Centrum voor Wiskunde en Informatica

REPORT RAPPORT

Image segmentation by connectivity preserving relinking in hierarchical graph structures

P.F.M. Nacken

Department of Operations Research, Statistics, and System Theory

Report BS-R9417 May 1994

CWI is the National Research Institute for Mathematics and Computer Science. CWI is part of the Stichting Mathematisch Centrum (SMC), the Dutch foundation for promotion of mathematics and computer science and their applications.
SMC is sponsored by the Netherlands Organization for Scientific Research (NWO). CWI is a member of ERCIM, the European Research Consortium for Informatics and Mathematics.

P.O. Box 94079, 1090 GB Amsterdam (NL)
Kruislaan 413, 1098 SJ Amsterdam (NL)
Telephone +31 20 592 9333
Telefax +31 20 592 4199

Image Segmentation by Connectivity Preserving Relinking in Hierarchical Graph Structures

Peter F.M. Nacken

TNO Institute for Human Factors
Kampweg 5, 3769 DE Soesterberg, The Netherlands

Abstract

The method of image segmentation by pyramid relinking is extended to the formalism of hierarchies of region adjacency graphs. This approach has a number of advantages: (1) resulting regions are connected, (2) the method is adaptive, and therefore artifacts caused by a regular grid are avoided, and (3) information on regions and boundaries between regions can be combined to guide the segmentation procedure. The method is evaluated by the segmentation of a number of synthetic and natural images.

AMS Subject Classification (1991): 68U10
Keywords & Phrases: image segmentation, hierarchies of graphs, relinking, connectivity
Note: This research was supported by the Foundation for Computer Science in the Netherlands (SION) with financial support from the Netherlands Organisation for Scientific Research (NWO). This research was part of a project in which the TNO Institute for Human Factors, CWI and the University of Amsterdam cooperate. Most of this work was performed while the author was a guest at the Technical University of Vienna, Austria, in the framework of the Erasmus exchange program of the European Community.

1. Introduction

In this paper, the extension of pyramid relinking to hierarchies of graphs is developed. This approach allows the segmentation of an image into connected regions and the use of boundary information in the relinking process.

Pyramid relinking, which was originally described by Burt *et al.* [5], is a powerful and conceptually attractive method for image segmentation. In the past, a number of techniques have been proposed to reduce some of its weaknesses (e.g. [1, 2]). A number of extensions have been made, which allow the application of relinking to other image types, such as flow fields [8] or textured images [16].

Conventional image segmentation by relinking uses a regular pyramid structure, as illustrated in figure 1.1. This is a stack of regular grids of sizes $2^n \times 2^n$, $2^{n-1} \times 2^{n-1}$, $\ldots$, 1×1. In the lowest level of the pyramid, each cell corresponds to a pixel in the image grid. Each cell in level $i+1$ represents a cluster of cells in level i. The cells which may be contained in such

Report BS-R9417
ISSN 0924-0659
CWI
P.O. Box 94079, 1090 GB Amsterdam, The Netherlands

a cluster form a 4×4 square in level i. These squares overlap in such a way that each cell on level i can belong to one out of four clusters. The cells in the cluster represented by a given cell are called the *children* of this cell; the representing cell is called the *parent* of its children. Each cell has one parent. Note that, for cells near the image border, the number of possible parents and children of a cell is smaller than 4 and 16, respectively.

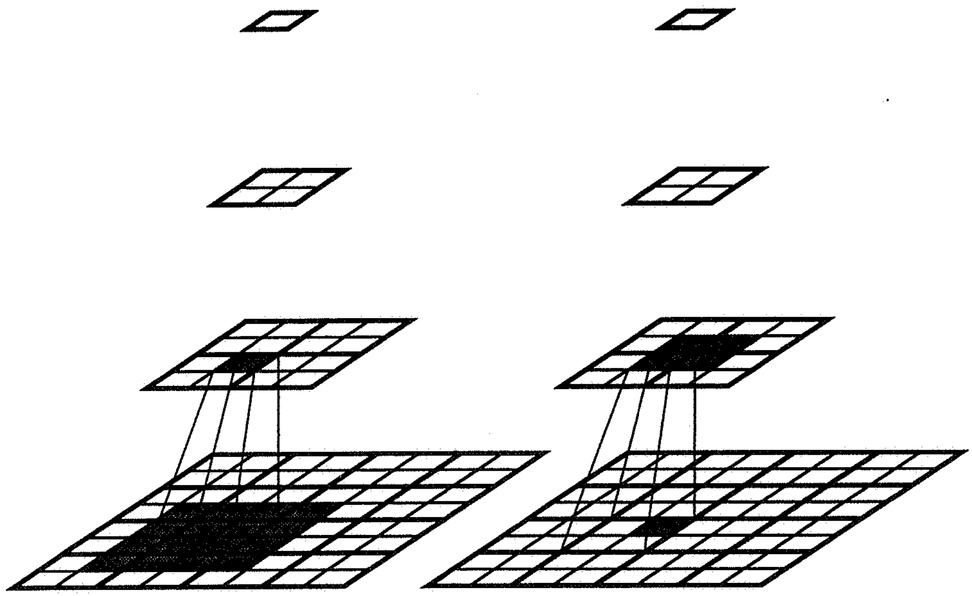


Figure 1.1: The regular pyramid structure used in conventional relinking. The lines between levels show the relative positions of cells in different levels. On the left, the sixteen shaded cells in the lower level are the candidate children for the shaded cell in the higher level. On the right, the four shaded cells in the higher level are the candidate parents for the shaded cell in the lower level.

The parent-child links induce a tree structure in the hierarchy. Thus, each cell in the pyramid represents a region in the image plane, which can be found by tracking all series of parent-child links leaving from a given cell, in a downward direction.

Segmentation by relinking is performed by iteratively updating the cluster membership of cells, i.e. by adapting parent-child links. This is done in such a way that the standard deviation of the grey levels of the regions represented in each cell decreases. Thus, the relinking procedure converges towards a state in which the region represented by each vertex is as homogeneous as possible.

In this paper, some drawbacks of the relinking method are discussed and a relinking scheme based on hierarchies of graphs is presented, which solves these problems. The drawbacks originate from the fact that the levels in a relinking pyramid do not represent a region adjacency graph, but merely a subdivision of the image points in a predefined number of classes. Moreover, not all subdivisions corresponding to a subdivision of the image in the correct number of classes, can be represented [4].

The first problem is the fact that the clusters represented by a cell need not correspond to connected regions in the image plane. The relinking process takes the spatial structure of the image into consideration by allowing only a fixed set of possible children for each cell. The algorithm does not use connectivity: cells which are adjacent in some higher level grid of the pyramid need not represent adjacent regions in the image plane, or vice versa. This can cause the creation of regions which are scattered over the image plane and which consist of many connected components. If the pyramid structure is adapted by increasing the number of candidate children for each cell, regions can become increasingly scattered and the process becomes similar to isodata clustering of grey values, in which no account is taken of spatial structure. (See Kasif and Rosenfeld [11] for a discussion of the relations between pyramid relinking and isodata clustering.)

The second problem is caused by the regularity of the grid of cells in each level of the image grid and the associated set of 16 possible children for each cell. In such a configuration, not all possible subdivisions of the image plane can be represented, as was shown by Bister *et al.* [4]. Therefore, artifacts can occur in the segmentation of particular shapes such as elongated ones. This can be repaired by allowing irregular structures, in which the number of levels and the number of neighbors for each vertex is not fixed in advance.

The third problem is related to the first one. As the concept of a connected region can not be represented in the conventional relinking pyramid, it is also not possible to manipulate or represent the boundary between two classes or regions. Therefore, it is not possible to use information on boundaries between regions, such as length and average response of an edge detection filter, in a relinking based segmentation method.

In this paper, we attack these three difficulties by using the hierarchy of graphs formalism. In section 2, we describe new relinking rules that force the regions represented by each cell to be connected.

However, this relinking strategy fails to detect strongly elongated objects, such as spirals, as a single region. Moreover, the number of regions represented by the hierarchy is fixed. In section 3, a method for the adaptive construction of subsequent levels in a hierarchy of graphs is presented, in which arbitrary image subdivisions can be represented.

The methods discussed in sections 2 and 3 require the application of the hierarchy of graphs formalism. In this formalism, information on the boundaries between regions can be represented as the attributes of the edges of the region adjacency graphs. In section 4, it is described how this possibility can be exploited for the combination of region and boundary information in order to improve the segmentation process.

In section 5, the integration in a single system of the methods described in the previous section is described and some segmentation results are shown.

Section 6 presents the conclusions of this paper.

2. Connectivity Preserving Relinking

In the classical relinking method [5], the spatial arrangement of cells in a regular grid in higher levels of the pyramid does not reflect the spatial arrangement or connectivity of the regions represented by such cells. In this section, an adaptation of the relinking rules is proposed. If these new rules are used, regions are guaranteed to be connected and the adjacency relations between the regions can be represented by a region adjacency graph. The proper framework in this section is the *hierarchy of graphs.* From now on, graphs will be denoted as $G = (V, E)$, where V is the vertex set of the graph and E is the edge set.

