Unsupervised Image Segmentation by Stochastic Reconstruction

Volker Metzler[1], Ralf Vandenhouten[2], Jörg Krone[3], Reinhard Grebe[4]

[1]Institute of Medical Informatics and [2]Institute of Physiology
Aachen University of Technology (RWTH), D–52074 Aachen, Germany

[3]Iserlohn Polytechnical School (MFH), D–58644 Iserlohn, Germany

[4]Génie Biologique, Technical University of Compiegne (UTC), F–60205 Compiegne, France

ABSTRACT

To segment complex and versatile image data from different modalities it is almost impossible to achieve satisfying results without the consideration of contextual information. In this approach, image segmentation is regarded as a high–dimensional optimization task, that can be solved by stochastical methods like evolutionary algorithms (EA). Initially, the iterative algorithm is provided with a set of good–quality sample segmentations. An efficient EA–based learning strategy generates a segmentation for a given target image from the provided samples. This two–level process consists of a global image–based optimization whose convergence is enhanced by locally operating pixel–based Boltzmann processes which restrict the search space to reasonable subsets. The stochastic reconstruction extracts the relevant information from the samples in order to adapt it onto the current segmentation problem, which results in a consistent labeling for the target image. The algorithm works unsupervised, because the range of possible labels and their contextual interpretation is provided implicitly by the sample segmentations. To prove the usefulness of the method experimental results based on both, reproducible phantom images and pysiological NMR scans are presented. Moreover, an analysis of the basic segmentation and convergence properties is provided.

Keywords: Image segmentation, unsupervised learning, stochastic optimization, evolutionary algorithm, Boltzmann process, classification, stochastic reconstruction

1. INTRODUCTION

The *automatic segmentation* of a given image is one of the central tasks in image processing. Because of the complexity of the problem, due to the huge amount of different image domains and user requirements, a method capable to segment all kinds of images with satisfying quality has not yet been developed.[1] Hence, to achieve good results, sophisticated segmentation techniques will usually focus on a clearly defined image domain to make use of *a–priori information*.[2] Therefore, certain image properties, such as topological or morphological features, occuring in all images of the domain, can be presupposed and utilized for the segmentation process.

The way of incorporating such contextual properties is crucial to the efficiency of the method. The simplest technique is to code the a–priori information directly in the algorithm, but this leads to rather inflexible methods that cannot adapt to trends, the coded features may be subject to. More advanced approaches will perform some kind of learning of characteristic features, which results in adaptive methods.[3] Learning strategies such as neural nets often require special training phases, where the characteristics are extracted from sample feature vectors.[4]

Alternatively, *knowledge–based* techniques can be used to make contextual information available to the algorithm.[5,6] In conjunction with a problem independent inference strategy a general problem solving concept can be set up, that does require neither feature extration nor a training phase. Besides the classic logic–based inference mechanisms there were other approaches introduced.[7,8] The idea to adapt reliable information to a new segmentation task by deformable models leads to the use of optimization techniques.[9,10]

Send correspondence to: Volker Metzler, Institut für Medizinische Informatik, RWTH Aachen, D-52057 Aachen, Germany
E–mail:metzler@imib.rwth-aachen.de, Phone:+49-241-80-88382, FAX:+49-241-8888-426

In fact, segmentation may be interpreted as a *high–dimensional optimization task*. Such problems cannot be solved by conventional analytic approaches in reasonable time, but may be tackled by stochastic optimization techniques such as *Evolutionary Algorithms* (EA)[11,12] or *Simulated Annealing* (SA).[13] On the other hand, an evolutionary process can be regarded as an inference system that generates new information in order to adapt it to a global target on the basis of a generation–and–selection process.

Recently, some researchers have focused on segmentation by means of EA techniques. One approach is to employ evolutionary optimization as a meta strategy to adapt external parameters to various image domains and conditions,[14,15] or to guide given segmentation techniques by optimizing their results according to a global quality measure.[16] Another strategy is to code the image segmentation requirements directly into the target function to be optimized.[17,18] A variant of the EA–framework, the *Classifier Systems*, represent an algorithmic link between evolutionary optimization and dynamic rule–based knowledge generation.[19] They have been applied to image segmentation recently.[20]

Compared to these concepts our method follows a different way. The target function is set up problem–independent by measuring the level of reconstruction of the target image from a set of provided segmentation samples. Therefore, the actual segmentation is generated indirectly via optimization of the graylevel reconstruction process.

Our two–level approach consists of a global image–based optimization technique (Sec. 2) based on an evolutionary algorithm, which is supported by a local pixel–based manipulation operator (Sec. 3), directed by a *Boltzmann decision process*.[21] The knowledge base contains examples of high–quality segmentations from the target's domain. An efficient learning strategy extracts the relevant information in order to adapt it to the current segmentation problem. The idea is to reconstruct the target image from the pixels of the samples whereby a consistent labeling is generated implicitly. The algorithm works unsupervised, because the possible range of labels and their contextual interpretation is provided implicitly by the sample segmentations.

The interaction between the two optimization levels is a central aspect of the algorithm. The global process generates and assesses new potential solutions, while the local process restricts the functionality of the image manipulation to improve the convergence by reducing the search space. Therefore, the occurence of an actual pixel manipulation depends on its impact on different low–level image features, (e.g. topology or contour) that have been modeled in advance and are generally domain dependent (Sec. 3.3).

2. THE STOCHASTIC RECONSTRUCTION ALGORITHM

The term *evolutionary algorithms* (EA) subsumes different computational models of optimization properties of the natural evolution such as genetic algorithms[11] or evolution strategies.[12] The basic idea of such iterative optimization processes is Dawin's well–known principle of "the survival of the fittest".

An individual $\vec{a}(t) \in \mathbb{R}$ is a point in d–dimensional search space, representing one setting for the d free parameters. The vector $\vec{a}(t)$ represents a solution of the optimization problem. The initial population $P(0) = \{\vec{a}_1(0), \cdots, \vec{a}_n(0)\}$ is the starting point of the iterative optimization. The generation $P(n+1)$ is generated from $P(n)$ by application of special manipulation operators, such as mutation (Sec. 3.1) and recombination (Sec. 3.2). A problem dependent target function $f(\vec{a}(t))$ evaluates the quality of the potential solutions and attaches a selection propability (fitness) to them. The better adapted the setting $\vec{a}(t)$ is, the better is its selection propability. This guarantees, that reliable solutions are more likely to generate offsprings than poor ones. After the complete population is reproduced, a selection operator chooses the "fittest" individuals for the next reproduction. This technique applies an optimnisation pressure to the population, that leads to continuously raising fitnesses. The currently best fitness value $f^*(\vec{a}(t))$ converges towards a maximum as $t \rightsquigarrow \infty$.

