JOURNAL OF MEDICAL INFORMATICS & TECHNOLOGIES Vol.3/2002, ISSN 1642-6037

image segmentation
hierarchical watershed transform
scale space

Miloš ŠRÁMEK*, Leonid I. DIMITROV*

SEGMENTATION OF TOMOGRAPHIC DATA BY HIERARCHICAL WATERSHED TRANSFORM

The aim of the proposed watershed based image segmentation technique is to split images into spatially homogeneous regions, which can be further processed by different image analysis tools. The advantage of such approach, in comparison to pixel oriented processing, is its lower sensitivity to superimposed noise due to averaging of regions properties over their area. The watershed segmentation technique is based on interpretation of an image as a topographic relief and on simulation of flow of water along steepest descent paths called downstreams. Thus, for each local minimum of the image, a drainage region is defined, which, if computed for a gradient image, represents an area with approximately constant properties. The segmentation technique is further extended for multi-scale image analysis by means of Gaussian smoothing. The aim of smoothing is to suppress image details that are smaller than standard deviation of the Gaussian. However, smoothing results not only in the desired increase of region size, but it also affects position of region boundaries, at least for larger standard deviations of the Gaussian filter. Therefore a new technique is proposed, based on region hierarchies, which enables to transfer region contours with precise position from the levels with low smoothing to levels with higher smoothing. Thus, segmentation of an image into large regions, but with exact contours, is obtained.

1. INTRODUCTION

In medical imaging, in order to visualize and/or quantify structures of interest, it is necessary to segment the data in meaningful regions. The segmentation task has already been addressed by many authors, and plentitude of techniques, both automatic and interactive, were designed [4,1,5,7]. In spite of this effort, the task is still far from an ultimate solution, mainly due to the unpredictable complexity of medical objects, the ever-increasing resolution of scanners and the high demands on segmentation precision, since the results are often used for diagnostic and treatment purposes.

The basic idea underlying the proposed region-based approach is that classification of approximately homogeneous regions is less sensitive to superimposed noise than pixel-based classification, due to the averaging of their properties over their area. The approach consists of several subsequent processing steps: (i) partition of the image in a set of basic homogeneous regions, (ii) region merging, based on some similarity criterion and, finally, (iii) region classification, either supervised (eventually interactive) or unsupervised. In this paper, we deal with

* Commision for Scientific Visualization (VISKOM), Austrian Academy of Sciences, Vienna, Austria
{milos.sramek | leonid.dimitrov}@oeaw.ac.at

the first two steps, which in our case stems from a scale-space based hierarchy of regions obtained by the watershed transform.

The paper is divided in four sections. In the Section 2, we introduce scale-space based watershed hierarchies, which enable to set the size of the basic regions according to demands (e.g., according to the dimension of the smallest details of interest). Section 3 deals with region merging and, finally, in Section 4, we give implementation details, results and formulate directions of our future research.

1.1. THE SCALE-SPACE CONCEPT AND THE WATERSHED TRANSFORM

Objects in the real world are perceived only over certain ranges of scale. A typical example of this fact is cartography: a map of the world depicts continents, big islands, rivers and maybe some of the major cities, while a city guide shows streets, buildings, parks and other details. From this point of view, the atlas can be understood as a multiscale symbolic representation of the world around us.

Transferring this concept into the area of image processing and vision, we could represent an image in the form of a sequence of smoothed images, showing gradually less details. Theory has shown, that Gaussian

$$G(x,y,\sigma) = \frac{1}{2\pi\sigma^2} \exp\left[-\frac{x^2+y^2}{2\sigma^2}\right] \tag{1}$$

is the smoothing operator[3], that should be used to generate this derived sequence, due to its linearity and spatial shift invariance as well as the fact that it introduces no new accidental structures (e.g. local extrema — at least in 1D). The level of smoothing can be characterized by the parameter σ, the standard deviation, approximately defining the dimensions of suppressed details (Figure 1).

The goal of low-level vision operations is to identify important image features, usually edges or regions, that can be used later for image interpretation. Therefore, the linear smoothing by Gaussian kernel should be combined with some other operator(s) in order to get more explicit descriptors of the scene geometry. The Canny's edge detector is a well known example of this approach[2].

Another possibility for extracting edges and regions from digital images, attracting the interest of the vision community in the recent years, is the concept of *watersheds*, adopted from topography[6]. From the point of view of this approach, grayscale images are considered to be topographic reliefs. A *catchment basin* $C(M)$ is defined around each local minimum M of an image, such that each of its points can be connected with the minimum M by a descending path, called *downstream*. Lines, separating the different catchment basins are called *watersheds* (Figure 2b,c).