Definition 2.1 *A hierarchy of graphs is a sequence* $G_i = (V_i, E_i)$ *(with* $i = 0, \ldots, n$*) of graphs and a sequence* $(\pi_0, \ldots, \pi_{n-1})$ *of mappings* $\pi_i : V_i \to V_{i+1}$ *such that:*
(1) for $i = 0, \ldots, n-1$, $\pi_i(V_i) = V_{i+1}$ *;*
(2) for each $i = 0, \ldots, n-1$ *and each* $x \in V_{i+1}$ *,* $\pi_i^{-1}(x)$ *is a connected subset of* G_i *;*
(3) for each $i = 0, \ldots, n-1$ *and* $x, y \in V_{i+1}$ *,* $(x, y) \in E_{i+1}$ *if and only if there are* $x' \in \pi_i^{-1}(x)$ *and* $y' \in \pi_i^{-1}(y)$ *such that* $(x', y') \in E_i$ *.*

A hierarchy of graphs is illustared in figure 2.1. The mapping $\pi_i : V_i \to V_{i+1}$ assigns to each vertex in V_i a *parent* in V_{i+1}. The mapping $\kappa_i : V_i \to \mathcal{P}(V_{i-1})$ assigns to each vertex $v \in V_i$ its children $\{w \in V_{i-1} \mid \pi_{i-1}(w) = v\}$. Where no confusion can occur, the subscripts of π and κ are omitted.

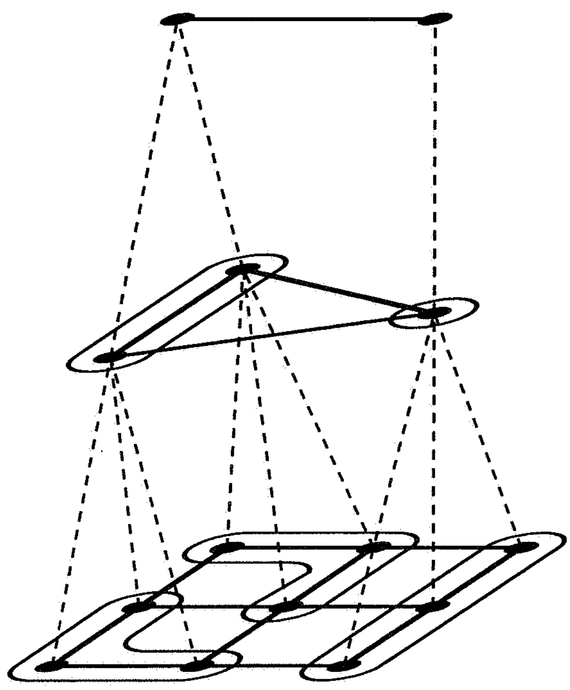


Figure 2.1: An example of a hierarchy of graphs, consisting of three levels. Edges in each level are indicated by solid lines and parent-child relations by dashed lines. The groups of children of each parent are encircled; these groups are represented by their parent on the next level.

We will use notations such as $\kappa(X)$ for $\{\kappa(v) \mid v \in X\}$, where X is a set of vertices. Sometimes $\kappa(v)$ will be identified with the subgraph consisting of the children of v. We will write $N(v)$ for the set of neighbors of v in the graph. If $v \in V_i$, the set $\kappa_1\kappa_2 \ldots \kappa_i(v)$ is the set of vertices in the base level graph which corresponds to the region in the image represented by vertex v. This region is called the *receptive field* of v and is denoted by $R(v)$. If $v \in V_i$, then $\pi^k(v)$ will denote $\pi^{i+k-1} \ldots \pi^{i+1}\pi^i(v)$ and $\kappa^k(v)$ will denote $\kappa^{i-k+1} \ldots \kappa^{i-1}\kappa^i(v)$.

The lowest level graph $G_0 = (V_0, E_0)$ corresponds to the image grid, which is considered as a 4-connected graph. It is also possible to use a 6-connected or 8-connected grid. The latter has the disadvantage that it is not a planar graph. This is not a problem in the present paper, but it would be a problem for other methods, which will be described elsewhere [15].

The edges in the higher levels represent the adjacency structure for the receptive fields of the vertices in that level. These edges depend on the parent-child relations and on the structure of the base level graph. Two vertices v and w on level i are connected by an edge if they have children $v' \in \kappa(v)$ and $w' \in \kappa(w)$ which are connected in level $i-1$. Recursive application of this criterion allows for the construction of all levels of the graph from the base level.

During the discussion of the connectivity of regions, the following result will be used.

Theorem 2.2 *Consider a hierarchy of graphs $G_i = (V_i, E_i)$ with parent assignments π_i and child assignments κ_i. Suppose that the edges of G_0 are fixed and that the edges in higher levels* *are induced by the base level and the parent assignments. Then the following statements are equivalent.*

1. *For each i and for each vertex $v \in V_i$, the subgraph of G_{i-1} induced by the vertices in $\kappa_i(v)$ is connected.*
2. *Each vertex v has a connected receptive field.*

PROOF. Suppose that the first statement holds. It will be shown with induction in the level i that each vertex $v \in V_i$ has a connected receptive field in G_0. For $i = 0$, this is obviously true: each vertex in the base level is its own receptive field. Now suppose that the first statement holds for level $i-1$ and let $v \in V_i$. Consider two vertices w_1, w_2 in the receptive field $R(v)$ of v.

As $\kappa_i(v)$ is connected, there is a path from $\pi^{i-1}(w_1)$ to $\pi^{i-1}(w_2)$ in $\kappa(v) \subset V_{i-1}$. Denote this path by $u_1, \ldots, u_n$. For each edge (u_j, u_{j+1}) in this path there is an edge (u'_j, u''_{j+1}) in the base level with $u'_j \in R(u_j)$ and $u''_j \in R(u_j)$. By the induction hypothesis, each $R(u_j)$ is connected. As u'_j and u''_j are both in $R(u_j)$, there are paths in the base level from u''_j to u'_j. For the same reason, there are paths from w_1 to u'_1 and from u''_n to w_2 in the base level. Concatenation of these paths yields a path from w_1 to w_2 in the base level, which is entirely contained in $R(v)$. Thus, there is a path from w_1 to w_2 in the receptive field $R(v)$ and $R(v)$ is connected.

Now suppose that the second statement holds. Consider a vertex $v \in V_i$. Let w_1 and w_2 be two vertices in $\kappa_i(v) \subset V_{i-1}$. It will be shown that there is a path in $\kappa(v)$ from w_1 to w_2. Let $w'_1 \in \kappa^{i-1}(w_1)$ and $w'_2 \in \kappa^{i-1}(w_2)$ be two vertices in the base level. Then there is a path $w'_1 = u_1, u_2, \ldots, u_n = w'_2$ in $R(v)$ from w'_1 to w'_2. Consider the vertices $u'_j = \pi^{i-2} \ldots \pi^1\pi^0(u_j)$ in V_{i-1}. For each pair (u_j, u_{j+1}), either $u'_j = u'_{j+1}$ or (u'_j, u'_{j+1}) is an edge in $\kappa(v)$. Thus, by deleting repetitions, the sequence $u'_1, \ldots, u'_n$ yields a path in $\kappa(v)$ from $w_1 = u'_1$ to $w_2 = u'_n$. ∎

The second property in theorem 2.2 (connectivity of receptive fields) is the one we are interested in; yet the first one (connectivity of the set of children of a given vertex) is the most manageable one, because it is a local property. Therefore, in the sequel, only connectivity of sets of children will be discussed.

Consider a hierarchy of graphs G_i in which each vertex represents a connected region in the base level. Suppose $\pi(v) = p_{\text{old}}$ for some $v \in V_k$, for some k. Consider the adapted stack of graphs G'_i which is constructed from G_i by putting $\pi(v) = p_{\text{new}}$ and adapting the edge structure accordingly. In a relinking process, the hierarchy of graphs is adapted repeatedly in this way, until an optimal segmentation is achieved. All vertices can be relinked to a new parent simultaneously or they can be relinked one at a time.

We now discuss, for which vertices the connectivity of the receptive field is lost by this relinking step. Edges change only in levels k and higher and the receptive fields of vertices in the levels $i \leq k$ do not change. Vertices in level $k+1$ and vertices in levels $i > k+1$ will be discussed separately.

2.1. Connectivity Preservation in the Parent Level

Suppose a vertex v in level k with parent p_{old} is relinked to a new parent p_{new}. In level $k+1$, the receptive fields of p_{old} and p_{new} change. After the relinking step, the children of p_{old} are $\kappa(p_{\text{old}}) \setminus \{v\}$; the children of p_{new} are $\kappa(p_{\text{new}}) \cup \{v\}$. The receptive fields of other vertices in this level do not change.