2.1. The optimization framework

In contrast to common EAs, our approach doesn't start with a randomly generated inital population. Instead, $P(0)$ consists of high–quality segmentation samples from the same image domain as the target image $\mathbf{T}$. Hence, $P(0)$ can be viewed as a *knowledge–base* for the evolutionary learning process.

The k^{th} individual of generation $P(t)$ is denoted by

$$\mathbf{A}_k(t) = (\mathbf{G}_k(t), \mathbf{L}_k(t)) \quad \in \quad \mathbb{N}^{\mathrm{m}\times\mathrm{n}} \times \mathbb{N}^{\mathrm{m}\times\mathrm{n}}. \tag{1}$$

$\mathbf{A}_k(t)$ consists of a graylevel matrix

$$\mathbf{G}_k(t) \quad \text{with} \quad g_k(i,j) \in \{0,1,\ldots,G-1\} \tag{2}$$

and the appropriate label matrix (the segmentation of $\mathbf{G}_k(t)$)

$$\mathbf{L}_k(t) \quad \text{with} \quad l_k(i,j) \in \{0,1,\ldots,L-1\}, \tag{3}$$

with G being the number of graylevels and L being the number of the different possible labels (Fig. 1). A gene of the individual $\mathbf{A}_k(t)$ is denoted by

$$a_k(i,j) = (g_k(i,j), l_k(i,j)) \tag{4}$$

where $i \in \{0,\ldots,(m-1)\}$ and $j \in \{0,\ldots,(n-1)\}$, consists of the graylevel at pixel (i,j) and its attached label.

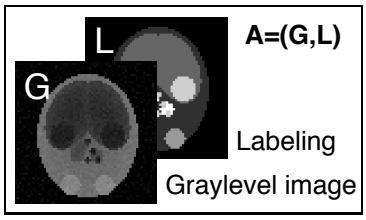

Figure 1. The individual **A** consists of a graylevel image **G** and an appropriate segmentation **L**, that applies an independent class index (label) to each pixel of **G**.

The sample segmentations $\mathbf{A}_k(0)$ should be as well–segmented as possible, because they represent the only source of information for the learning process. Ideally, they are manually segmented by an expert.

The fitness of an individual $\mathbf{A}_k(t)$ is defined as measure of similarity between the graylevel image $\mathbf{G}_k(t)$ and the target image

$$\mathbf{T} \quad \text{with} \quad t(i,j) \in \{0,1,\ldots,G-1\}, \tag{5}$$

i.e. the graylevel image to be segmented. The more similar these two images, the higher is the propability that $\mathbf{L}_k(t)$ is a reasonable segmentation for $\mathbf{T}$. This leads to the central assumption of the algorithm, the linear relationship between the fitness $f(\mathbf{A}_k(t))$ of an individual (defined via $\mathbf{G}_k(t)$) and the segmentation error $q(L_k(t))$:

$$f(A_k(t+1)) \leq f(A_k(t)) \Rightarrow q(L_k(t+1)) \leq q(L_k(t)) \tag{6}$$

with q being the number of false labels, if $L_k(t)$ is taken as segmentation for $\mathbf{T}$. This only holds true if the manipulation operators (recombination and mutation) process information in a reasonable way. The generation of offsprings for $P(t+1)$ out of the reliable information from $P(t)$ by special manipulation operators represents the local optimization process (Sec. 3).

As fitness function f we have chosen the Euclidean distance between two images, since this has two advantages. On the one hand $f(\mathbf{A}_k(t))$ is easy and fast to calculate and on the other hand it provides a means to determine the contribution of each single gene to the total fitness of the individual, which is of importance for the local optimization (Sec. 3.3.2).

$$f(\mathbf{A}_k(t)) = \left(\sum_{i=0}^{m} \sum_{j=0}^{n} (g_k(i,j) - t(i,j))^2 \right)^{1/2} \tag{7}$$

Fig. 2 overviews the algorithm. After each iteration, the individual $\mathbf{A}_k^*(t)$ with the currently best fitness f^* (the smallest distance between $\mathbf{G}_k(t)$ and $\mathbf{T}$ in generation $P(t)$) provides its label matrix $\mathbf{L}_k(t)$ as the current segmentation $\mathbf{S}(t+1)$ for $\mathbf{T}$. Hence, the required labeling of $\mathbf{T}$ is generated implicitly by optimizing the graylevel–based distance function $f(\mathbf{A}_k(t)) = d(\mathbf{G}_k(t), \mathbf{T})$.

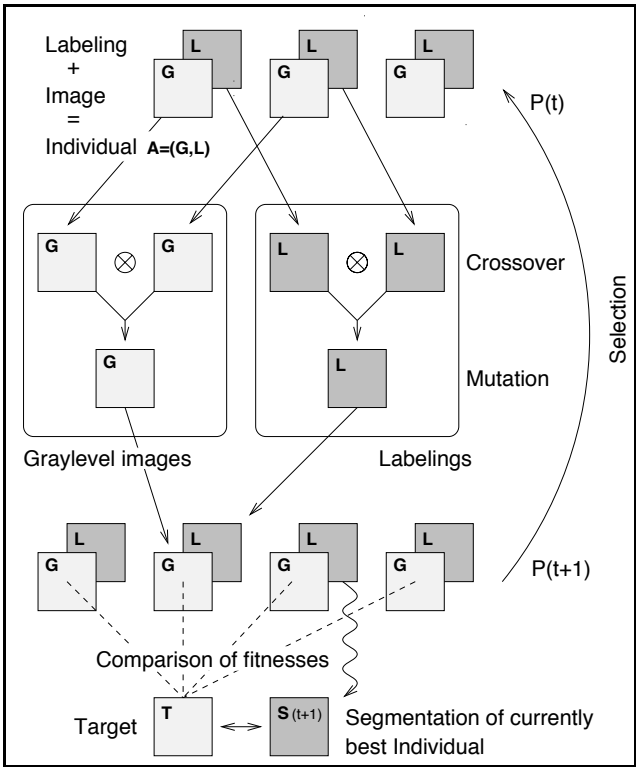

Figure 2. The basic scheme of the stochastic reconstruction algorithm.

2.2. Selection

The selection applies a certain optimisation pressure to the individuals. To be able to evaluate its impact on the optimization results, a generalized selection operator was implemented whose optimization pressure is calculated as the ratio between the number of parents μ and the number of offsprings λ. The smaller the ratio μ/λ is, the higher is the optimization pressure. The selection pressure itself is a matter of optimization, because low pressure is as harmful to the optimization process as high pressure is. As basic selection scheme the $(\mu+\lambda)$–strategy is used,[12] which selects the μ best individuals from the $\mu+\lambda$ individuals of $P(t)$ to be the parents in $P(t+1)$. These selected individuals generate λ offsprings and the selection is performed again on the current $\mu+\lambda$ individuals. Hence, the individuals are potentially immortal, because their lifespan depends exclusively on their fitness.

