pixels in the distance image that coincide with the edge projections. The average value thus found is called the edge distance. A perfect fit between two sets of edges results in an edge distance zero, since each pixel from of the projected model contour will be at zero distance from an image contour.

Contour matching is generally computationally expensive, since it involves the generation and evaluation of a large number of different model instances. The best match is found by minimizing the evaluation function. Unless the matching situation is very simple, this function will have many local minima apart from the global one. To find a global minimum, the search has to be started close to the optimal match position. Unless this position is known, searches must be started in a large number of different positions, so that at least one is close to the optimal location. Each search is rather costly in computing resources.

*Combinatorial optimization approach*

Bengtsson (1990) presented a contour matching scheme based on a stochastic optimization method that is related to SA. It was applied to match a template contour of a target to edges in a digital image, using an edge distance measure as an evaluation function. The method was limited to a three parameter space: two translation parameters and one rotation parameter. The size and shape of the template contour were fixed. In a pilot test, the stochastic method was succesful in 4 out of 8 trials, while a deterministic approach failed in all cases to find the correct match.

Friedland and Rosenfeld (1992) presented a compact target recognition scheme based on a cyclic 1D Markov Random Field polar model representation. The evaluation function

consists of $(i)$ a low-level part that restricts the model's shape (contour smoothness) and measures the local edge strength in the image (i.e. the evidence from the image data), and $(ii)$ a high-level part that matches a model instantiation to a library of existing target configurations. The low-level part initially serves to construct an initial approximation of the target's contour. At a later stage in the optimization process, as its confidence in its match increases, the high-level part (the selection of an appropriate model type) becomes more dominant, and the low-level part ensures that the high-level part does not create artifacts that are not corroborated by the image data. SA was used to find a best match. The center of the model must be initialized at a location within the target that is to be classified. To avoid detection errors due to incorrect initialization, it was suggested to divide the image into many overlapping regions, and start the recognition scheme in each region. It was found that the algorithm converged well, even for very noisy data, and required a relatively short convergence time due to the use of a 1D contour model (in stead of the full 2D template).

Toet (1994) recently introduced a genetic contour matching algorithm that involves 3 stages. First, a binary (cartoon-like) representation of the input image is created by thresholding the output of a simple difference operator. Second, this binary edge image is transformed into a distance image. Third, a GA minimizes an average of the values of the pixels in the distance image that coincide with the projection of an instantiation of a parameterized 2D contour model. The contour model employed represents the hull of the star-shaped 2D projection of 3D compact targets, so that hidden edges are no problem. The objective function evaluates the degree to which local image evidence supports any particular model instance. The GA first creates an initial population of random model instances. By $(i)$ sampling the objective function in many regions of the parameter space simultaneously and by $(ii)$ focussing the search process on those parts of the solution space that are most likely to contain a global extremum of the objective function, the GA converges to a solution. The result is a best fit of the parameterized contour model to the edge image. In a pilot study it was shown that the outcome of the GA is an instantiation of the contour model which closely agrees with the image data. In this scheme, the size and shape of the model contour adapt to the local image evidence. Also, the method is less likely to suffer from initialization sensitivity, since the GA generates model instances at all possible locations in the image plane, while avoiding being trapped in local minima. In a companion report (Toet, 1994) it is suggested that this contour matching scheme may be a suitable testbed for the evaluation of the relative performance of the optimization techniques studied in the CALMA RTP 6.4 project.

*Comparative studies*

No comparative studies have been performed to evaluate the performance of different optimization techniques on the contour matching problem.

*Visual reconstruction* is the reduction of noisy visual data to stable descriptions. Piecewise reconstruction can be formulated in terms of nonconvex optimization problems. Both stochastic and deterministic (relaxation) algorithms have been used for visual reconstruction with discontinuities.

