

Hierarchical Cell Structures for Segmentation of Voxel Images^{*}

Lutz Priese, Patrick Sturm, and Haojun Wang

Institute for Computational Visualitics,
University Koblenz-Landau, Koblenz, Germany

Abstract. We compare three hierarchical structures, S_{15} , C_{15} , C_{19} , that are used to steer a segmentation process in 3d voxel images. There is an important topological difference between C_{19} and both others that we will study. A quantitative evaluation of the quality of the three segmentation techniques based on several hundred experiments is presented.

1 Introduction

Segmentation of 3d voxel images is a rather new and important discipline as the request for 3d images is increasing in industry and medicine. Most standard segmentation or edge detecting techniques of 2d have been adapted to 3d, such as morphologic operations (e.g., [1]), watersheds (e.g., [2]), level sets (e.g., [3]), B-spline snakes (e.g., [4]), anisotropic diffusion filters (e.g., [2]), Sobel filters and gradient maps (e.g., [5]), etc.

However, generalizing the fast and robust 2d Color Structure Code (CSC) [6] leads to an interesting geometric challenge. The CSC is a hierarchical, inherently parallel, elaborated region growing technique that works with partially overlapping sub-segments: at level $n+1$ all neighbored and similar partial segments of level n inside a so-called island of level $n+1$ are merged into a new segment of level $n+1$, whereas the non similar but overlapping segments are splitted. This is done in parallel for all islands of the image. The island structure steers this homogeneous growing or splitting of segments all over the image. The quality of the segmentation depends heavily on certain topological properties of this underlying island structure.

In an ongoing research project of several research groups this 2d CSC is generalized to 3d voxel images. Thus, the 2d hierarchical overlapping island structure has to be generalized to a 3d hierarchical, overlapping 'cell' structure. Unfortunately, in 3d there is no cell structure with all the nice properties of the 2d island structure, complicating the generalization of the CSC.

We very briefly introduce the 2d island structure, I_7 , and the CSC based on I_7 . Then three generalizations of I_7 to 3d cells, S_{15} , C_{15} and C_{19} , are presented

^{*} This work was founded by the BMBF under grant 01/IRC01B (research project 3D-RETISEG).

and some topological properties are mentioned. We discuss the differences between S_{15} , C_{15} and C_{19} . A measurement of the coverability and error rates in more than thousand segmented 3d images compared with the 'perfect segmentation' gives a quantitative analysis of a CSC segmentation based on S_{15} , C_{15} , or C_{19} cells.

1.1 The Island Structure

The 2d island structure, I_7 , is from [7]. An island of level 0 in a hexagonal pixel structure consists of a center pixel and its six neighbor pixels. An island of level $n+1$ consists of a center island plus its six neighbor islands, each of level n . Two islands of level n are called neighbored if they overlap (in an island of level $n-1$), see Figure 1. Some properties of this structure are:

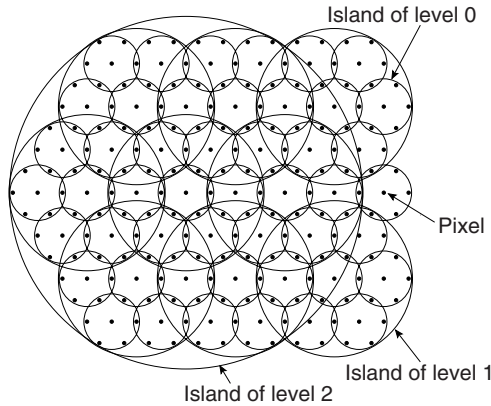


Fig. 1. The hierarchical hexagonal island structure I_7

- 6-Neighborhood: each island overlaps with six islands of the same level.
- Plainness: two islands of the same level overlap in at most one island of one level lower.
- Coverability: each island (except the topmost island) is a sub-island of a (parent) island of one level higher.
- Density: two neighbored islands are sub-islands of one common parent island of one level higher.

For a deeper discussion see [8]. This island structure steers a homogeneous region growing (the CSC) as follows:

Do for any island I of level 0: form a segmentation within the seven pixels of I .