There is an additional optimization pressure applied by the recombination, because an individual is only capable to propagate its information to the next generation if it is selected for recombination. Two methods of calculating the probability $p_s(\mathbf{A})$ of being chosen for recombination were considered.[22,23] Their influence on the segmentation result is discussed in Sec.4.1.

2.2.1. Proportional selection

The selection probability $p_s(\mathbf{A}_k(t))$ is calculated proportionally to $\mathbf{A}$'s fitness and relative to the mean fitness of the current population:

$$p_s(\mathbf{A}_k(t)) = \frac{f(\mathbf{A}_k(t))}{\sum_{i=1}^{\mu+\lambda} f(\mathbf{A}_i(t))} \tag{8}$$

By this strategy, the fit individuals pass their information to the next generation more often than the less fit ones. This is a desired effect but yields the disadvantage that a super–individual might be created whose offsprings dominate the population and suppress further progress due to a lack of genetic diversity.

2.2.2. Uniform selection

In contrast to the previously described strategy, the uniform selection pressure does not apply an additional selection pressure to the individuals. The probability to generate offsprings is equally distributed within the population:

$$p_s(\mathbf{A}_i(t)) = \frac{1}{\mu+\lambda} \tag{9}$$

Hence, in this variant an individual'a probability to pass its information to the next generation depends only on the number of parents μ and offsprings λ.

2.3. The termination criterion

The algorithm terminates in generation t with $f_t^* = f(\mathbf{S}(t))$ being the best fitness in the current generation $P(t)$, if one of the following four conditions is satisfied:

(1) The user defined maximum number of iteration steps is reached: $t \geq t^{\max}$.

(2) The user defined lower bound of the graylevel fitness is reached: $f_t^* \leq f^{\min}$.

(3) The best fitness in population $P(t)$ is equal to the mean fitness: $f_t^* = \bar{f}_t$. This indicates that the genetic diversity is not sufficient any more for further improvements.

(4) Compared to generation $t-1$ the best fitness f_t^* of the population has not improved, while their mean fitness $\bar{f}_t$ has been decreased. This condition had to be included for extremely high mutation rates p_m (Sec. 3.1). In that case condition (3) might not be met even though the algorithm is unable to generate better solutions.

the maximal calculation time is limited by the user specified parameters $t^{\max}$ and $f^{\min}$. Conditions (3) and (4) showed to be suitable criteria for automatic controlled termination.

3. LOCAL MANIPULATION OPERATORS

The fitness f, implemented as Euclidean distance (Eq. 7) is a unimodal function that does not contain local optima. Hence, the classic mutation,[24] where the whole search space is scanned by arbitrary changes of the genetic information, is inappropriate here. Each gene has a well–defined contribution to fitness f and segmentation error q. Therefore, correlations among genes need not to be considered. In terms of GA, we neglected *co–adapted allels* *. In pre–studies we have seen that mutations of single genes (pixels) are more efficient than mutations of subimages. Consequently, genes will only be copied as a whole in all manipulations. Other techniques that first manipulate the graylevel and then determine an appropriate label on the basis of Bayes–modelled local information turned out to be inadequate in pre–studies.

3.1. Coupled mutation

A mutation $m(\mathbf{A}_k(t)) = \mathbf{A}'_k(t)$ manipulates genes by a certain mutation probability p_m. Fig. 3 exemplifies the principle of mutating single genes. The central gene $a_k(i,j)$ is replaced by a (3×3)–neighbouring gene according to its contribution to the defined low–level image properties E_i, which act like constraints to the mutation process (Sec. 3.3). Those neighbours that influence these properties (and therefore the segmentation quality) in a positive way are more likely chosen to replace the gene $a_k(i,j)$ (Fig, 3). The influence of p_m on the convergence is discussed in Sec. 4.1.

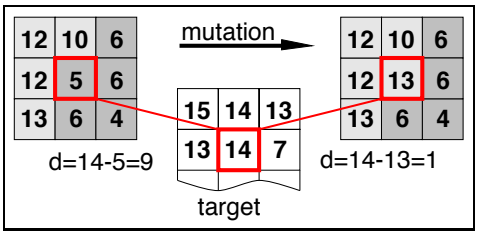

Figure 3. The general mutation operator. If, for example, the only constraint was the improvement of the fitness f, the sketched mutation would most likely occur, for the gene $a(i+1, j-1) = (13, \text{LIGHT})$ fits best to the corresponding target graylevel $t(i,j) = 14$. The mutation reduced the distance between the genes from $d = 9$ to $d = 1$. A side effect of this mutation is the implicit change of $l_k(i,j)$ from DARK to LIGHT.

3.2. Uniform recombination

Pointwise recombination is known as *uniform crossover* (GA)[25] or discrete recombination (ES).[12] These techniques have equal functionality and were used as models for the recombination operator $c(\mathbf{A}_f, \mathbf{A}_m)$. The usual negative consequences of uniform recombination,[24] can be excluded here, because the genes have no co–adapted (correlated) allels as instances. The operator can be described as:

$$c(\mathbf{A}_f(t), \mathbf{A}_m(t)) = \mathbf{A}_k(t) \quad \text{with} \quad 1 \le f, m \le \mu \le k \le \mu + \lambda \tag{10}$$

The two parents $\mathbf{A}_f$ and $\mathbf{A}_m$ are selected according to Sec 2.2. To generate an offspring $\mathbf{A}_k(t)$, a decision has to be made for each gene $a_k(i,j)$, whether $a_k(i,j) = a_f(i,j)$ or $a_k(i,j) = a_m(i,j)$ should be set. Again, this decision is drawn with respect to its effect on the low–level features E_i (Sec. 3.3). At this stage two recombination models can be applied alternatively. Either the gene with the best contribution to the low–level image features E_i is chosen to

*An allel is an instance of a gene. The correct setting of co–adapted allels yields more progress as the same number of "ordinary" uncorrelated allels.

replace $a(i,j)$, or the probabilities of the two candidates $a_m(i,j)$ and $a_f(i,j)$ are set proportional to their contribution to E_i. The two strategies are compared in Sec. 4.1.

This manipulation concept is very fast, because only a few comparisons are required to perform a mutation or crossover. The gene replacements are target–directed and therefore based on local information which decreases the error probability compared to more global or arbitrarily manipulating operators which are usually being used in classical GAs.[24] The recombination can be regarded as a special case of the general mutation operator.