Blake (1989) provides a comparison of graduated nonconvexity (GNC, see: Blake, 1983; Blake and Zisserman, 1987) and three variants of the SA algorithm, using the reconstruction of 1D signals as a benchmark. The SA variants investigated were respectively the "heatbath" (HB, Geman and Geman, 1984), the "Metropolis-heatbath" [MHB, a hybrid of the Metropolis algorithm (1953) and the heatbath algorithm], and "mixed-annealing" (MA, a hybrid of a deterministic update rule and the Metropolis algorithm). The performance of the three stochastic algorithms depends on the relative noise in the data. At low noise levels, the MA algorithm fails to converge to a stable state at large scales of the reconstruction. At high noise levels, the HB and MHB algorithms fail at large scales, and the MA algorithm fails at all scales. MA appears therefore less powerfull than HB or MHB. MHB is a little more efficient than HB. Blake therefore compared the MHB with the GNC algorithm. Comparative results were obtained as a function of both the scale of the reconstruction and the level of noise in the data. Variation of scale, at moderately high noise levels, shows that up to a certain critical scale MHB requires between 10–20 times more iterations than GNC does. Beyond the critical scale MHB fails altogether to find a solution (within 8000 iterations).

Summarizing, for reconstruction the three stochastic algorithms proved considerably less efficient than the deterministic GNC algorithm. Furthermore, they are considerably less powerful in that they cannot deliver correct solutions for problems involving moderately high levels of noise (and which therefore require large scale reconstructions). These conclusions hold both for serial and parallel implementations of the algorithms.

Heitz *et al.*(1994) recently introduced a new multiscale relaxation algorithm to find the global minimum of an objective function. The global optimization problem is solved within a sequence of nested subspaces of the original space of possible solutions. Those subspaces contain constrained configurations describing the expected solutions at different scales. Each subspace defines a new "coarse" energy function whose parameters are derived from the original (full resolution) objective function. They compared this multiscale relaxation (MSR) algorithm to standard monoresolution, multigrid (MGR), deterministic (DR) and stochastic (SR) relaxation schemes, using a highly nonlinear optimization problem as a benchmark. The issue was to estimate long-range motion in an image sequence. The problem is solved on a discrete state space by minimizing a global energy function which involves two components. The first corresponds to a similarity function that matches brightness patterns. The second term is a standard regularization term. This problem is known to generate a very complex energy landscape (Konrad and Dubois, 1988), and is therefore a good benchmark for comparing different methods. It appeared that SR leads to the lowest energy values but requires a large number of iterations to converge. The new

MSR algorithm yielded results that were close to those provided by SR, with a gain of almost two orders of magnitude in the number of iterations. The standard methods (DR, MGR) became stuck at rather high energy local minima. In some cases, SR performed even worse than MSR. A very significant gain in the cpu time was also observed for the MSR method, compared to all other methods studied. This is due to a faster convergence at the finest resolution were most of the cpu time is consumed. For a large number of different test sequences, it was found that, on average, the MSR method finds solutions close to the best estimates obtained by SR, with an average gain of about 60, both on the number of iterations, and on the cpu time.

# 6 CONCLUDING REMARKS

A large number of problems in image analysis have been formulated as optimization problems. The formulation of a problem as an optimization problem reveals our lack of understanding the physical principles of the image formation process, and is definitely not desirable for many practical purposes. However, in many cases it is so far the only possibility. Therefore, it is of great importance to have good and efficient optimization techniques, that produce high-quality solutions in short time.

*Current status*

Of all modern optimization techniques, *simulated annealing* is most frequently applied to solve large combinatorial problems in image analysis. It produces good quality solutions, but converges rather slowly (Bhandarkar *et al.*, 1994; Hérault and Horaud, 1993; Hérault and Niez, 1989).

*Genetic algorithms* produce good quality solutions when the population size fully represents the diversity of the underlying search space. They converge faster when the local search (mutation) procedure is constrained by a priori knowledge of the image structure (Bhandarkar *et al.*, 1994).

*Local search* generally yields the fastest rate of convergence, but also produces the lowest quality solutions because the algorithms easily get trapped in local minima (Bhandarkar *et al.*, 1994).

*Mean field annealing* is considerably more efficient than simulated annealing, and produces nearly the same quality of the final solution (Bout and Miller, 1990; Hérault and Horaud, 1993; Hérault and Niez, 1989).

*Hierarchical neural nets* are particularly efficient and robust in graph matching and recognition (Mjolsness *et al.*, 1991).

*Deterministic methods* proved considerably more efficient and robust than stochastic methods in visual reconstruction (Blake, 1989) and long-range motion estimation (Heitz *et al.*, 1994).