Do for any island I of level $n+1$:

- Merge neighbored and similar segments of level n in I into a new segment of level $n+1$, and

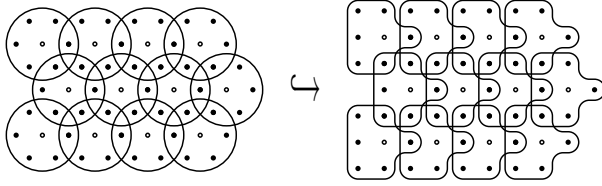


Fig. 2. Transformation of the hexagonal structure I_7 into the orthogonal grid

- If two neighbored segments s_1, s_2 of level n within I are not similar enough then split their common sub-region $s_{1,2}$ optimally into two sub-regions $s_{1,2}^1, s_{1,2}^2$ and merge $s_{1,2}^i$ with s_i .

Note, a merging of partial segments of level n is solely done within an island of level $n+1$. Thus, if such an island structure would violate 'coverability' all already created partial segments within an island that possesses no parent island would become completely lost for a further region growing. If 'density' is violated already created partial segments of two neighbored islands without a common parent island cannot become merged. Thus, a violation of 'coverability' or 'density' results principally in too many too small segments, an over-segmentation.

Two partial segments are only merged if they already possess a common sub-segment of one level lower. This hierarchical overlapping is the key for the quality of a CSC segmentation.

As pixels in real images are given in an orthogonal topology the hexagonal topology of the island structure has to be transformed. I.e., an island of level 0 now becomes simply a window as shown in Figure 2.

2 3D Cell Structures

We follow the idea of the most dense sphere packing. The 3d space is organized into 2d layers and the spheres must be placed onto a layer. Each sphere touches 6 spheres of the same layer (in the known hexagonal topology) and three spheres in each layer 'above' and 'below'. Thus, Sph_{13} consists of a center sphere and 12 neighbored spheres.

Any cell (large enough) can create a hierarchical overlapping cell structure by declaring any second cell of level n of any second macro row in any second macro layer as a center of a cell of level $n+1$ (of the same shape as the level 0 cell), compare Figure 1. In this hierarchical topology two cells of level $n+1$ are neighbored if their center cells of level n have a (macro) distance of 2 (measured in cells of level n). Thus, two neighbored cells of level $n+1$ possess common sub-cells of level n ; an overlapping is achieved. Unfortunately, such a Sph_{13} cell structure even violates coverability: on each level 1/8 of all cells don't possess a parent cell.

In a corrected version a Sph_{15} cell consists of a center cell plus its six neighbors in the same layer plus four cells in each layer above and below. Sph_{15} can be transformed in several ways in the orthogonal topology. A very compact

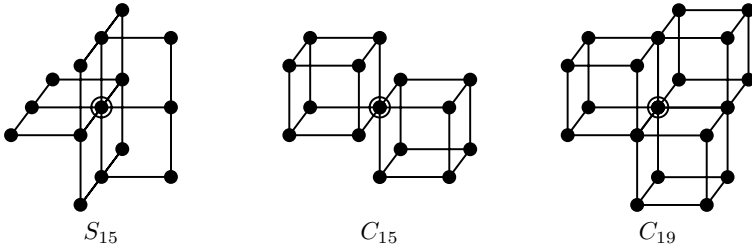


Fig. 3. The three cell structures

transformation is S_{15} , another one, more similar to a cube with holes, is C_{15} , see Figure 3. The Manhattan distance of the center to its remotest corner is 2 in S_{15} and 3 in C_{15} . Both cell structures, created by S_{15} and C_{15} , possess the wanted properties of 14-neighborhood (each cell overlaps with 14 cells of the same level), plainness (two cells overlap in 0 or 1 sub-cells of one level below), and coverability, but they violate density. Thus, a 3d CSC steered by a S_{15} or C_{15} cell structure must lead to over-segmentations.

This over-segmentation is avoided by a C_{19} cell as presented also in Figure 3. A C_{19} cell structure fulfills coverability and density. However, it violates plainness and the neighborhood property:

- Two cells may overlap in 0, 1 or 2 sub-cells of the level below,
- The center cell doesn't overlap with all other 18 sub-cells of a parent cell.