3.3. The constraints

The proposed reconstruction strategy yields sensible results only if the following two central assumptions can be guaranteed:

- ▷ A linear coupling between the "conventional" fitness $f(\mathbf{A})$ and the quality of its labelling $q(\mathbf{L}(t))$ (interpreted as semgentation error) must exist. For the user it must be reliable that if f^* inproves, the quality of $\mathbf{S}(t+1)$ (which is actually unknown to the user) improves accordingly (Eq. 6).
- ▷ The initial graylevel/label relations in the knowledge base must not be destroyed by application of pixel–based manipulation, because this would correspond to the destruction of reliable information. The risk of losing information during recombination or mutation exists, because the aim of the optimization algorithm (to minimize the graylevel distance $d(\mathbf{G},\mathbf{T})$) is different from the user's aim to increase the segmentation quality $q(\mathbf{A})$. If a mutation improves the fitness $f(\mathbf{A})$ of an individual but assigns an unsuitable label to the mutated graylevel, q becomes worse (loss of information) although f improves. Consequently, only sensible graylevel/label pairs (genes) should be allowed to be generated by a manipulation.[20]

This can be achieved by using directed operators that incorporate a manipulations effect on pre–defined image features in their decision. Nontheless, a mutation operator strictly based on the local gradient, such as correlated mutation[12] is not practical here, because an increase of f needs not necessarily result in a decrease of q. E.g. if $f(\mathbf{A})$ improves to fast, the coupling might get lost which results in the divergence of $q(\mathbf{A})$. Conversely, the initial individuals $\mathbf{A}_i(0)$ from the knowledge base yield already good approximations of the desired segmentation $\mathbf{S}^*$. Hence, the operators have to be functionally restricted in order to minimize the risk of destructing the linear coupling between f and q. Hence, the restriction of the seach space yields two advantages: The generation of individuals that are obviously not images of the current domain is prevented, and the algorithm's convergence is speed up.

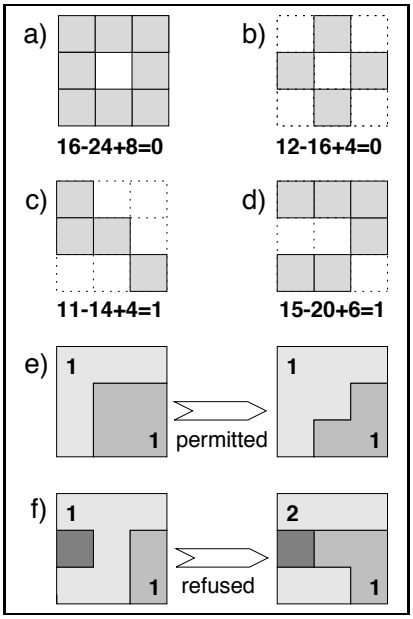

Figure 4. Local topology evaluated by the Euler relation.

3.3.1. Topology invariance

In a real–world application the knowledge base will usually contain images showing objects that are of equal (or very similar) topology. Therefore, a simple technique was introduced in the decision process that improves segmentation quality and convergence speed. Only such gene manipulations are permitted, which preserve the topology of the participating regions. Mathematically, two objects are *topologically equal* if one object can be transformed continuously into the other one, without the emergence of new holes. The *Euler relation* (Eq. 11) provides a means to detect changes of the local topology of discrete objects.[26] For all discrete objects with equal topology we have:

$$C - E - P = constant \tag{11}$$

with C being the number of corners, E the number of edges, and P the number of points of the object. The objects (a) and (b), and the objects (c) and (d) in Fig. 4 are topologically equal as indicated by their Euler numbers. Accordingly, the gene mutation (e) will be permitted and (f) will be refused by the topology check. A label change in a gene is only possible at the edges of the regions, therefore, the occurence of isolated labels (noise) can be prevented. The topology check is a powerful means to restrict the search space.

3.3.2. Directed manipulation by energy models

Replacing a gene $a(i,j)$ by a neighbour (mutation), or by the corresponding gene of the other parent (crossover) respectively, is interpreted as a change of state s_ν. ν is the index of the candidate gene, and s_0 denotes the constant gene replacement, that causes no change of state. To draw the decision which gene replaces the current one, all propabilities p_ν have to be calculated. This is done by evaluating energy terms E_i, that model different low–level image features. A positive effect of a gene replacement on a certain feature results in an energy release $\Delta E_i < 0$. Negative contributions result in an energy input $\Delta E_i > 0$. This corresponds to the physical model, because a solid always tends to reach low energy levels (e.g. mechanical energies like curvature or surface tension). The following four low–level features, modeled by different energy terms E_i, have shown to be suitable as restictions of the search space.

Graylevel energy. The energy term $\Delta E_1(s_\nu)$ models the effect of a gene manipulation on the fitness $f(\mathbf{A})$ of an individual.

$$\Delta E_1(s_\nu) = |g_\nu(i,j) - t(i,j)|^2 - |g(i,j) - t(i,j)|^2 \tag{12}$$

An improvement of an individual's fitness is viewed as desirable change in energy which results in $\Delta E_1(s_\nu) < 0$.

Label energy. $\Delta E_2(s_\nu)$ models a change in an individual's label image $\mathbf{L}$. The occurance of the two labels $l_k^0(i,j)$ and $l_k^\nu(i,j)$ that are subject to manipulation is compared. If $l_k^\nu(i,j)$ is occuring more often in the 5×5 environment of $a_k(i,j)$ than $l_k^0(i,j)$, the manipulation s_ν has a desired contribution to the label feature and therefore causes an energy release ($\Delta E_2 < 0$).

$$\Delta E_2(s_\nu) = \#l_k^\nu(i,j) - \#l_k(i,j) \tag{13}$$

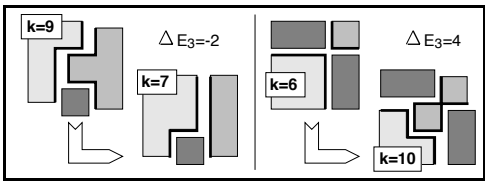

Figure 5. Determination of the surface energy ΔE_3.

Surface energy. $\Delta E_3(s_\nu)$ models the change of surface of the regions that are effcted by a manipulation. If the combined length of surface is increased by the manipulation, this is interpreted as a negative influence on the surface energy ($\Delta E_3 < 0$). Fig. 5 sketches the principle. The left example shows a positive manipulation, because the total number of edges of the two affected regions within the 3×3 environment is reduced from $k = 9$ to $k = 7$.

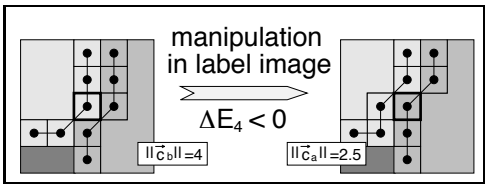

Figure 6. The curvature energy ΔE_4 results from the change of local curvature. The exemplified manipulation yields $\Delta E_4 = -1.5$ and is likely to occur.