Connectivity of receptive fields in level $k+1$ is lost in two cases (see figure 2.2). The first case (shown to the left) is the situation where $\kappa(p_{\text{new}}) \cup \{v\}$ is not connected. This happens when v is not connected to some vertex in $\kappa(p_{\text{new}})$, i.e. when p_{new} is not the parent of some neighbor of v.

Figure 2.2: The two ways in which connectivity can be lost at the parent level by relinking. The receptive fields that become disconnected by a relinking step are marked.

The second case (shown to the right) is the situation where removing v from $\kappa(p_{\mathrm{old}})$ changes the number of connected components of $\kappa(p_{\mathrm{old}})$. Such a vertex v is called an *articulation point* or *cut point* [7] of $\kappa(p_{\mathrm{old}})$. Note that this also includes the case where v is the only child of p_{old}, in which case p_{old} would have an empty receptive field after the relinking step.

The time required for the computation of the articulation points of the subgraph $\kappa(p_{\mathrm{old}})$ is $\mathcal{O}(|\kappa(p_{\mathrm{old}})|)$ [17]. They need only be recomputed when $\kappa(v)$ is changed by a relinking step.

Summarizing, retaining the connectivity of the receptive fields of vertices in level $k+1$ imposes two conditions on the relinking rules:

1. A vertex v which is an articulation point of $\kappa(\pi(v))$ may not be relinked to a new parent.
2. A vertex v may choose only a new parent from the set $\pi(N(v))$, i.e. a parent of a neighbor of v.

Wharton [20] has tried to find a connectivity preserving relinking method, but overlooked the second criterion, and the possible loss of connectivity on higher levels of the pyramid which will be discussed in subsection 2.2.

It may be desirable to perform relinking for many vertices of a level in parallel. Then, extra care must be taken in order not to loose connectivity: two vertices which can be relinked individually from a given situation may not always be relinked simultaneously. Only a subset of the vertices in a level can be relinked to a new parent in each step. A safe strategy is to relink a vertex v only if none of the other vertices in $\kappa(\pi(v))$ or $N(v)$ is relinked.

A suitable subset of vertices which may be relinked in parallel can be computed by a stochastic procedure, similar to stochastic decimation [12, 14]. Each vertex v for which a relinking might be performed draws a random variable from some distribution. If that random number is larger than that drawn by all the vertices in $\kappa(\pi(v)) \cup N(v)$ which may not be relinked simultaneously, the vertex v 'wins the right' to be relinked, and the other vertices in $\kappa(\pi(v)) \cup N(v)$ are prohibited to relink. New random numbers can be drawn and new vertices selected as long as there are vertices which are neither selected for relinking nor prohibited to relink.

2.2. Connectivity Preservation in Higher Levels

Thus far, we have ignored the connectivity of receptive fields of vertices at levels $i > k+1$. If the pyramid is constructed bottom-up, like in section 3, such levels do not exist at the moment of relinking: first, levels 0 and 1 are built and the parent-child links between levels 0 and 1 are processed by relinking. Then level 2 is constructed and the parent-child links between levels 1 and 2 are processed by relinking, et cetera. In this case, no higher levels are present during the relinking steps.

If relinking is performed when higher levels are present, the connectivity of receptive fields in higher levels of the pyramid can be lost, as illustrated in figure 2.3. Suppose a vertex $v \in V_k$ with parent $p_{\mathrm{old}} \in V_{k+1}$ is relinked to a new parent $p_{\mathrm{new}} \in V_{k+1}$. For a vertex $w \in V_{k+n+1}$,

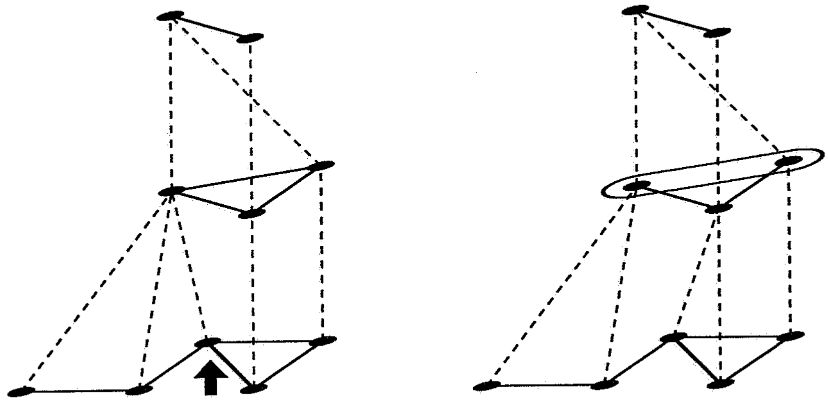


Figure 2.3: If edges in the parent level disappear by a relinking step, the connectivity in some higher level can be lost.

$n > 1$, the set $\kappa(w)$ is not changed by relinking. Connectivity of the receptive field of w can only be lost by the removal of edges between vertices in $\kappa(w)$. On level $k+n$, an edge which might be removed has the form $(\pi^n(v), \pi^n(v'))$, where $v' \in N(v)$. Removal of this edge may destroy the connectivity of $R(\pi^{n+1}(v))$. If the connectivity of a connected graph is lost by the deletion of a single edge, this edge is called a bridge [7]. The edge is indeed removed if all the edges in level k between $\kappa^n\pi^n(v)$ and $\kappa^n\pi^n(v')$ have v as an end point, i.e. $\kappa^n\pi^n(v) \cap N(\kappa^n\pi^n(v')) = \{v\}$. Moreover, v must not end up in the receptive field of $\kappa^n(v)$ or $\kappa^n(v')$ after the relinking step. Therefore, the condition $\pi^{n-1}(p_{\text{new}}) \notin \{\pi^n(v), \pi^n(w)\}$ must hold.

Summarizing, relinking of vertex $v \in V_k$ to $p_{\text{new}} = \pi(v')$, with $v' \in N(v)$, destroys the connectivity of the receptive field of $\pi^{n+1}(v)$ if the following conditions are satisfied:

1. $(\pi^n(v), \pi^n(v'))$ is a bridge in $\kappa\pi^{n+1}(v)$.
2. $\kappa^n\pi^n(v) \cap N(\kappa^n\pi^n(v')) = \{v\}$.
3. $\pi^{n-1}(p_{\text{new}}) \notin \{\pi^n(v), \pi^n(v')\}$.

These conditions can be tested for each vertex v and new parent p_{new}. For each vertex v this requires inspection of a set of vertices of the form $\pi^m(v)$, up to the level where $|\pi^m(N(v) \cup \{v\})| \leq 2$. Using these conditions, the set of possible parents for each vertex can be restricted in such a way that connectivity of receptive fields is not destroyed by relinking.

Just like we described in the previous subsection, it is possible to generate sets of vertices which can be relinked in parallel. Again, there are combinations of vertices which cannot be relinked simultaneously. As before, a random procedure can be used to compute such a set. In the present situation, however, vertices which may not be relinked simultaneously can be far apart, and propagation of information through the levels of the pyramid is required to compare the random labels for all mutually exclusive pairs. This makes the resulting algorithm is rather complicated.

2.3. The Relinking Method

In the previous subsections, it has been described how a set of candidate parents for a vertex can be computed. This subsection describes the relinking procedure. In the present method, vertices are visited one by one. For each vertex, a new parent is chosen from the set of allowed candidate parents. The vertex is relinked to the new parent and the graph structure and attributes of vertices are updated accordingly. This procedure is repeated until a stable configuration is reached.

The new parent p_{new} is chosen for each vertex v such that the difference between $g(v)$ and $g(p_{\text{new}})$ is minimized.

The rules which determine the selection of a new parent for a vertex can be formulated as an energy minimization problem. For each level m in the hierarchy, an energy

$$E_{\text{region}}(m) = \sum_{v \in V_m} n(v)\left[g(v) - g(\pi(v))\right]^2, \tag{2.1}$$

can be defined. Here, $n(v)$ denotes the area of the receptive field of v and $g(v)$ denotes its grey level.

The choice method used here is a steepest descent method: a vertex $v \in V_k$ is relinked to a new parent in such a way that $E_{\text{region}}(k)$ is reduced as much as possible in each step. This can cause the energy to converge in a local minimum. Spann [18] proposed a stochastic relinking algorithm, similar to simulated annealing, which tries to avoid local minima.

The relinking method can be shown to converge by an argument due to Cibulskis and Dyer [6]. Whenever a vertex in level k is relinked, the energies $E_{\text{region}}(i)$ with $i < k$ are not changed, while the energy $E_{\text{region}}(k)$ is reduced. This implies that the energy of the base level is non decreasing. As there is only a finite number of possible configuration for the links between the two lower levels, these links must reach a stable configuration in a finite number of steps. When this has happened, $E_{\text{region}}(0)$ will remain constant and the energy $E_{\text{region}}(1)$ will start to decrease monotonically. Therefore, the links between G_1 and G_2 will finally reach a stable configuration, et cetera.