Further, a cell may be a sub-cell of 1, 2 or 3 parent cells. Those three deficiencies heavily complicate an implementation, but they don't lead to an over-segmentation.

3 Evaluation

All CSC versions (based on S_{15}, C_{15}, C_{19} cells) have been implemented. A S_{15} -CSC version will exist soon in hardware. We have applied all versions to real MR and CT (T1 and T2) medical images. A first inspection hinted to slightly better segmentation results for the S_{15} -CSC than for the C_{15} -CSC (possibly due to the more compact form of the S_{15} , although both are orthogonal realizations of the same cell Sph_{15}). But both are beaten by the C_{19} -CSC. A quantitative measurement with real MR and CT data is almost impossible as the theoretical correct segmentations are unknown and a hand segmentation by humans may be worse than by any CSC version. Therefore, we had to create artificial test images where the correct segmentation is known in advance.

3.1 Test Images

We created 16 test images of 256^3 voxels of 8 bit gray depth, the 'original images'. 8 images possess a background of gray value 88 and two 3d trees, plexus p_1, p_2 , of gray value 122 and 133. Both plexuses p_1 and p_2 don't touch each other but

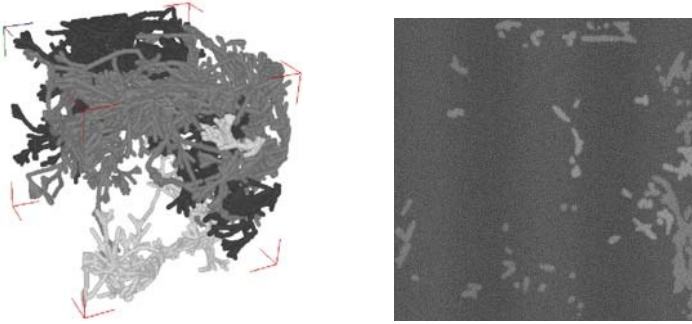


Fig. 4. Three plexuses plus a slice through the blurred image

have a minimal distance of 1 (4 images) or 2 (4 images) voxels. The diameter of each plexus is five voxels and their density varies. Further 8 images are created in the same way but with three plexuses in each of a gray value 122, 133, and 143. Thus, the correct classification, which voxel belongs to plexus p or to the background, is known.¹ Now, all images are blurred:

- A sinus curve along the z-axis adds values between -5.3 and +5.3,
- A 3d binomial filter (a threefold convolution with the vector $1/4(1\ 2\ 1)$ along the x-, y- and z-axis) is applied, and
- A Gaussian noise with $\sigma = 2$ (with $\sigma = 4$ and $\sigma = 6$, respectively) is added.

This gives us 48 'blurred images'. Figure 4 present an original image with three plexuses (where each plexus is visualized in an artificial clearly distinguished gray value, and the gray value of the background is set to transparent). A slice through its blurred form with $\sigma = 6$ in the original gray values is also shown in this Figure.

3.2 Statistics

In all CSC versions we declare two neighbored partial segments to be similar enough to become merged if the difference of their mean gray values is below a certain threshold t . All 48 test images (with 120 different plexuses) are segmented by all three CSC versions with different thresholds ranging from 6 to 17. This gives us more than 1.000 segmented images with some ten thousand segments. The perfect segmentation for each plexus is known, namely the plexus itself within the corresponding original image. Thus, for any plexus p and any set s of voxels we can compute in a blurred image the following:

The *coverability rate* $CR_p(s)$ of p by s , i.e. the number of voxels of s that belong to p , divided through the number of voxels inside p ,

The *error rate* of s (with respect to p), i.e. the number of voxels of s that do not belong to p , divided by the number of voxels inside s .

¹ Those images are free available under 'Download-Section: 3d plexus images' on <http://www.uni-koblenz.de/~lb>

We want to measure the quality of a segmented image. There are many possible ways to do this. The following two methods follow the idea that one wants to conclude from the segmented blurred image to the position of a plexus p in the (usually unknown) original image. Thus, the rates $CR_p(s)$ and $ER_p(s)$ will be unknown in practice. However, it is reasonable to assume that one can detect the correspondence between a segment s and a plexus p if $CR_p(s)$ is high and $ER_p(s)$ is low.