Regions, identified in the segmentation process, should represent areas with some level of density homogeneity. Therefore, a gradient operator E is usually applied beforehand in order to enhance the inhomogeneities (edges) and the image is subjected to the watershed transform W (Figure 2a). Thus, taking into account the aforementioned scale-space concept, the segmentation by watersheds can be expressed as

$$S = W * E * G, \tag{2}$$

where the standard deviation σ of the smoothing kernel *G* defines the level of details and thus the size of the detected regions. Gaussian and gradient filters together can be replaced by the single magnitude of the gradient of the Gaussian (Gabor filter):

$$GA(x, y, \sigma) = \sqrt{\left[\frac{\partial G(x, y, \sigma)}{\partial x}\right]^2 + \left[\frac{\partial G(x, y, \sigma)}{\partial y}\right]^2} \tag{3}$$

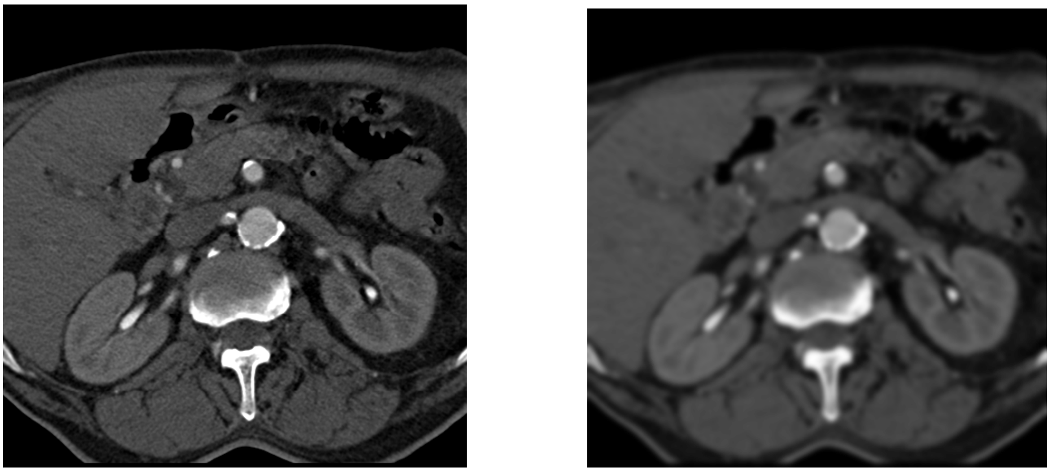

Fig.1. (a) Original image and (b)image smoothed by Gaussian with σ=3.0.

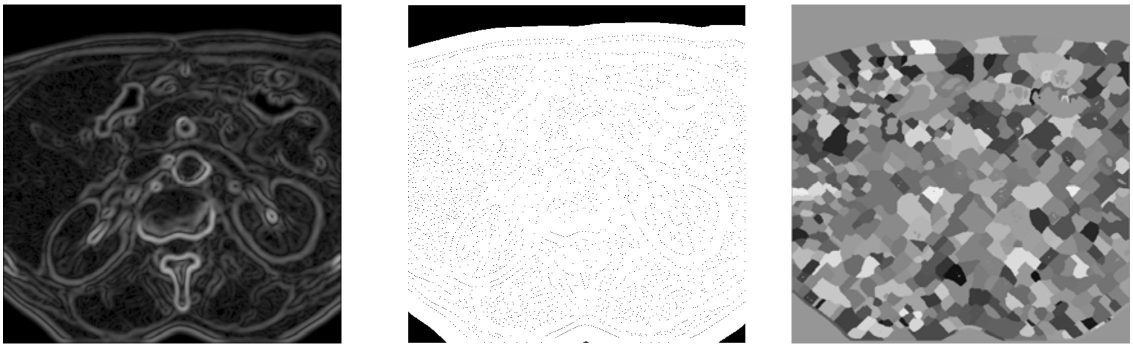

Fig.2. Computation of watersheds: (a)Edge detection by the Sobel operator (from Figure 1b), (b)local minima and (c)regions detected by the steepest descend technique.

1.2. WATERSHED HIERARCHIES

As we already mentioned in the previous section, smoothing by the Gaussian filter results in the desired larger regions. However, region contours are affected, too: they become also smoother and do not follow the edges exactly. To avoid this drawback, we propose a technique, based on watershed hierarchies, i.e. on a sequence of regions obtained by smoothing the image by different Gaussians with increasing σ.