A different way of formulating the region based relinking criteria uses the score Q_{region} defined by

$$Q_{\text{region}} = -|g(v) - g(p_{\text{new}})|. \tag{2.2}$$

The new parent will be the one which maximizes this score. When the maximal score equals $-|g(v) - g(p_{\text{old}})|$ for all vertices, the relinking procedure has converged.

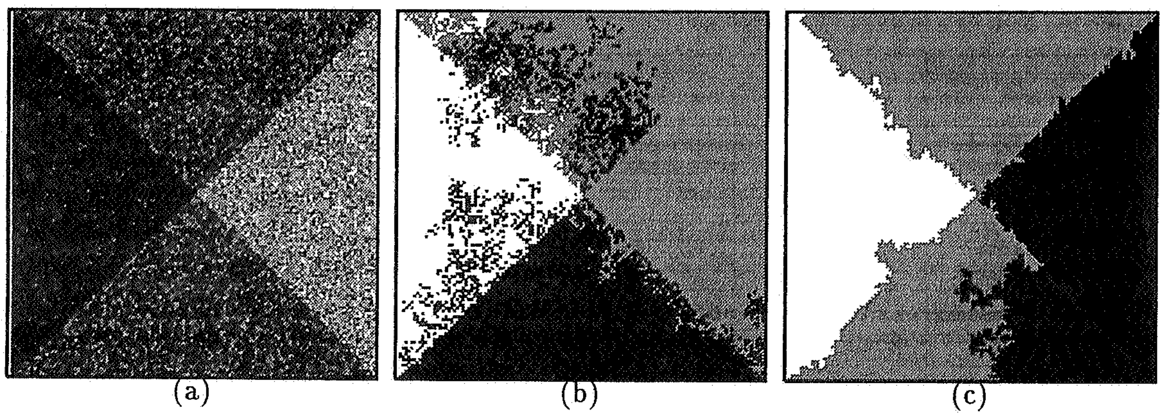


Figure 2.4: With the classical relinking algorithm, noise can scatter points of each class in the image. With the new method, classes are connected. From left to right: a synthetic image, corrupted with noise; its segmentation by the classical algorithm; its segmentation by the connectivity preserving algorithm.

The effect of connectivity preservation in relinking can be seen from figure 2.4. The initial image contains four regions of grey values 64, 128, 128 and 192, corrupted with Gaussian noise with standard deviation $\sigma = 32$. If the image were divided in four squares along horizontal and vertical lines, the initial configuration of the parent-child links would already represent the correct segmentation, and no conclusions can be drawn from the segmentation of such an image. Therefore, the image is segmented along diagonal lines. The middle image shows, in false grey scale, the subdivision in 4 classes computed by the classical algorithm of Burt *et al.* [5]. The image to the right shows the classification into four classes, obtained by connectivity preserving relinking. In can be seen that Burts algorithm scatters the points in each class over the image, while the connectivity preserving method generates connected classes, which are less affected by noise.

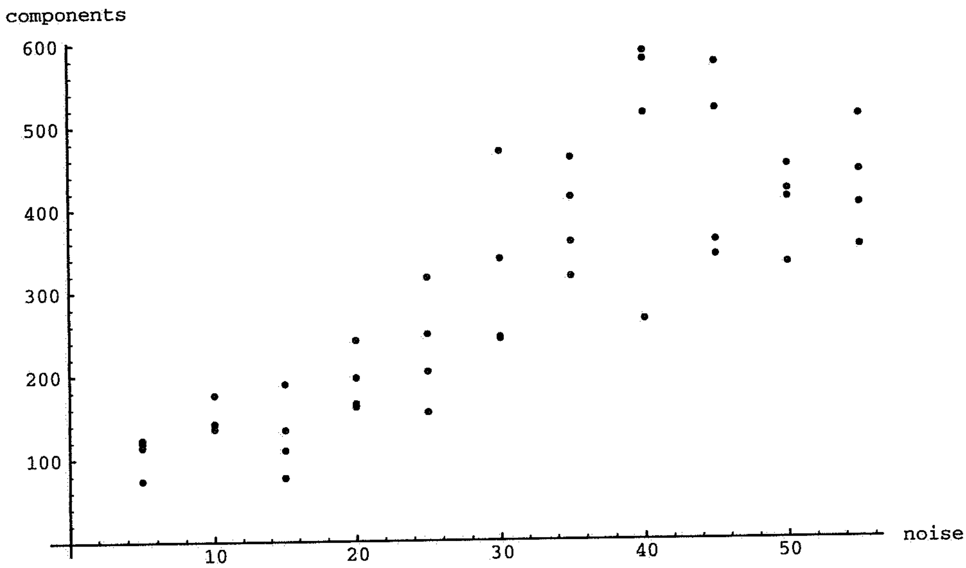


Figure 2.5: The number of connected components generated by Burt's algorithm at various noise levels

The number of 8-connected components generated by Burts relinking algorithm depends on the amount of noise added. In figure 2.5, the number of connected components is shown for a number of realizations of Gaussian noise with varying standard deviations. It can be seen that the number of connected components is large, even for very low noise levels. Moreover, the number of connected components for a given noise level assumes values in a rather broad range.

It can also be seen that the boundaries detected by the connectivity preserving method are not always smooth. This is caused by the fact that relinking is a steepest descent procedure. In order to remove the indentations present in some parts of the boundary, a number of relinking steps would have to be performed in which the difference between parent and child grey values is incremented temporarily.

3. Adaptive Construction of Successive Levels

As mentioned before, classical relinking suffers from a number of problems. Firstly, not all segmentations can be represented in a regular structure: elongated objects cannot be represented [4]. Moreover, the number of regions in each level is fixed. Originally, relinking was used for images containing a single light object on a dark background, for which this restriction is not a problem, but for segmentation of more general images, the number of regions must be adapted to image contents. Various authors [2, 10, 19] have described methods for root detection, i.e. the marking of cells in the pyramid as the representatives for a region in the final segmentation. This can improve the results, but the methods still operate in a regular structure.

These problems can be avoided if a hierarchy of region adjacency graphs is built level by level, adaptive to image contents. This idea has been used by Montanvert *et al.* [14] for image segmentation.

In this section, level by level construction of a hierarchy is combined with relinking. This is done as follows: from the base level G_0, the next level G_1 is constructed and the parent-child links between these levels are initialized. The parent-child links which are created in this phase have the same role as the regular structure in the classical relinking scheme: they serve as an initial configuration which is adapted by relinking. This adaptation is performed by the procedure described in section 2.

Then the second level G_2 is constructed. The parent-child links between G_1 and G_2 are initialized and updated by relinking, et cetera.

In order to create the vertices of a new level in the hierarchy, the vertices of the lower level are partitioned into a number of connected sets or *clusters*. Each of these clusters will be the set of children of a vertex on the next level. Like in graph decimation [14], each cluster will consist of central vertex and a number of its neighbors.

In order to avoid adjacent, but dissimilar regions to be merged into a single cluster, such vertices are forbidden to become members of a single cluster. Dissimilarity of adjacent vertices is defined using an edge strength measure. We consider two choices for the edge strength. The first one is defined by

$$S^{\mathrm{sub}}_{\mathrm{global}}(v,w) = |g(v) - g(w)| - \frac{1}{2}(\sigma(v) + \sigma(w)), \tag{3.1}$$

where $g(v)$ is the average grey value within the receptive field of a vertex and $\sigma(v)$ is the standard deviation of the grey value. The second choice is

$$S^{\mathrm{div}}_{\mathrm{global}}(v,w) = \frac{|g(v) - g(w)|}{1 + \frac{1}{2}(\sigma(v) + \sigma(w))}. \tag{3.2}$$

Both measures depend on the difference of the grey values of the regions, corrected for variations within each region. The measure S_{div} has the advantage that it is dimensionless. On the other hand, the estimation for $\sigma(v)$ and $\sigma(w)$ can be bad for small regions, yielding a large uncertainty in the value of $S^{\mathrm{div}}_{\mathrm{local}}$.

If each vertex stores the size, the average grey value and the standard deviation of its receptive field, these values can be recomputed for each vertex by considering only the corresponding values of its children after each relinking step. Therefore, local communication in the hierarchy suffices for the execution of these computations.

In the lower levels of the hierarchy, receptive fields are very small, maybe just a single pixel. Small receptive fields are relatively homogeneous, so the corresponding edge strengths are high. Therefore, the strength of the boundary between small receptive fields must be corrected for region size. This is achieved by multiplying all edge strengths with a geometry factor of the form

$$\left(1 - \frac{1}{A(v) + \delta}\right) \times \left(1 - \frac{1}{A(w) + \delta}\right), \tag{3.3}$$

where $A(v)$ is the area of the receptive field of v and δ is some small number. This factor approaches 1 for large regions, but is small for small regions. For single pixel receptive fields it is approximately equal to δ^2. Thus, survival of small noise regions to higher levels of the hierarchy is suppressed. Vertices are said to be dissimilar when the strength of the edge between them is larger than some predefined value.