Curvature energy. $\Delta E_4(s_\nu)$ models the local curvature of a contour segment. This energy term was motivated by the *active–contours* segmentation method, that uses a similar model for segmentation purposes.[27] The change of curvature at a certain point is determined from the contour segment in its 5×5 environment. Fig. 6 demonstrates this principle. For each point in both paths the resulting curvature vector of the segment is calculated. Their weighted sum yields the curvature vector. The curvature energy ΔE_4 is calculated as the difference of the sum of curvature vectors before ($||\vec{c}_b||$) and after ($||\vec{c}_a||$) the manipulation.

3.3.3. Boltzmann processes

A *Boltzmann process* describes the probability of a change of state in thermodynamic systems as

$$p(s) = \exp\left(\frac{-\Delta E(s)}{kT}\right), \tag{14}$$

where T is the current temperatur in the system. The Boltzmann constant k can be ommited for our purposes. Such processes are used in *Simulated Annealing* techniques to reach a stabil state by reducing the temperature slowly.[21]

A Boltzmann process is used to determine an energy dependent measure $\varepsilon(s_\nu)$ for all potential gene replacements s_ν. The single energy terms ΔE_i can be weighted independently by w_i.

$$\varepsilon(s_\nu) = \exp\left(-\left(\frac{\sum_{i=1}^{n} w_i \Delta E_i(s_\nu)}{T}\right)\right) \tag{15}$$

In our case is $n = 4$, because we have modelled four different energy terms. Further energy terms can be defined if appropriate. A change of state s_ν that changes the local topology will be set to $\varepsilon(s_\nu) = 0$. To draw a decision for a manipulation, the total manipulation probability

$$p(s_\nu) = p_m \cdot \frac{\varepsilon(s_\nu)}{\sum_{i=0}^{\kappa} \varepsilon(s_i)} \tag{16}$$

is normalized to 1. For mutations holds $\kappa = 8$ and for crossover holds $\kappa = 1$. The user–specified temperature T serves as tuning parameter for the optimization process. For temperatures $T \rightsquigarrow 0$ always the best manipulation available is performed, while for $T \rightsquigarrow \infty$ all potential manipulations are equally likely to occur.

4. EXPERIMENTAL RESULTS

4.1. Analysis of algorithmic properties

To analyse the algorithm's capacity, synthetically generated phantom images were used. The knowledge of the optimal segmentation $\mathbf{S}^*$ of the target image $\mathbf{T}$ enables an objective evaluation of results, because the influence of poor contextual information is minimized. The segmentation error $q(\mathbf{L})$ is given as the percentage of mislabeled pixels. Fig. 7 shows the target image $\mathbf{T}$ and its optimal labeling $\mathbf{S}^*$. Note, that the six present regions cannot be determined from the target's bimodal histogram $h(\mathbf{T})$. The segmentation errors $q(\mathbf{L}_i(0))$ indicate the segmentation qualities of the knowledge–base $P(0) \ni \mathbf{A}_i(0)$.

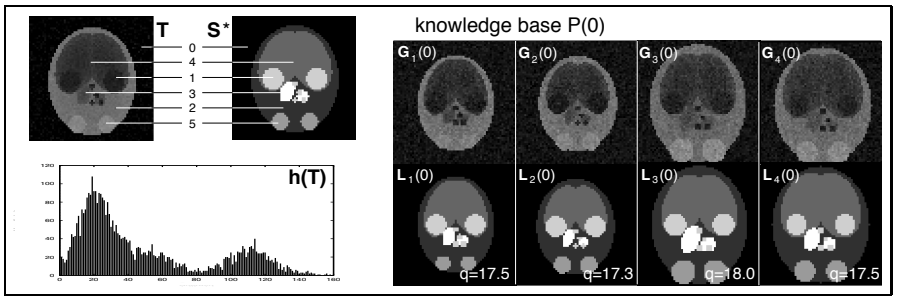

Figure 7. The algorithm's convergence was analysed with the target image $\mathbf{T}$, and the knowledge base $P(0)$. The six regions cannot be determined from the histogram $h(\mathbf{T})$. The errors $q(\mathbf{L}_i(0))$ give the percentage of mislabeled pixels compared with the optimal solution $\mathbf{S}^*$.

The segmentation results shown in Fig. 8 exemplify two aspects of the algorithm, the effect of the local topology check (Sec. 3.3.1) and the crossover operator (Sec. 3.2). All examples were generated with optimized settings for temperature T and energy weights w_i, that were determined in pre–studies. Even though the segmentation error was equal in Fig. 8a and 8b, the effect of the topology check is obvious. The consideration of topology leads to better subjective results, because the occurance of isloated labels is prevented. Topology invariance is a poweful tool to restrict the search space and inprove results.

The experiments in Fig. 8c and 8d prove the efficiency of recombination. The difference images $\mathbf{S}(t) - \mathbf{S}^*$ show the mislabeled parts. If the described crossover is used, the segmentation error is three times less, while the convergence speed is doubled. The crossover provides the transfer of information within the population and guarantees genetic divercity. The fitness values $f(\mathbf{A})$ converged towards the limits $f = 524$ (Fig. 8c) and $f = 178$ (Fig. 8d). The ratios $f(\mathbf{A}(t))/q(\mathbf{S}(t))$ are similar in both cases, which is a first evidence that the postulated linear coupling between fitness $f(\mathbf{A})$ and quality $q(\mathbf{L})$ of an individual is preserved by the algorithm (Eq. 6).

The mutation rate p_m represents the propability that a gene is mutated. For mutation rates $18 \leq p_m \leq 20$ the results reach a minimal segmentation error as shown in Fig. 9. The fact that the segmentation error is still acceptable even if the mutation rate is $p_m \geq 70$ is noticable. This behaviour results from the restrictions of the search space,

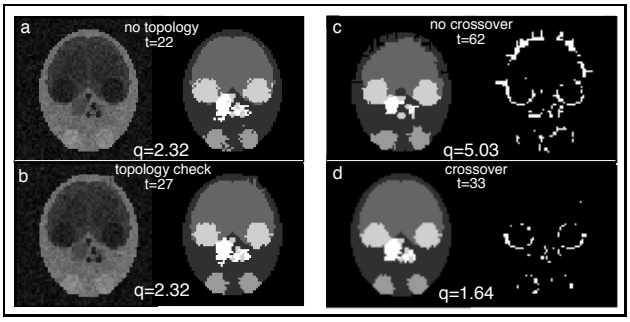

Figure 8. The experiments (a) and (b) show the effect of the local topology check. The left image is the reconstructed greylevel image $\mathbf{G}_k(t)$, while the according label image $\mathbf{S}^*(t) = \mathbf{L}_k(t)$ is displayed on the right. In (c) and (d) the influence of crossover is shown. The generated segmentations $\mathbf{S}(t)$ (left) contain mislabeled pixels, that are visualized by the difference images $\mathbf{S}(t) - \mathbf{S}^*$ (right).

because they guide the individuals in desired directions in the search space, even for extremely high mutation rates. Ordinary EAs, which do not restrict the search space by image specific constraints, yield poor results for such high mutation rates, because information is destroyed rather than newly generated.[28] As expected, the convergence time decreases exponentially as the mutation rate rises.