*Further research*

Efficient image analysis routines can only be developed when the physical principles underlying the image formation process are fully known. Hence, further research is needed to increase our understanding of these principles.

In the meantime, the relative efficiency and applicability of different optimization techniques should be established. Also, it should be investigated in which cases newly developed deterministic methods (Blake, 1989; Heitz *et al.*, 1994) are more efficient and robust than stochastic methods.

# REFERENCES


Acton, S.T. and Bovik, A.C. 91992). Anisotropic edge detection using mean field annealing. Proc. IEEE Intl. Conf. Acoustics, Speech and Signal Processing, vol. II, pp. 393–396.

Andrey, P. and Tarroux, P. (1994). Unsupervised image segmentation using a distributed genetic algorithm. Pattern Recognition 27, pp. 659–673.

Ankenbrandt, C.A., Buckles, B.P. and Petry, F.E. (1990). Scene recognition using genetic algorithms with semantic nets. Patt. Rec. Lett. 11, pp. 285–293.

Bala, J. and Wechsler, H. (1993). Shape analysis using genetic algorithms. Pattern Recognitin Letters 14, pp. 965–973.

Ballard, D.H. and Brown, C.M. (1982). Computer Vision, Prentice-Hall, Englewood Cliffs, New Jersey.

Beaulieu, J.-M. and Goldberg, M. (1989). Hierarchy in picture segmentation: a stepwise optimization approach. IEEE Tr. on Pattern Analysis and Machine Intelligence PAMI-11, pp. 150–163.

Bengtsson, M. (1990). Stochastic optimization algorithms — an application to pattern matching. Pattern Recognition Letters 11, pp. 717–724.

Bhandarkar, S., Zhang, Y. and Potter, W. (1994). An edge detection technique using genetic algorithm-based optimization. Pattern Recognition 27, pp. 1159–1180.

Bienenstock, E. (1988). Neural-like graph matching techniques for image-processing. In: Organisation of neural networks, von Seelen, W., Shaw, G. and Leinhos, U. (eds.), pp. 211–238. VCH Verlag, Weinheim.

Bienenstock, E. and von der Malsburg, C. (1987). A neural network for invariant pattern recognition. Europhysics Letters 4, pp. 121–126.

Blake, A. (1989). The least disturbance principle and weak constraints. Pattern Recognition Letters 1, pp. 393–399.

Blake, A. (1989). Comparison of the efficiency of deterministic and stochastic algorithms for visual reconstruction. IEEE Tr. on Pattern Analysis and Machine Intelligence PAMI-11, pp. 2–12.

Blake, A. and Zisserman, A. (1987). Visual Reconstruction. MIT Press, Cambridge, MA.

Bout, D.E. van den, and Miller, T.K. (1990). Graph partitioning using annealed neural networks. IEEE Tr. Neural Networks 1, pp. 192–203.

Carpenter, G.A., Grossberg, S. and Mehanian, C. (1989). Invariant recognition of cluttered scenes by a self-organizing ART architecture: CORT-X boundary segmentation. Neural Networks 2, pp. 169-181.

Crama, Y., Kolen, A. and Pesch, E. (1992). Local search in combinatorial optimization. Report RM 92-020, Faculty of Economics, Limburg University, Maastricht, The Netherlands.

Davis, L. (1991). A genetic algorithms tutorial. In: Handbook of Genetic Algorithms, Davis, L. (ed.), Van Nostrand Reinhold, New York.

Dreyfus, G. and Zippelius, A. (1989). Graph recognition by neural networks. In: Neural networks: from models to applications, Personnaz, L. and Dreyfus, G. (eds.), pp. 483–492. I.D.S.E.T., Paris.

Friedland, N.S. and Rosenfeld, A. (1992). Compact object recognition using energy-function-based optimization. IEEE Tr. on Pattern Analysis and Machine Intelligence PAMI-14, pp. 770–781.