Now that dissimilarity of vertices has been defined, the determination of clusters can be performed. Clusters will have the following properties:

1. For each cluster $C \subset V$, there is a center vertex $c \in C$ such that $C \subset N(c) \cup \{c\}$.
2. No cluster contains a pair of adjacent, dissimilar vertices.
3. Two center vertices c_1 and c_2 can be neighbors only if the cluster of c_1 contains a dissimilar neighbor of c_2 or vice versa.

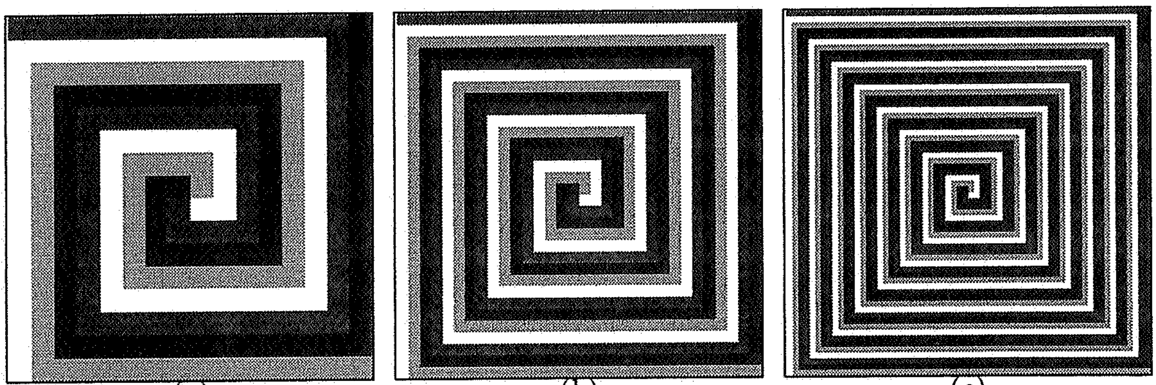
The difference with graph decimation is that, in graph decimation, pairs of dissimilar vertices in a cluster are forbidden only if one of these vertices is the center of the cluster, while here, all pairs in a cluster are to be considered. If conventional graph decimation were used here, undesirable merges of regions can occur when there is a small region on the boundary between two large dissimilar regions, and this small region is selected as the center of a cluster.

The computation of clusters is similar to stochastic decimation, but some changes have to be made in order to avoid dissimilar vertices to be put in a single cluster. Clusters are computed by repeated application of the following steps

1. Every vertex v_i which is not yet part of a cluster is given some label λ_i.
2. Every vertex whose label is larger than that of all of its similar neighbors is selected as the center of a new cluster.
3. For each newly selected center c, a maximal subset of $N(c)$ containing no dissimilar pairs is added to complete the cluster.

The label λ_i can be a random number, but it can also be some image dependent value. Throughout this paper, the area of the receptive field of a vertex is used as its label. An additional random number is assigned to each vertex for resolving ties between regions of equal size.

The difference with the stochastic decimation procedure is in the order of the steps. In stochastic decimation, the computation of a maximal independent set by repeated selection of local maxima is completed before the clusters are computed by assignments of neighbors; here, a number of clusters is computed in each selection of local maxima.



(a) (b) (c)

Figure 3.1: Three 128×128 images, each containing four regions.

The effect of adaptive construction of subsequent levels is shown by comparing the segmentations of some spiral images (figure 3.1). These are 128×128 images, each containing four regions. Figure 3.2 shows the segmentation of these images. The top row shows the segmentation according to the method described in the previous section; the bottom row shows the segmentation computed with adaptive construction.

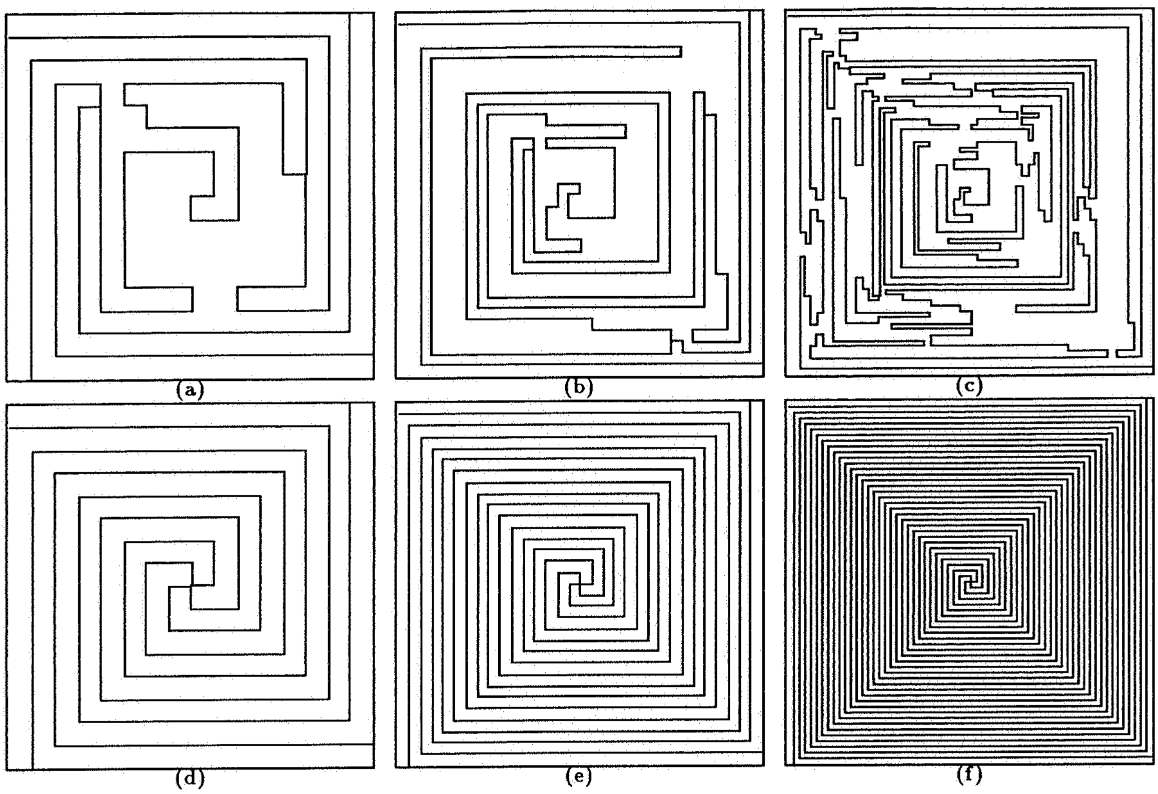


Figure 3.2: Segmentation of images with four connected regions using the regular pyramid (top row) and adaptive construction of levels (bottom row).

In can be seen that the regular pyramid is too rigid for the detection of strongly elongated regions. The adaptively constructed pyramid, on the other hand, finds these regions exactly.

4. Combining Region and Edge Information

Image segmentations can be based on a variety of criteria. Two important groups of criteria are those using properties of regions and those using properties of the boundaries between regions. In a hierarchy of graphs formalism, both types of information can be represented. In this section, it is described how region and boundary information can be combined in a relinking scheme.

Region based methods presuppose a model for homogeneous regions. It can for example be assumed that each region has a uniform grey value, or that the grey value of each pixel is drawn from some distribution, which is the same for each pixel in a homogeneous region. Models for different types of images, such as flow fields [9] or textured images [16], can also be used. In relinking methods using region information, the new parent is chosen in such a way that the fit of the receptive fields of the old and new parent vertices are improved in each step.

Segmentation methods based on boundary information detect differences between the properties of adjacent locations in the image, for example by the application of an edge detection filter. Boundaries between regions are then detected, for example by the application of a peak filter on the gradient or by the computation of the zero crossings of a second order difference operator, which correspond to the extrema of a first order difference operator. Segmentation methods using peak filters can have problems with the generation of compact regions, because if a small part of the boundary between two regions is missing, these regions will be merged. Segmentation methods based on zero crossings do generate compact regions, but have problems with situations where three regions meet in a point.

In conventional relinking schemes, there is no representation available for connected regions and the boundaries between such regions. Therefore, these schemes cannot handle boundary information. Such information can be represented in a hierarchy of graphs. This was done for the first time by Montanvert and Bertolino [13], using a randomized edge detector which can detect edges at all scales simultaneously [3].

The segmentation method presented in the previous section consists of two parts: the relinking procedure and the procedure for the adaptive construction of new levels. In both parts, the combination of region and edge information can be used.

4.1. Combining Region and Boundary Information for Relinking

In the previous section, we described the selection of a new parent in terms of energy minimization. The same approach will be taken here in order to incorporate boundary information in the relinking procedure.

In grey level based relinking as described in section 2, the energy being minimized was

$$E_{\text{region}}(m) = \sum_{v \in V_m} n(v)\left[g(v) - g(\pi(v))\right]^2, \tag{4.1}$$

which corresponds with a choice criterion using the score

$$Q_{\text{region}} = -|g(v) - g(p_{\text{new}})|. \tag{4.2}$$

Boundary-based segmentation evaluates regions by considering the edges surrounding each region. The response of an edge detector should be strong on the boundary of a region, while it should be weak in the interior of a region.