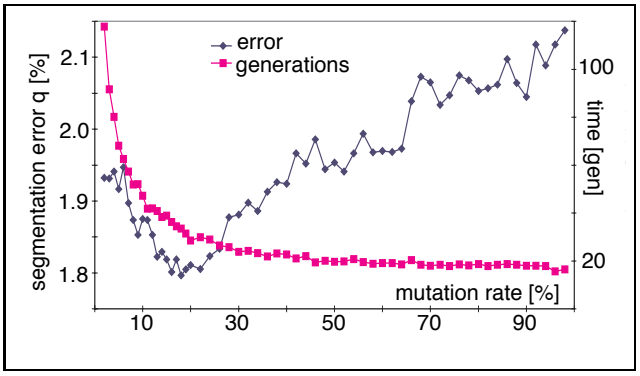

Figure 9. Segmentation error q and convergence time t are dependent on the mutation rate p_m.

Two other important parameters for convergence and segmentation error $q(\mathbf{L})$ are the selection variant and the selection pressure applied to the individuals. The left diagram in Fig. 10 shows the segmentation error $q(\mathbf{L})$ for proportional (Sec. 2.2.1) and uniform selection (Sec. 2.2.2), dependent on the selection pressure $\mu/24$. The number of offsprings was set to $\lambda = 24$, because pre–studies showed, that larger populations do not lead to significantly better results. The use of potentially immortal individuals seems adequate, because the search space does not contain prominent local optima that might catch the individuals. Although the design of the gene manipulation surely introduces further local optima into the initially homogeneous search space, this assumption still holds true.

As expected, the proportional selection scheme yields better results than the uniform one. The optimal selection pressure setteled in both cases between $\mu/\lambda \approx 1/2$ and $\mu/\lambda \approx 2/3$. The right diagram of Fig. 10 presents the results of the two examined recombination variants (Sec. 3.2) dependent on the mutation rate p_m. The best–choice graph has a distinct optimum at $p_m = 10$, while the proportional–choice recombination yields segmentation errors, that seem to be independent on the mutation rate p_m.

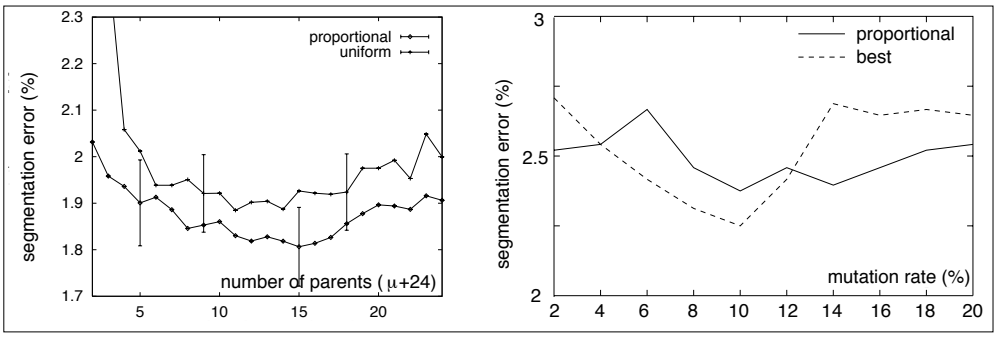

Figure 10. Left digram: The selection pressure $\mu/24$ is evaluated for both, proportional and uniform selection. Right diagram: The recombination types (proportional, best) are evaluated dependent on the mutation rate p_m.

4.1.1. Convergence of the algorithm

Fig. 11 shows a sample convergence of the scaled fitness $f^*(\mathbf{A}_i(t))/100$ (Eq. 7) of the best individuals and the corresponding segmentation errors $q(\mathbf{L}_i(t)) = q(\mathbf{S}(t))$. The larger diagram (with both axes logarithmically scaled) illustrates the linear graylevel/label coupling (Eq. 6). The postulated linear coupling does not hold until final convergence, for the segmentation error q decreases monotonously only until step $t = 8$. The linear decrease of both graphs in the smaller diagram (logarithmic y–axis) signals exponential convergence (for q until $t = 8$, and for f until $t = 13$). The algorithm terminates in generation $t = 33$ with a segmentation error $q(\mathbf{S}(33)) = 1.64$ that is not significantly better than $q(\mathbf{S}(8))$. Hence, the conclusion from this experimant is that the linear coupling obviously holds as long as both graphs, the fitness f and the segmentation error q, converge exponentially.

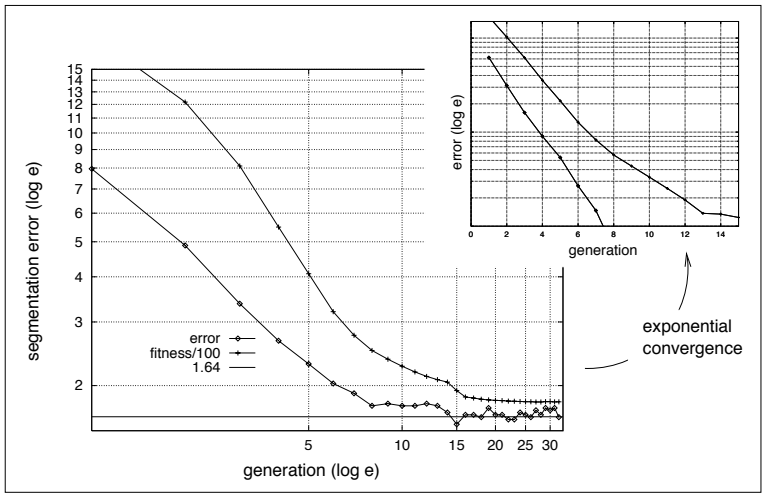

Figure 11. Convergence characteristics of the EA. Both diagrams show the scaled fitness $f/100$ (upper graph) and the segmentation error q (lower graph) along the t–axis. The bigger diagram (double logarithmic) illustrates the linear coupling of the graphs until $t = 8$. Both graphs converge exponentially, which can be said from the qualitative linear course (after subtracting the limit) in the smaller y–logarithmic diagram.