Let I_σ be a segmented image smoothed by a Gaussian with kernel size σ and $S(\sigma_1,\sigma_n)$ is a sequence of n segmented images with $\sigma\in\{\sigma_1=\sigma_{min},\sigma_2,\ldots,\sigma_n=\sigma_{max}\}$. We tested two possibilities, for building such a sequence:

1. $\sigma_{i+1}=\sigma_i+\sigma_s$ (additive sequence) and
2. $\sigma_{i+1}=\sqrt{2}\sigma_i$ (multiplicative sequence).

Experiments showed that in the case of the multiplicative sequence the size of the region approximately doubles in one step, which indicates merging of neighboring pairs. Since we considered this property advantageous, only multiplicative sequences were used.

Images with small σ have precise contours but small mean region size, while images with large σ value have larger regions with imprecise contours. The idea of the watershed hierarchy based segmentation is to transfer the (precise) contours from the low levels to the larger regions at the higher levels of the hierarchy by means of *region overlapping*.

Let us consider two segmentations at levels σ_i and σ_{i+1}.

1. For each region j at level σ_i a region k at level σ_{i+1} exhibiting the largest number of common pixels is found. Since the region contours between two neighboring levels are shifted only insignificantly, there is usually one region k, which dominates in the number of common pixels.
2. A temporary image is created by assigning to each region j at the level σ_i the label of the corresponding region k of level σ_{i+1}. Since there are usually two or several regions at σ_i which overlap with the region k, a new region is defined, with size corresponding to k at level σ_{i+1} but with contours from level σ_i.
3. The segmentation at level σ_{i+1} is replaced by the temporary image. Thus, the precise contours from level σ_i are transferred to the level σ_{i+1}.

In order to build the region hierarchy $H(\sigma_1,\sigma_n)$ corresponding to the sequence $S(\sigma_1,\sigma_n)$, but with precise contours, we have to process all levels, sequentially starting from σ_1:

```
for(i=1;i<n;i=i+1)
    I_{i+1}=overlap(I_i,I_{i+1});
end for
```

It is necessary to point out that both i-th images in $S(\sigma_1,\sigma_n)$ and $H(\sigma_1,\sigma_n)$ have the same number of regions. Figure 3 shows an example of two segmentations of an image, with different smoothing and with and without region overlapping. We can see the perfect coincidence of contours on the overlapped image and shifted contours on the image without the overlap step.

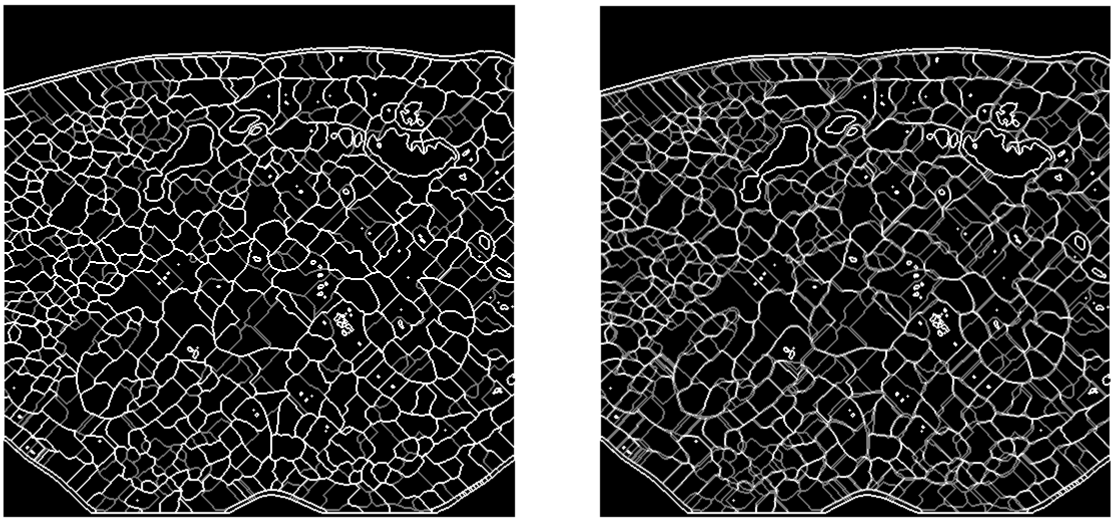

Fig.3. Coincidence of region contours smoothed with σ=2.8 and σ=4. Bright values represent pixels belonging to contours at both levels, darker only to one contour at one level.