The strength of each edge is measured as the average response of an edge detector along the boundary represented by the edge. For the results presented here, the Sobel edge detector was used. Boundary-based image segmentation in a hierarchical structure can be performed by minimization of the sum over all regions of the average edge strength over the boundary of each region. When each edge carries its length and its average edge detector response as attributes, the average edge response of an edge can be computed recursively. When relinking is performed for a given level in a hierarchy, each vertex on the upper level corresponds to a region, and the boundaries of such regions are composed of active edges in the lower level.

Let $\eta(R(v))$ denote the average response of an edge detection filter along the receptive field of a vertex v. Our boundary based relinking criterion will be based on the minimization of the energy

$$E_{\text{boundary}}(m) = \sum_{v \in V_{m+1}} \eta(R(v)), \tag{4.3}$$

In each level of the hierarchy, the boundary around a vertex v, represented by edges between v and its neighbors, consist of boundary parts on the previous level. These boundary parts correspond to edges between vertices in $\kappa(v)$ and vertices not in $\kappa(v)$. An edge (v, w) for which $\pi(v) = \pi(w)$ is not part of some boundary in the next level; such an edge is called inactive. If $\pi(v) \neq \pi(w)$, the edge (v, w) represents a boundary fragment which is part of the boundary between the receptive fields of $\pi(v)$ and $\pi(w)$. Such an edge is called an active edge.

If a vertex v is relinked from an old parent p_{old} to a new parent p_{new}, some active edges become inactive and vice versa. The edges which become inactive are the edges (v, v') with $\pi(v') = p_{\text{new}}$; the edges which become active are the edges (v, v') with $\pi(v') = p_{\text{old}}$.

The strength of each edge is measured as the average response of an edge detector along the boundary represented by the edge. For the results presented here, the Sobel edge detector was used. Boundary-based image segmentation in a hierarchical structure can be performed by minimization of the sum over all regions of the average edge strengths over the boundary of each region. When each edge carries its length and its average edge detector response as attributes, the average edge response of an edge can be computed recursively. When relinking is performed for a given level in a hierarchy, each vertex on the upper level corresponds to a region, and the boundaries of such regions are composed of active edges in the lower level.

When a vertex v is relinked from p_{old} to p_{new}, the only average edge strengths which change are those of the receptive fields of p_{old} and p_{new}. Thus, the total change in the average edge strength is

$$\eta\big(R(p_{\text{old}}) \setminus R(v)\big) + \eta\big(R(p_{\text{new}})\big) - \eta\big(R(p_{\text{old}})\big) - \eta\big(R(p_{\text{new}}) \setminus R(v)\big), \tag{4.4}$$

where $\eta(\cdot)$ denotes the average edge strength along the boundary of a region.

In order to maximize the average edge strength, in each time the new parent must be chosen such that the score

$$Q_{\text{edge}} = \eta\big(R(p_{\text{new}})\big) - \eta\big(R(p_{\text{new}}) \setminus R(v)\big) \tag{4.5}$$

is maximal. When the maximal score is equal to $\eta\big(R(p_{\text{old}})\big) - \eta\big(R(p_{\text{old}}) \setminus R(v)\big)$, the relinking procedure has converged. Convergence of this relinking procedure can be proven with the argument used before.

Combining region and edge information in a relinking method can be done by using a combined score

$$Q_\alpha = \alpha Q_{\text{edge}} + (1-\alpha) Q_{\text{region}}, \tag{4.6}$$

where α is a weight factor between 0 and 1. Again, the relinking method can be shown to converge with the argument used before [6]

4.2. Combining Region and Boundary Information for Construction of New Levels

The second part of our segmentation method is the construction of new levels, which is performed adaptively to image contents through the influence of vertex dissimilarities. It will now be shown how region and boundary information can be combined in the definition of these dissimilarities. This can be done simply by choosing a region dissimilarity measure which depends on both region and edge information.

In section 3 we considered the edge strengths

$$S^{\text{sub}}_{\text{global}} = |g(v) - g(w)| - \frac{1}{2}(\sigma(v) + \sigma(w)) \tag{4.7}$$

and

$$S^{\text{div}}_{\text{global}} = \frac{|g(v) - g(v)|}{1 + \frac{1}{2}(\sigma(v) + \sigma(w))} \tag{4.8}$$

These edge strengths are based on the average grey values of adjacent regions, i.e. on global properties. Local properties of the boundaries between adjacent regions can be measured by considering the average response of an edge detector along this boundary. Let $s_2(v,w)$ denote the average response of an edge detector along the boundary of the receptive fields $R(v)$ an $R(w)$ and let $s_1(v)$ denote the average response within the receptive field $R(v)$.

Then the strength of an edge, corrected for the structure within regions, can be expressed as

$$S^{\text{sub}}_{\text{local}}(v,w) = s_2(v,w) - \frac{1}{2}(s_1(v) + s_1(w)) \tag{4.9}$$

or as

$$S^{\text{div}}_{\text{local}}(v,w) = \frac{s_2(v,w)}{1 + \frac{1}{2}(s_1(v) + s_1(w))}. \tag{4.10}$$

A combined dissimilarity measure can then be defined as

$$S^{\xi}_{\alpha}(v,w) = \alpha S^{\xi}_{\text{local}}(v,w) + (1-\alpha) S^{\xi}_{\text{global}}(v,w), \tag{4.11}$$

where ξ denotes either 'div' or 'sub'. In all cases, the value of α used here will be the same values as the one used previously for weighing scores in the relinking method described earlier in this section.

5. Results

The elements described in the previous sections have been combined in a segmentation algorithm. A hierarchy is constructed in a bottom-up order. Alternatingly, a new level is constructed by decimation of a region adjacency graph, and the parent-child links between the lower and the upper level are updated by relinking. Relinking is performed in a connectivity-preserving manner, and region and edge information are combined both in the construction of the similarity graph and during relinking. This section presents the segmentation results for some synthetic and natural images.

The effect of combining region and edge information is illustrated in figure 5.1. It shows an 128×128 image consisting of four ramps with grey values from 112 to 144, contaminated with Gaussian noise with standard deviation $\sigma = 8$. The four regions cannot be discerned based on their average grey values, because these are all the same. Therefore, boundary information must be used. On the other hand, in the lower levels, information on average grey values is more suitable for the classification of pixels, because boundary information is rather noisy for small regions. α with edge strength S^{sub}_{α}. Therefore, some intermediate value of α must be used for the optimal segmentation of the ramps image.

The segmentations, computed with edge strength S^{sub}_{α}, are shown for $\alpha = 0.95$, $\alpha = 0.5$ and $\alpha = 0.1$. If α is too large, the region boundaries are disturbed because of disturbances in the lower levels of the hierarchy. If, on the other hand, α is too small, the average grey level criterion dominates and the regions are split. Increasing the threshold with low values of α causes all the regions to be merged, possibly leaving some small noise segments.

Figures 5.2 and 5.3 show the number of misclassified pixels (with respect to the original four bands image) for various values of α and for different realizations of the noise. Figure 5.2 shows the situation for the strength measure S^{sub}_{α} and figure 5.3 shows the situation for S^{div}_{α}. It can be seen that the quality of the segmentation is good for a wide range of α values. Note that, in one case in figure 5.2, the segmentation with $\alpha = 0.65$ has a large number of misclassified pixels. This is caused by a 'wrong' link in a high level of the hierarchy, in which a large receptive field is involved.