Fukushima, K. (1980). Neocognitron: a self-organizing neural network model for a mechanism of

pattern recognition unaffected by shift in position. Biol. Cybern. 36, pp. 193-202.
Garey, M.R. and Johnson, D.S. (1979). Computers and intractability - A guide to the theory of NP-completeness. W.H. Freeman, San Fransisco, CA.
Geman, S., Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. IEEE Tr. on Pattern Analysis and Machine Intelligence PAMI-6, pp. 721–741.
Geman, D., Geman, S., Graffigne, C., and Dong, P. (1990). Boundary detection by constrained optimization. IEEE Tr. on Pattern Analysis and Machine Intelligence PAMI-12, pp. 609–628.
Goldberg, D.E. (1989). Genetic algorithms in search, optimization, and machine learning. Addison-Wesley, New York.
Gondran, M. and Minoux, M. (1984). Graphs and algorithms. John Wiley & Sons, Chicester.
Grimson, W.E. (1991). The combinatorics of heuristic search termination for object recognition in cluttered environments. IEEE Tr. on Pattern Analysis and Machine Intelligence PAMI-13, pp. 920–935.
Gupta, L. and Sayeh, M. (1988). Neural networks for planar shape classification. In: IEEE Proc. Int. Conf. Acoust. Speech and Signal Proc. ICASSP-88, pp. 936-939.
Hall, E.L. (1979). Computer Image Processing. Academic Press, New York.
Haralick, R.M. and Shapiro, L.G. (1991). Glossary of computer vision terms. Pattern Recognition 24, pp. 69–93.
Harary, F. (1969). Graph Theory. Addison-Wesley, New York.
Hecht-Nielsen, R. (1988). Neurocomputing: picking the human brain. In: Artificial neural networks: theoretical concepts. Vemuri, V. (ed.). IEEE Comp. Soc. Press, Washington.
Heitz, F., Perez, P. and Bouthemy, P. (1994). Multiscale minimization of global energy functions in some visual recovery problems. CVGIP: Image Understanding 59, pp. 125–134.
Heijmans, H.J.A.M., Nacken, P., Toet, A. and Vincent, L. (1992). Graph morphology. J. of Visual Comm. and Im. Representation 3, pp. 24-38.
Hérault, L. and Horaud, R. (1993). Figure-ground discrimination by mean field annealing. In: Computer Vision — ECCV '92, Sandini, G. (Ed.), Springer Verlag, Berlin.
Hérault, L. and Horaud, R. (1993). Figure-ground discrimination: a combinatorial optimization approach. IEEE Tr. on Pattern Analysis and Machine Intelligence PAMI-15, pp. 899–914.
Hérault, L. and Niez, J.J. (1989). Neural networks and graph K-partitioning. Complex Systems 3, pp. 531-575.
Hill, A., Taylor, C.J. (1992). Model-based image interpretation using genetic algorithms. Image and Vision Computing 10, pp. 295–300.
Hill, A., Taylor, C.J. and Cootes, T. (1992). Object recognition by flexible template matching using genetic algorithms. In: Computer Vision - ECCV'92, Sandini, G. (Ed.), pp. 852–856. Springer Verlag, Berlin.
Holland, J.H. (1992a). Adaptation in natural and artificial systems. The MIT Press, Cambridge, Mass.
Holland, J.H. (1992b). Genetic algorithms. Scientific American vol. 267, pp. 44–50.
Hopfield, J.J. Neural networks and physical systems with emergent collective computational abilities. Proc. Nat. Acad. Sci. U.S. 79, pp. 2554–2558.
Hopfield, J.J. and Tank, D.W. (1986). Collective computation with continuous variables. In: Disordered systems and biological organization, Bienenstock, E., Fogelman Solié, F. and Weisbuch, G. (eds.), NATO ASI Series, vol. F20, pp. 155–170. Springer Verlag, Berlin.
Horaud, R. Veillon, F. and Skordas, T. (1990). Finding geometric and relational structures in an image. In: Computer Vision - ECCV 90. Faugeras, O. (ed.). Springer Verlag, New York.
Jain, A.K. and Dubes, R.C. (1988). Algorithms for clustering data. Prentice Hall. Englewood Cliffs,