4.2. Applications to physiological image data

After the algorithm's segmentation capacity was demonstrated in the previous sections, it is applied to two different physiological domains. In the experiment shown in Fig. 12 a set of 19 transversal NMR scans of a human brain was segmented into seven anatomical sections by an expert. The 10^{th} slice was used as target image $\mathbf{T} = \mathbf{G}_{10}(0)$ for the segmentation process. Hence, the initial population consisted of 18 individuals, those four that are most similar to the target are shown in Fig. 12. The knowledge base in the experiment in Fig. 13 contained five similar sagittal head slices (from data sets of 128 images) of two different persons. The segmentations into four regions were obtained automatically, but had to be enhanced manually afterwards, to suit the algorithms capability. This was necessary, because exact topolgical information is most important for our algorithm and is provided by very few automatic segmentation techniques. As target image a similar slice was taken from one of the data sets.

When interpreting the results from Fig. 12 and Fig. 13, one has to keep in mind that both knowledge bases were poor in two ways. Firstly the segmentations are based on anatomical structures and do not necessarily correspond to certain graylevel distributions that could be extracted automatically by an algorithm, because of the incorporation of subjective information from the human observers who segmented the images. Secondly the number of individuals that carried usefull information were small in both cases, because the majority of slices were to different from the target images. In both experiments, comprehensive knowledge bases could not be used, because this would mean that similar data sets from several patients are provided, which wasn't the case. Under these conditions, the presented results are very interesting, because they demonstrate two noticable properties of the algorithm.

- ▷ The available information was recombined where appropriate, and conserved where no better information could be found. There is no indication, that information was destructed. This behaviour results from the constraints on the local manipulation operators (Sec.3.3), which keep the individuals in reasonable areas of the search space. Furthermore, unnecessary information was reduced as far as possible (limited only by topology invariance) as can be seen especially in Fig. 13.

▷ The optimal weights w_i, used for Fig. 12a and Fig. 13a were obtained heuristically. The boundaries of their regions are smooth because surface energy ΔE_3 and curvature energy ΔE_4 have a strong influence. Even though Fig. 12b and Fig. 13b were obtained with dominating graylevel energy ΔE_1, yielding irregular and unnatural region contours, the according results are topologically very similar. (The weights of Fig. 12c and Fig. 13c represent a compromise between optimal weights and dominating graylevels.) Obviously, the weights w_i control the general shape and look of the regions but not necessarily their location and extend. This is given by the samples of the knowledge base.

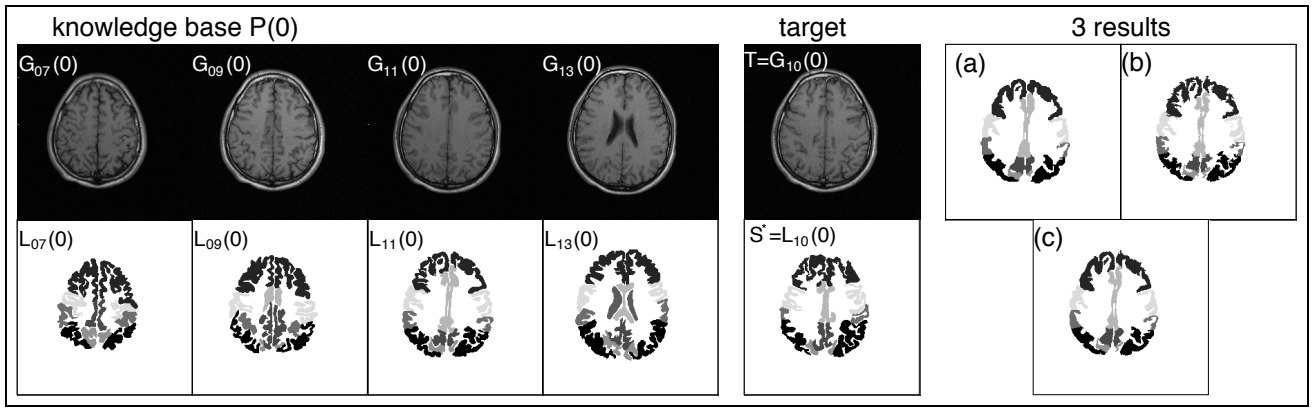

Figure 12. Four out of 18 manually segmented NMR scans of a human brain are displayed to represent the knowledge base $P(0)$. The middle slice $\mathbf{G}_{10}(0)$ was chosen as target $\mathbf{T}$, it contains 7 different anatomical regions . The two results were yielded with different parameter settings. One with the (heuristically determined) optimal energy weights w_i and the second with dominating graylevel energy.

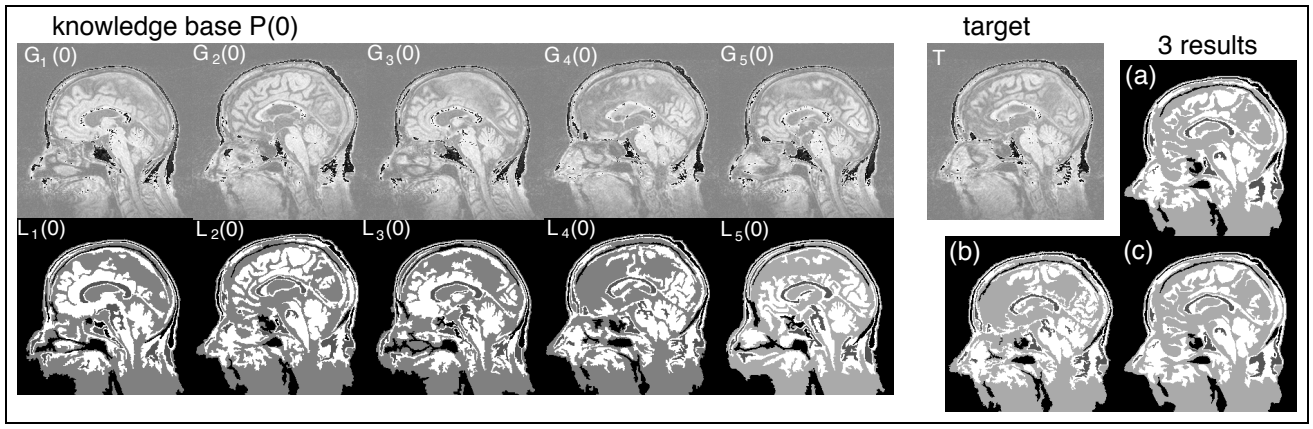

Figure 13. Five sagittal NMR head slices from two different persons made up the initial population $P(0)$. The automatically obtained segmentations were topologically enhanced by an expert to meet the algorithms requirements. The target slice $\mathbf{T}$ was from a similar plane as the initial individuals. For the generation of the results different sets of energy weights w_i were used.