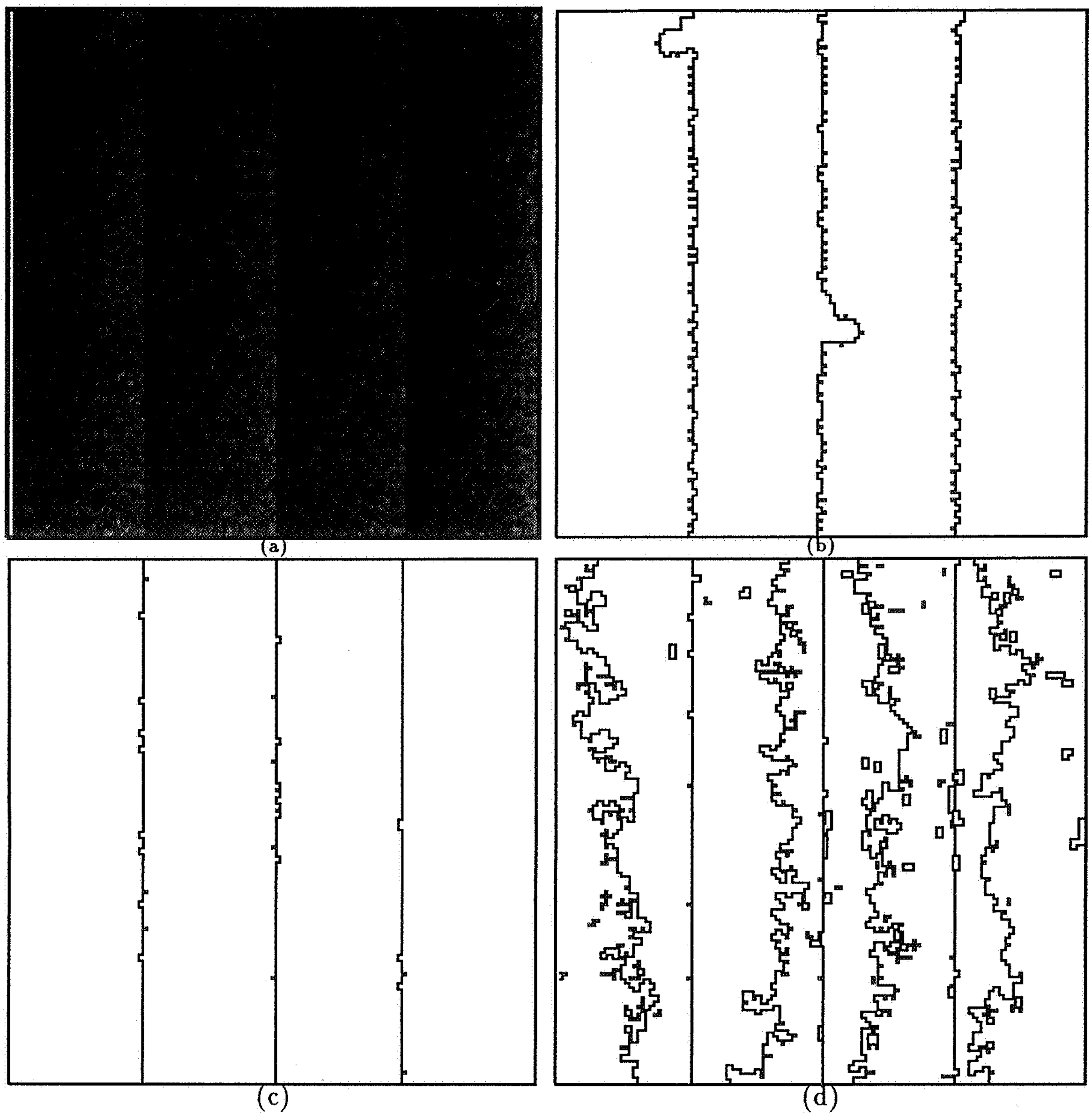


Figure 5.1: A synthetic image (a), contaminated with noise and its segmentations with $\alpha = .95$ (b), $\alpha = .50$ (c) and $\alpha = .10$ (d).

When S_α^{div} is used, a satisfactory segmentation can be obtained with $\alpha = 1$. On the other hand, the segmentation method breaks down for $\alpha < 0.5$. A wider range of possible values for α is available when S_α^{sub} is used.

Figure 5.4 shows the results obtained for an MRI image of a head, both for the edge strength $S_{0.5}^{\text{div}}$ (figure 5.4a, threshold 1.5) and the edge strength $S_{0.5}^{\text{sub}}$. (figure 5.4b, threshold 20).

Figure 5.5 shows the segmentation result for an image of muscle tissue without (a) and with (b) relinking. As the individual muscle fibers do not have a strictly homogeneous grey value, edge information must be emphasized more in the segmentation process. Therefore, $\alpha = 0.75$ was used.

It can be seen that the construction without relinking extracts some of the structures from

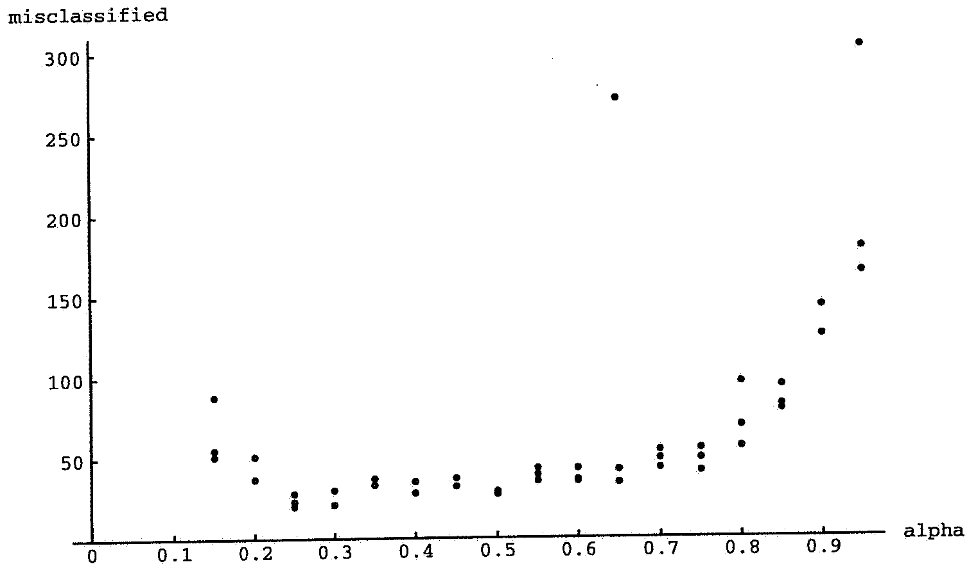


Figure 5.2: The number of misclassified pixels as a function of the parameter α with edge strength S_α^{sub}.

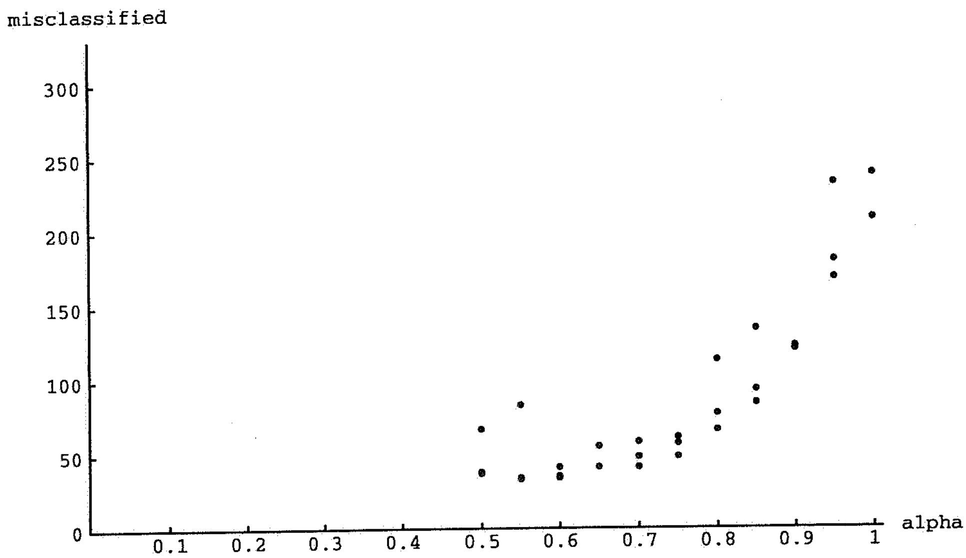
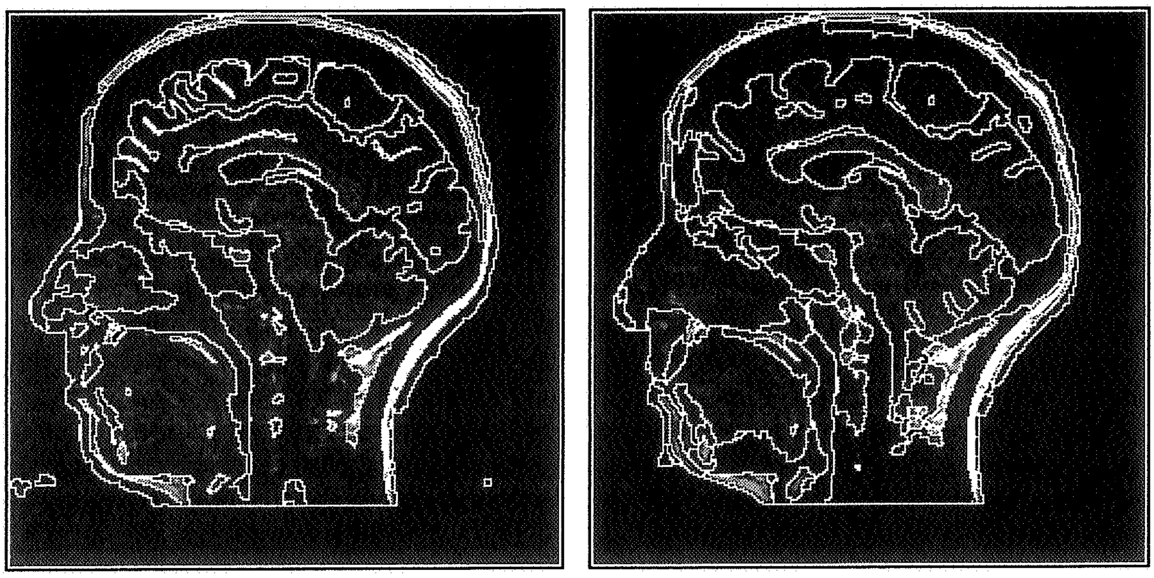