Quality Measure I: We call a segment s *fair* (with respect to a plexus p) if $ER_p(s) \leq 20\%$, i.e. more than 80% of the segment s must belong to p . For each p we regard at most the 25 segments s_i with the highest rates $CR_p(s_i)$.

For any segmented blurred image I' , for any plexus p in the original image I of I' do:

- if there exists a fair segment s with $CR_p(s) \geq 90\%$ then set output := s , otherwise:
- if there exists a union s of fair segments with $CR_p(s) \geq 90\%$ then choose such a union s with minimal error rate and set output := s , otherwise:
- if there exists a fair segment s with $CR_p(s) \geq 80\%$ then set output := s , otherwise:
- if there exists a union s of fair segments with $CR_p(s) \geq 80\%$ then choose such a union s with minimal error rate and set output := s , otherwise:
- if there exists a fair segment s with $CR_p(s) \geq 70\%$ then set output := s , otherwise:
- if there exists a union s of fair segments with $CR_p(s) \geq 70\%$ then choose such a union s with minimal error rate and set output := s , otherwise:
- choose a segment s with maximal $CR_p(s) - ER_p(s)$, set output := s .

For output s we measure $CR := CR_p(s)$, $ER := ER_p(s)$, $\# := \#_p(s)$, where $\#(s)$ is the number of segments used to produce s .

We group our images into 6 classes (d, σ) , for $d \in \{1, 2\}$, $\sigma \in \{2, 4, 6\}$. Each class (x, y) consists of all images with a minimal distance of x between two plexuses where a Gaussian noise with $\sigma = y$ was applied. Each class possesses 8 images with 20 plexuses. We present the mean values CR , ER , and $\#$ in each class for some thresholds with reasonable results. Using a threshold > 8 in the $d = 1$ classes, e.g., would result in error rates above 10%.

The line $\frac{d \ \sigma \ \text{cell} \ t \ CR \ ER \ \#}{2 \ 2 \ S_{15} \ 13 \ 92.6 \ 3.2 \ 11.6}$ of Table 1 reads as: in average 11.6 segments are required to cover in average 92.6% of a plexus with a mean error rate of 3.2% in the (2,2)-class if we have made a S_{15} -CSC segmentation with a threshold 13.

The idea of this measure is to find out how few segments are needed to cover a reasonable rate of a plexus. However, the last line of the algorithm, to use a segment with maximal difference of the coverability and error rate, $CR_p(s) - ER_p(s)$ is maximal, may lead to a segment of the background that touches relatively few voxels of p . In our next quality measure we will correct

Table 1. Statistics for Quality Measure I

d	σ	cell	t	CR	ER	#	d	σ	cell	t	CR	ER	#
2	2	S_{15}	13	92.6	3.2	11.6	1	2	S_{15}	8	77.7	0.6	19.8
2	2	C_{15}	13	92.7	3.5	9.8	1	2	C_{15}	8	77.8	0.6	19.9
2	2	C_{19}	13	98.6	3.1	1.0	1	2	C_{19}	6	76.9	0.6	14.2
2	4	S_{15}	13	90.8	4.9	13.1	1	4	S_{15}	8	73.4	9.2	19.5
2	4	C_{15}	13	91.1	5.2	11.8	1	4	C_{15}	8	73.6	9.7	21.0
2	4	C_{19}	13	96.9	5.0	1.0	1	4	C_{19}	8	79.1	16.8	11.7
2	6	S_{15}	10	73.9	4.2	17.1	1	6	S_{15}	8	64.6	1.6	21.7
2	6	C_{15}	10	73.6	4.4	17.1	1	6	C_{15}	8	64.1	1.7	22.6
2	6	C_{19}	10	83.1	6.6	13.8	1	6	C_{19}	6	55.9	2.4	25.0

this final step and will simplify the first steps: we just try to cover as much of p as possible, however only with 'reasonable' segments.

Quality Measure II: We call a segment s *reasonable* (with respect to p) if s is fair and $CR_p(s) > 0.1\%$, i.e. it is large enough.