5. CONCLUSIONS

The proposed method turned out to be a powerful technique for knowledge–based image segmentation as could be proved by experimental results from both, synthetic and physiological data sets. Compared to other algorithms its main advantages is its applicability in a wide range of fields as soon as an apropriate knowledge–base is provided. However, for routine use (e.g. in medical image processing), further development is required, especially by enhancing

the termination criterion a general speed up could be achieved. Also, more sophisticated fitness functions, that weight the influence of the genes independently might lead to better results. However, the next developing stages need to be analysed by an application of the algorithm to a physiological domain, based on comprehensive knowledge.

In order to meet the requirements of particular image domains in detail, a library of energy models for different low–level image features, might be a useful contribution. The individual weights w_i can be determined in advance by an analysis of the provided knowledge–base. Since the topology invariance allows changes of labels just at the edges of regions, an edge–based approach might reduce memory requirements without touching the philosophy of the algorithm. Since EAs have a very clear hierarchical structure, many sophisticated parallel strategies for enhancing calculation were suggested[29] which could further exhaust the problem solving potential of the stochastic reconstruction algorithm.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Elena Rota–Kops (Research Center Jülich/Germany) for the kind providence of the NMR sequences, used in Sec. 4.2, Fig. 13.

REFERENCES

1. N. Pal and S. Pal, "A review on image segmentation methods," *IEEE Trans. Pattern Recognition* **26**, pp. 1277–1294, 1991.
2. R. Haralick and L. Shapiro, "Image segmentation techniques," *Computer Vision, Graphic and Image Processing* **29**, pp. 100–132, 1985.
3. W. Wells III, W. Grimson, R. Kikinis, and F. Jolesz, "Adaptive segmentation of MRI data," *IEEE Trans. Medical Imging* **15**(4), pp. 429–442, 1996.
4. M. Fukumi, S. Omatu, F. Takeda, and T. Kosaka, "Rotation–invariant neural pattern recognition system with application to coin recognition," *IEEE Trans. Neural Networks* **3**(2), pp. 272–277, 1992.
5. L. Chunlin, D. Goldof, and L. Hall, "Knowledge–based classification and tissue labelling of MR images of human brain," *IEEE Trans. Medical Imaging* **12**, pp. 740–750, 1993.
6. D. Paulus, A. Winzen, and H. Niemann, "Knowledge based object recognition and model generation," in *Procs. SPIE*, vol. 1989, pp. 382–393, 1993.
7. S. Raya, "Low–level segmentation of 3–D magnetic resonance brain images: A rule–based system," *IEEE Trans. Medical Imaging* **9**, pp. 327–337, 1990.
8. M. Ozkan, B. Dawant, and R. Maciunas, "Neural–network–based segmentation of multi–modal medical images: A comparative and prospective studie," *IEEE Trans. Medical Imaging* **12**, pp. 534–544, 1993.
9. S. Sandor and R. Leahy, "Surface–based labeling of cortical anatomy using a deformable atlas," *IEEE Trans. Image Processing* **6**(1), pp. 41–54, 1997.
10. H. Tagare, "Deformable 2–D template matching using orthogonal curves," *IEEE Trans. Medical Imaging* **16**(1), pp. 108–117, 1997.
11. D. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, Reading, MA, 1989.
12. H.-P. Schwefel, *Evolution and Optimum Seeking*, John Wiley & Sons, New York, 1995.
13. L. Davis, ed., *Genetic Algorithms and Simulated Annealing*, Morgan Kaufmann, San Mateo, CA, 1988.
14. B. Bhanu, S. Lee, and J. Ming, "Self–optimizing image segmentation using a genetic algorithm," in *Proc. of* 4^{th} *Int. Conf. on Genetic Algorithms*, pp. 362–369, 1991.
15. B. Bhanu, S. Lee, and J. Ming, "Adaptive image segmentation using a genetic algorithm," *IEEE Trans. Systems, Man, and Cybernetics* **25**(12), pp. 1543–1567, 1995.
16. D. Chun and H. Yang, "Robust image segmentation using genetic algorithms with a fuzzy measure," *Pattern Recognition* **29**(7), pp. 1195–1211, 1996.
17. Y. Huang, K. Palanippan, X. Yhuang, and J. Cavanaugh, "Optic flow field segmentation and motion estimation using a robust genetic paritioning algorithm," *IEEE Trans. Pattern Analysis and Machine Intelligence* **17**(12), pp. 1177–1190, 1995.
18. M. Sonka, S. Tadikonda, and S. Collins, "Knowledge–based interpretation of MR brain images," *IEEE Trans. Medical Imaging* **15**(4), pp. 443–452, 1996.
19. L. Booker, D. Goldberg, and J. Holland, "Classifier systems and genetic algorithms," *Artificiall Intelligence* **40**, pp. 235–282, 1989.
20. P. Andrey and P. Tarroux, "Unsupervised image segmentation using a distributed genetic algorithm," *Pattern Recognition* **27**(5), pp. 659–673, 1994.
21. E. Aarts and J. Korst, *Simulated annealing and Boltzmann machines*, Wiley, Chichester, 1989.
22. T. Bäck and F. Hoffmeister, "Extended selection mechanisms in genetic algorithms," in *Proc. of the* 4^{th} *Int. Conf. on Genetic Algorithms*, R. Belew and L. Booker, eds., pp. 92–99, Morgan Kaufmann, San Mateo CA, 1991.
23. D. Goldberg and K. Deb, "A comperative analysis of selection schemes used in genetic algorithms," in *Foundations of Genetic Algorithms*, G. Rawlins, ed., pp. 69–93, Morgan Kaufmann, San Mateo, CA, 1990.
24. J. Holland, *Adaptation in Natural and Artificial Systems*, University of Michigan Press, Ann Arbor, 1975.
25. G. Syswerda, "Uniform crossover in genetic algorithms," in *Proc. Third Int. Conf. on Genetic Algorithms*, J. Shaffer, ed., pp. 2–9, Morgan Kaufmann, San Mateo, CA, 1989.
26. J. Serra, *Image Analysis and Mathematical Morphology*, ch. 4, Digital Morphology. Academic Press, London, 1982.
27. S. Lobregt and A. Viergever, "A discrete dynamic contour model," *IEEE Trans. Medical Imaging* **14**(1), pp. 12–24, 1995.
28. J. Hesser and R. Männer, "Towards an optimal mutation propability in genetic algorithms," in *Parallel Problem Solving from Nature*, H.-P. Schwefel and R. Männer, eds., Lecture Notes in Computer Science 496, Springer, Berlin, 1991.
29. H. Mühlenbein, "Parallel genetic algorithms, population genetics and combinatorial optimization," in *Proc. Third Int. Conf. on Genetic Algorithms*, J. Shaffer, ed., pp. 416–412, Morgan Kaufmann, San Mateo, CA, 1989.