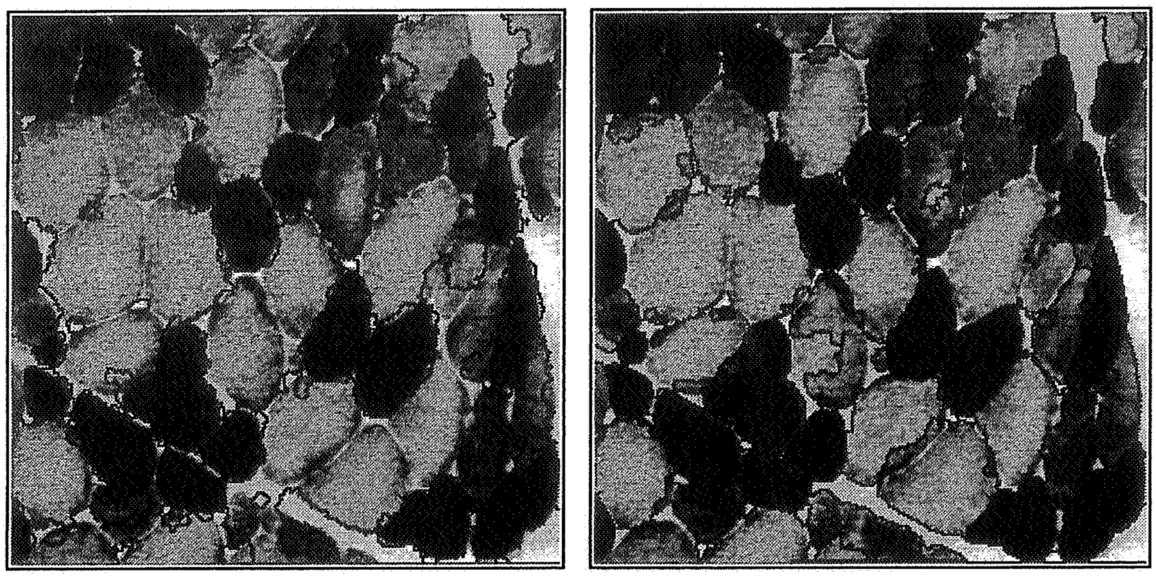
Figure 5.3: The number of misclassified pixels as a function of the parameter α with edge strength S_α^{div}.

the image, but improvement can be made by applying relinking. The segmentation obtained when relinking is used shows more individual fibers than the segmentation obtained without relinking. This can be understood as follows. Each time a level is constructed by graph decimation, the boundaries between the resulting regions are not located optimally. Therefore, the response of an edge detection filter will not be maximal on the boundaries found by the procedure, but more in the interior of some regions. Similarly, the regions found by graph decimation do not correspond exactly with the boundaries between homogeneous regions in the image. Therefore, a region detected by the decimation process can contain pixels from different regions in the image. This causes an increase in the standard deviation of the grey



(a) (b)

Figure 5.4: Segmentation with edges strengths $S_{0.5}^{\text{div}}$ (a) and $S_{0.5}^{\text{sub}}$ (b).



(a) (b)

Figure 5.5: Segmentation for the muscle image without (a) and with (b) relinking. The edge strength S_{α}^{div} was used with $\alpha = 0.75$ and threshold 1.25.

levels in a region.

If the boundaries are not corrected in the subsequent relinking phase, these effects reduce the edge strengths S_{α}^{div} and S_{α}^{sub}. More adjacent vertices will be similar and therefore, more regions will be merged in the construction of subsequent levels.

6. Conclusions

In this paper, we have presented the extension of pyramid relinking to hierarchies of graphs. This approach solves a number of problems associated with conventional relinking methods.

The first problem we solved was the problem of connectivity preservation. Conventional relinking does not produce connected regions, but classes which may be scattered in the image plane. We have been able to adapt the relinking rules in such a way that connectivity of regions is guaranteed. When the pyramid is built level by level, two criteria (illustrated in figure 2.2) must be used to select suitable relinking steps. When relinking is performed in a complete pyramid, in which all levels are present, a third criterion (illustrated in figure 2.3) must be used in order to avoid the destruction of the connectivity of receptive fields of vertices in higher levels.

The artifacts from which relinking in regular pyramids suffers were avoided by constructing the hierarchy level by level, adaptive to image contents. The construction of a new level can be seen as a first guess for the aggregation of region primitives which is improved by the relinking procedure.

The solution to the connectivity preservation problem required the introduction of a graph structure. The edges in this graph correspond with the boundaries between receptive fields in the image plane. This provides the possibility to represent information on these boundaries in the data structure and to combine region and boundary information in the segmentation process. It has been shown that the combination of region and edge information can be useful and that the corresponding segmentation scheme is robust under the change of the parameter α.

Some segmentation results for natural images have been presented, showing that satisfactory results can be obtained in practice.

Acknowledgements: Most of the work presented here was performed while the author was a guest of the department for Pattern Recognition and Image Processing of the Technical University of Vienna, Austria. The author thanks Prof. W.G. Kropatsch and his department for their hospitality and many interesting discussions on the subject. The advice of A. Toet, F. Groen and H. Heijmans is also gratefully acknowledged.

[1] H.J. Antonisse. Image segmentation in pyramids. *Computer Graphics and Image Processing*, 19:367–383, 1982.

[2] S. Baronti, A. Casini, F. Lotti, L. Favaro, and V. Roberto. Variable pyramid structures for image segmentation. *Computer Vision, Graphics and Image Processing*, 49:346–356, 1990.

[3] P. Bertolino and A. Montanvert. Edge detection for biomedical images: A self-adaptive and randomized operator. In *14th Annual International Conference of IEEE Engineering in Medicine and Biology Society*, pages 1904–1905, Paris, France, 1992.

[4] M. Bister, J. Cornelis, and A. Rosenfeld. A critical view of pyramid segmentation algorithms. *Pattern Recognition Letters*, 11:605–617, 1990.

[5] P. J. Burt, T. Hong, and A. Rosenfeld. Segmentation and estimation of image region properties through cooperative hierarchical computation. *IEEE Transactions on Systems, Man and Cybernetics*, 11(12):802–809, 1981.

[6] J. Cibulskis and C.R. Dyer. Node linking strategies in pyramids for image segmentation. In A. Rosenfeld, editor, *Multiresolution Image Processing and Analysis*, chapter 8, pages 109–120. Springer Verlag, Berlin etc., 1984.

[7] F. Harary. *Graph Theory*. Addison-Wesley Publishing Company, 1969.

[8] R. Hartley. *Multi-Scale Models in Image Analysis*. PhD thesis, University of Maryland, 1984.

[9] R. Hartley. Segmentation of optical flow fields. Technical Report TR86, University of Maryland, Institute of Automation Research, 1984.

[10] T.H. Hong and A. Rosenfeld. Compact region extraction using weighted pixel linking in a pyramid. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(2):222–229, 1984.

[11] S. Kasif and A. Rosenfeld. Pyramid linking is a special case of ISODATA. *IEEE Transactions on Systems, Man and Cybernetics*, 13(1):84–85, 1983.

[12] P. Meer. Stochastic image pyramids. *Computer Vision, Graphics and Image Processing*, 45:269–294, 1989.

[13] A. Montanvert and P. Bertolino. Irregular pyramids for parallel image segmentation. In *Proceedings of the 16tm ÖAGM Meeting*, Vienna, Austria, 1992.

[14] A. Montanvert, P. Meer, and A. Rosenfeld. Hierarchical image analysis using irregular tesselations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(4):307–316, 1991.

[15] P.F.M. Nacken. Top-down image analysis by cost minimisation in hierarchical graph structures. to be published.

[16] M. Pietikäinen and A. Rosenfeld. Image segmentation by texture using pyramid node linking. *IEEE Transactions on Systems, Man and Cybernetics*, 11:822–825, 1981.

[17] R. Sedgewick. *Algorithms, second edition*. Addison-Wesly Publishing Company, Inc., Reading, Mass., etc., 1988.

[18] M. Spann. Figure/ground separation using stochastic pyramid relinking. *Pattern recognition*, 24(10):993–1002, 1991.

[19] M. Spann and C. Horne. Image segmentation using a dynaic thresholding pyramid. *Pattern recognition*, 22(6):719–732, 1989.

[20] S. W. Wharton. PIPS: A procedure for interactive pyramid segmentation. In *Proceedings of IGARS '88 Symposium*, pages 197–200, Edinburgh, Scotland, 13–16 September 1988. ESA Publications Division.