For any segmented blurred image I' , for any plexus p in the original image I of I' do:

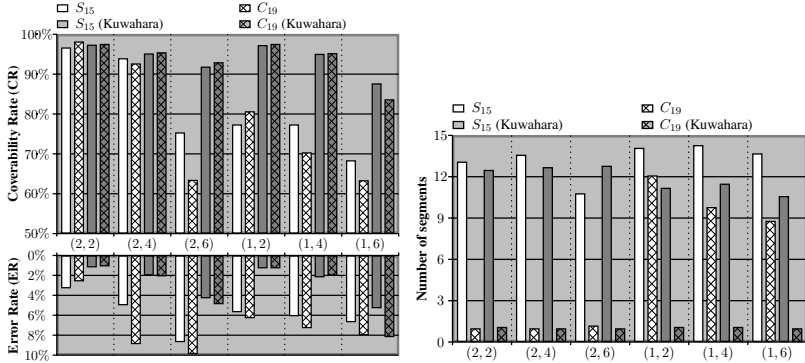
- build the union s of all reasonable segments;
- if $CR_p(s) \geq 60\%$ then set output := s , otherwise:
- if there exists a segment s with $ER_p(s) < 50\%$ choose such a segment s with a maximal coverability rate and set output := s , otherwise:
- set output := \emptyset .

For output $s \neq \emptyset$ we measure CR , ER , and $\#$, as before. For output = \emptyset we set $CR := 0$, $ER :=$ undefined, and $\# :=$ undefined.

The line $\frac{d \ \sigma \ \text{cell} \ t \ CR \ ER \ \#}{2 \ 2 \ S_{15} \ 13 \ 96.7 \ 3.3 \ 13.1}$ of Table 2 reads as: a S_{15} -CSC segmentation with a threshold 13 can cover with all reasonable segments in average 96.7% of each plexus in class (2, 2) with a mean error rate of 3.3%, where in average 13.1 segments are required. Table 2 is graphically visualized in Graph 1 and 2.

Table 2. Statistics for Quality Measure II

d	σ	cell	t	CR	ER	#	d	σ	cell	t	CR	ER	#
2	2	S_{15}	13	96.7	3.3	13.1	1	2	S_{15}	10	77.4	5.7	14.1
2	2	C_{15}	12	95.4	2.8	13.5	1	2	C_{15}	10	79.8	5.8	14.7
2	2	C_{19}	12	98.2	2.6	1.0	1	2	C_{19}	8	80.7	6.3	12.1
2	4	S_{15}	13	94.0	5.0	13.6	1	4	S_{15}	10	77.4	6.1	14.3
2	4	C_{15}	14	95.4	6.2	13.8	1	4	C_{15}	10	76.2	6.6	15.5
2	4	C_{19}	14	92.7	8.9	1.0	1	4	C_{19}	8	70.4	7.3	9.8
2	6	S_{15}	13	75.4	8.7	10.8	1	6	S_{15}	10	68.4	6.7	13.7
2	6	C_{15}	13	79.8	10.6	13.1	1	6	C_{15}	10	71.1	7.4	16.3
2	6	C_{19}	12	63.5	9.9	1.2	1	6	C_{19}	8	63.4	8.0	8.8



Graph 1. Number of segments in **Graph 2.** Coverability and error rate in each class (d, σ)

Table 3. Statistics for Quality Measure I for filtered Images

d	σ	cell	t	CR	ER	#	d	σ	cell	t	CR	ER	#
2	2	S_{15}	8	93.5	1.0	15	1	2	S_{15}	8	95.3	1.2	16.8
2	2	C_{15}	8	93.7	1.0	13.5	1	2	C_{15}	8	95.0	1.2	15.6
2	2	C_{19}	8	97.6	1.0	1.0	1	2	C_{19}	8	97.5	1.3	1.0
2	4	S_{15}	13	92.8	2.2	16.6	1	4	S_{15}	8	92.8	1.9	15.3
2	4	C_{15}	13	92.8	2.2	15.1	1	4	C_{15}	8	93.3	1.9	15.0
2	4	C_{19}	13	95.6	2.2	1.0	1	4	C_{19}	8	95.2	2.0	1.0
2	6	S_{15}	14	90.9	4.6	16.2	1	6	S_{15}	8	89.6	2.7	16.1
2	6	C_{15}	14	91.0	4.6	15.4	1	6	C_{15}	8	89.6	2.7	15.1
2	6	C_{19}	14	92.7	4.6	2.7	1	6	C_{19}	8	91.4	2.8	2.3