New Jersey.
Jeong, H. and Kim, C.I. (1992). Adaptive determination of filter scales for edge detection. IEEE Tr. on Pattern Analysis and Machine Intelligence PAMI-14, pp. 579–585.
Johnson, J.H. (1989). Pixel parts and picture wholes. In: From pixels to features. Simon, J.C. (ed.), pp. 347-360. North Holland. New York.
Kahn, P., Winkler, A., andChong, C.Y. (1990). Perceptual grouping as energy minimization. In: Proc. 1990 IEEE Int. Conf. on Systems, Man and Cybernetics, IEEE Computer Soc. Press.
Katz, A.J. and Thrift, P.R. (1994). Generating image filters for target recognition by genetic learning. IEEE Tr. on Pattern Analysis and Machine Intelligence PAMI-16, pp. 906–910.
Kim, I.Y. and Yang, H.S. (1994). A systematic way for region-based image segmentation based on Markov random field model. Pattern Recognition Letters 15, pp. 969–976.
Kirkpatrick, S., Gelatt, C. and Vecchi, M. (1983). Optimization by simulated annealing. Science 220, pp. 671–680.
Koenderink, J.J. (1984). The structure of images. Biological Cybernetics 50, pp. 363–370.
Koffka, K. (1935). Principles of Gestalt psychology. Hartcourt Brace, New York.
Kohler, W. (1970). Gestalt psychology. Liveright Publ. Corp. , New York.
Kong, T.Y. and Rosenfeld, A. (1989). Digital topology: introduction and survey. Comp. Vis. Graph. Im. Proc. 48, pp. 357–393.
Konrad, J. and Dubois, E. (1988). Multigrid Bayesian estimation of image motion fields using stochastic relaxation. In: Proc. 2nd Int. Conf. Computer Vision, Tarpon Springs, Florida, pp. 354–362.
Kovalevsky, V.A. (1989). Finite topology as applied to image analysis. Comp. Vis. Graph. Im. Proc. 46, pp. 141–161.
Kree, R. and Zippelius, A. (1988). Recognition of topological features of graphs and images in neural networks. J. Phys. A: Math. Gen. 21, pp. L813–L818.
Kuner, P. and Ueberreiter, B. (1988). Pattern recognition by graph matching — combinatorial versus continuous optimization. Int. J. Patt. Rec. Art. Int. 2, pp. 527–542.
Lakshmanan, S. and Derin, H. (1989). Simultaneous parameter estimation and segmentation of Gibbs random fields using simulated annealing. IEEE Tr. on Pattern Analysis and Machine Intelligence PAMI-11, pp. 799–813.
Lowe, D.G. (1985). Perceptual organization and visual recognition. Kluwer Academic Publ., Boston.
Malsburg, C. von der (1985). Nervous structures with dynamical links. Berichte der Bunsengesellschaft für Physikalische Chemie 89, pp. 703–710.
Malsburg, C. von der(1988). Pattern recognition by labeled graph matching. Neural Networks 1, pp. 141-148.
Malsburg, C. von der, and Bienenstock, E. (1986). Statistical coding and short-term synaptic plasticity: A scheme for knowledge representation in the brain. In: Disordered systems and biological organization, Bienenstock, E., Fogelman Solié, F. and Weisbuch, G. (eds.), NATO ASI Series, vol. F20, Springer Verlag, Berlin.
Malsburg, C. von der, and Bienenstock, E. (1987). A neural network for the retrieval of superimposed connection patterns. Europhysics Letters 3, pp. 1243–1249.
McCafferty, J.D. (1990). Human and machine vision — computing perceptual organization. Ellis Horwood Ltd. Chicester, UK.
Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E. (1953). Equation of state calculations by fast computing machines. J. Chem. Phys. vol. 6, p. 1087.
Michalewicz, Z. (1992). Genetic Algorithms + Data Structures = Evolution Programs. Springer