Table 4. Statistics for Quality Measure II for filtered Images

d	σ	cell	t	CR	ER	#	d	σ	cell	t	CR	ER	#
2	2	S_{15}	10	97.4	1.2	12.5	1	2	S_{15}	9	97.3	1.3	11.2
2	2	C_{15}	10	97.4	1.2	12.2	1	2	C_{15}	9	97.3	1.3	12.2
2	2	C_{19}	8	97.6	1.1	1.1	1	2	C_{19}	8	97.6	1.3	1.1
2	4	S_{15}	11	95.2	2.0	12.7	1	4	S_{15}	10	95.1	2.2	11.5
2	4	C_{15}	11	95.2	2.0	12.4	1	4	C_{15}	11	95.3	2.3	12.8
2	4	C_{19}	11	95.5	2.1	1.0	1	4	C_{19}	8	95.3	2.0	1.1
2	6	S_{15}	12	91.9	4.3	12.8	1	6	S_{15}	10	87.7	5.3	10.6
2	6	C_{15}	14	92.4	4.6	12.7	1	6	C_{15}	10	92.2	3.3	12.4
2	6	C_{19}	16	93.0	4.9	1.0	1	6	C_{19}	9	83.7	8.2	1.0

3.3 Preprocessing

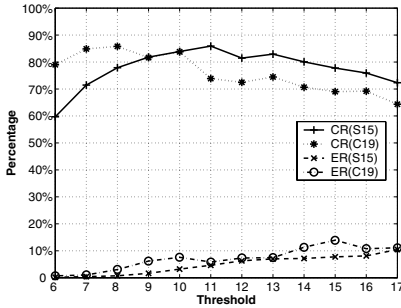
In practice it is important to apply before any segmentation a filter to smoothen the image within homogeneous regions and to enhance the contrast at bor-

ders. Very good such filters are the non-linear Harwood [9] or Kuwahara [10] filters. It is completely obvious how to generalize those known 2d filters to 3d. Tables 3 and 4 present some results for filtered images, where a single iteration of a 3d Kuwahara-Filter of size 3^3 is used. This considerably improves the results.

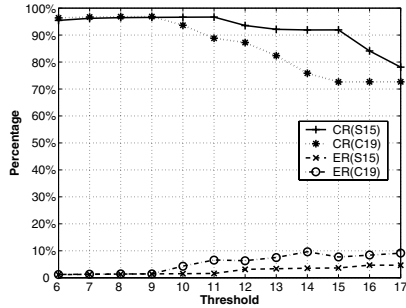
Table 4 is also visualized in Graph 1 and 2. The coverability rates (error rates) for the S_{15} and C_{19} segmentations with and without a Kuwahara-Filter are presented as upper (lower, resp.) bars in Graph 1 and the required number of segments is shown in Graph 2.

3.4 Statistics of all Thresholds

In tables 1 to 4 we have presented the coverability and error rates for the different CSC-segmentation variants S_{15} , C_{15} , C_{19} for some distinguished thresholds. Here we present the statistics of all used thresholds from 6 to 17 for all images in the classes (2, 2), (2, 4) and (1, 2). Graph 3 presents the coverability and error rates for the S_{15} - and C_{19} -segmentation variant (as C_{15} has very similar results to S_{15} we dropped it) without a filter and Graph 4 does the same with a Kuwahara-Filter.