Verlag, New York.
Mjolsness, E., Gindi, G. and Anandan, P. (1989). Optimization in model matching and perceptual organization. Neural Computation 1, pp. 218–229.
Mjolsness, E., Garrett, C.D. and Miranker, W.L. (1989). Multiscale optimization in neural nets. IEEE Tr. on Neural Networks 2, pp. 263–274.
Murray, D.W. and Buxton, B.F. (1987). Scene segmentation from visual motion using global optimization. IEEE Tr. on Pattern Analysis and Machine Intelligence PAMI-9, pp. 220–228.
Nacken, P.M. and Toet, A. (1994). Model-based bottom-up grouping of geometric image primitives. In: O, Ying-Lie, Toet, A., Heijmans, H.J.A.M., Foster, D.H. and Meer, P., SHAPE IN PICTURE - Mathematical descriptions of shape in greylevel images. NATO ASI Series F, Vol. 126, Springer Verlag, Berlin, pp. 549–558.
Pal, S.K., Bhandari, D. and Kundu, M.K. (1994). Genetic algorithms for optimal image enhancement. Pattern Recognition Letters 15, pp. 261–271.
Philips, W.A., Hancock, P.J.B., Willson, N.J. and Smith, L.S. (1988). On the acquisition of object concepts from sensory data. In: Neural computers, Eckmiller, R. and von der Malsburg, C. (eds.), NATO ASI Series, vol. F41, pp. 159–168. Springer Verlag, Berlin.
Potter, W.D., Miller, J.A., Tonn, B.E., Gandham, R.V. and Lapena, C.N. (1992). Improving the reliability of heuristic multiple fault diagnostics via the EC-based genetic algorithm. J. Appl. Intell. 2, pp. 5–19.
Pratt, W.K. (1991). Digital image processing, 2nd ed., Wiley, New York.
Reitbroeck, H.J. and Altmann, J. (1984). A model for size- and rotation- invariant pattern processing in the visual system. Biol. Cybern. 51, pp. 113-121.
Rosenblatt, F. (1961). Principles of neurodynamics: perceptrons and the theory of brain mechanisms. Spartan Books, Washington.
Rosenfeld, A. and Kak, A.C. (1982). Digital Picture Porcessing, 2nd. ed., Vol.1 and 2, Academic Press, New York.
Sanfeliu, A. and Fu, K.S. (1983). A distance measure between attributed relational graphs for pattern recognition. IEEE Tr. Syst. Man Cybern., vol. SMC-13, pp. 353–362.
Rosenfeld, A. and Kak, A.C. (1982). Digital Picture Porcessing, 2nd. ed., Vol.1 and 2, Academic Press, New York.
Tan, H.L., Gelfand, S.B. and Delp, E.J. (1992). A cost minimization appproach to edge detection using simulated annealing. IEEE Tr. on Pattern Analysis and Machine Intelligence PAMI-14, pp. 3–18.
Toet, A. (1991) Hierarchical clustering through morphological graph transformation. Pattern Rec. Letters 12, pp. 391–399.
Toet, A. (1994). Target detection and recognition through contour matching. TNO-report TNO-TM 1994 A-51. CALMA Report no. CALMA.TNO.WP31.AT.95b
Tüceryan, M. and Chorzempa, T. (1991). Relative sensitivity of a family of closest-point graphs in computer vision applications. Pattern Rec. 24, pp. 361–373.
Vaessens, R.J.M., Aarts, E.H.L., and Lenstra, J.K. (1992). A local search template. In: Männer, R. and Manderick, B. (Eds.), Parallel Problem Solving from Nature, Elsevier, New York.
Vemuri, V. (1988). Artificial neural networks: theoretical concepts. IEEE Comp. Soc. Press, Washington.
Wertheimer, M. (1935). Laws of organization in perceptual forms. In: A source book of Gestalt

psychology, Ellis, W.D. (ed.), pp. 71–88. Hartcourt Brace, New York.

Soesterberg,  December 11, 1995

Dr. A. Toet

30 woorden abstract

Er werd een literatuuronderzoek gedaan naar de bruikbaarheid van moderne optimzalizatietechnieken bij het oplossen van problemen in de beeldanalyse.

Simulated annealing (een methode die het proces simuleert waarbij atomen in een verhit stuk metaal stevige verbindingen kunnen vormen door een minimale energietoestand in te nemen tijdens de langzame afkoeling van het materiaal) wordt het meest toegepast, en levert goede resultaten, ten koste van veel rekentijd. Mean field annealing (een deterministische versie van simulated annealing) blijkt echter in de meeste gevallen aanzienlijk sneller te werken, en levert vergelijkbare resultaten.

Recentelijk ontwikkelde deterministische methoden werken mogelijk in veel gevallen efficienter en robuuster dan stochastische methoden.

*Descriptors*

Image Processing
Pattern Recognition

*Identifiers*

optimization