Graph 3. Statistics for Quality Measure II (without filter)



Graph 4. Statistics for Quality Measure II for filtered Images (Kuwahara-Filter)

4 Discussion

We have compared CSC segmentations based on the rather elegant and compact S_{15} cell structure, on the C_{15} cell structure, and on the much more complicated C_{19} cell structure. There is no significant difference between a S_{15} - or C_{15} -CSC segmentation. A C_{19} -CSC segmentation gives the best results and requires the fewest segments for the unfiltered images with a reasonable small Gaussian noise ($\sigma < 6$). However, for the highly blurred images ($\sigma = 6$) all methods have problems but the C_{19} -CSC segmentation becomes worst. This is probably due

to the larger size of a C_{19} cell. The $(d, 6)$ -classes are certainly unfair: the mean difference in the gray values of the plexuses in the unblurred images is just 10, and two plexuses may be rather close. Adding a Gaussian noise with $\sigma = 6$ means that we change the gray value of 32% of all voxels by a value ≤ -6 or $\geq +6$, and even 4.5% are changed by a value ≤ -12 or $\geq +12$. A high deviation in a plexus asks for a large threshold - that, on the other hand, may also melt blurred parts of two different close plexuses.

Applying a single Kuwahara-Filter significantly improves all results. In most cases a single segment of a C_{19} -CSC segmentation suffices to describe a searched structure in the original image. It may be argued that it is unfair to 'unblur' those artificially blurred images. But also in all our tests with real MR or CT images an application of a Kuwahara-Filter improves the segmentation.

A S_{15} - or C_{15} -CSC segmentation however requires always a further aggregation of several segments to describe a searched structure, due to the inherent over-segmentation in both methods. Therefore, a classification problem remains to be solved for the S_{15} - and C_{15} -CSC: which segments have to be aggregated to get the unknown original structure? A solution will probably depend on the kind of images and structures one wants to segment, and requires a further research. The C_{19} -CSC seems to avoid this classification problem by computing only very few reasonable segments. Also, dynamic thresholds have to be added to the CSC. Similarity of segments should depend on properties of local histograms of the images.

References

- [1] P. Dokladal, R. Urtasun, I. Bloch, and L. Garnero. Segmentation of 3d head mr images using morphological reconstruction under constraints and automatic selection of markers. In *IEEE International Conference on Image Processing 2001*, pages 1075–1078, 2001. ICIP-2001, Thessalonique, Greece, Vol. III.
- [2] J. Sijbers, P. Scheunders, M. Verhoye, A. Van der Linden, D. Van Dyck, and E. Raman. Watershed-based segmentation of 3d mr data for volume quantization. *Magnetic Resonance Imaging*, 15(6):679–688, 1997.
- [3] C. Baillard, P. Hellier, and C. Barillot. Segmentation of brain 3d mr images using level sets and dense registration. *Med. Image Anal.*, 5(3):185–194, 2001.
- [4] Tobias Stammberger, S. Rudert, Markus Michaelis, Maximilian Reiser, and Karl-Hans Englmeier. Segmentation of mr images with b-spline snakes. a multi-resolution approach using the distance transformation for model forces. In *Bildverarbeitung für die Medizin*. Springer, Heidelberg, 1998.
- [5] Christian D. Werner, Frank B. Sachse, Karsten Mühlmann, and Olaf Dössel. Modellbasierte segmentation klinischer mr-aufnahmen. In *Bildverarbeitung für die Medizin*. Springer, Berlin, 1998.
- [6] L. Priese, V. Rehrmann. A fast hybrid color segmentation method. In S.J. Pöppel, H. Handels, editor, *Proc. DAGM Symposium Mustererkennung, Informatik Fachberichte, Springer 1993*, pages 297–304, 1993.
- [7] G. Hartmann. Recognition of hierarchically encoded images by technical and biological systems. *Biological Cybernetics*, 57:73–84, 1987.

- [8] P.Sturm, L.Priese. 3d-color-structure-code. A hierarchical region growing method for segmentation of 3d-images. In J. Bigun, T. Gustavson, editor, *Proc. SCIA 2003, LNCS 2749*, pages 603–608, 2003.
- [9] D. Harwood, M. Subbarao, H. Hakalahti, L. Davis. A new class of edge preserving smoothing filters. *Pattern Recognition Letters*, 2:155–162, 1987.
- [10] M. Kuwahara, K. Hachimura, S. Eiho, and M. Kinoshita. Processing of ri-angiocardigraphic images. In K. Preston and M. Onoe, editors, *Digital Processing of Biomedical Images*, pages 187–202, 1976.