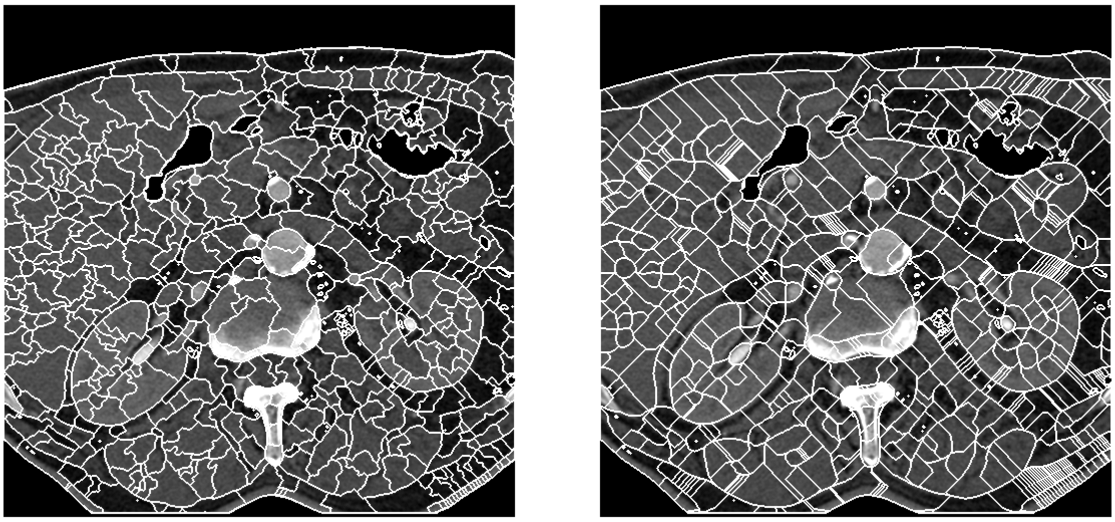

Fig.4. Region contours for σ=6.

We noticed that in segmented images, obtained by smoothing with larger σ (>6), series of narrow strip regions appear in long and narrow valleys, perpendicular to the valley axis. Their reason resides in the discrete nature of the processed image. If continuous, the smoothed image would have, of course, only one minimum at such a location. However, in our discrete case, due to the insufficient sampling rate given by the image resolution, a series of local minima are detected along the bottom of the valley, with very similar values resulting in the observed strip pattern. The problem could be probably solved by direct search of local minima by means of a gradient of the Gabor magnitude filter. This approach would be probably computationally very expensive and has not been implemented yet.

Since the strip artifacts do not appear at levels with small smoothing σ-s, and images at the same levels of both watershed hierarchy and segmentation sequence have the same number of regions, the strip-shaped regions disappear at the higher levels of the hierarchy and are replaced by regions with smaller size (Figure 4).

1.3. EFFICIENT COMPUTATION OF WATERSHED HIERARCHIES

The technique described above is computationally very expensive, since it is necessary to convolve the original image with Gaussian kernels of large size (e.g. 21×21). There are several possibilities how to speed up the computation:

1. The 2D (nD) Gaussian filter is separable, i.e. it can be expressed as a product of two components dependent only on the *x* and *y* coordinates respectively:

$$G(x,y,\sigma) = G_x(x,\sigma)G_y(y,\sigma) \tag{4}$$

$$G_x = \frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{x^2}{2\sigma^2}\right] \tag{5}$$

$$G_y = \frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{y^2}{2\sigma^2}\right] \tag{6}$$

   which allows to process rows of the image by the 1D filter G_x and subsequently columns by G_y. Thus the computational complexities reduce from $O(m^2n^2)$ to $O(2m^2n)$, where *m* is the dimension of the image and *n* is the filter kernel size.

2. In the case of a sequence of smoothed images, it is not necessary to process the original image at each level, but it is possible to use the smoothed image from the previous level. It can be shown (by means of multiplication of Fourier images of two Gaussians) that subsequent smoothing with σ_1 and σ_2 results into

$$\sigma = \sqrt{\sigma_1^2 + \sigma_2^2} \tag{7}$$

   So, if we have an image smoothed by σ_1 and we need smoothing by $\sigma = k\sigma_1$, we need to apply the filter with

$$\sigma_2 = \sqrt{k^2-1}\ \sigma_1. \tag{8}$$

   For $k = \sqrt{2}$ (our multiplicative sequence) we get $\sigma_2 = \sigma_1$.

2. REGION MERGING

One characteristic feature of the watershed transform is oversegmentation of the image into numerous small homogeneous regions. We showed in the previous section, how this number can be decreased by means of the watershed hierarchies. However, such regions still do not correspond to the homogeneous areas of the image since

- • the proposed segmentation technique, due to the rotational symmetry of the smoothing Gaussian kernel, results into approximately round regions, and
- • there is still significant oversegmentation in some areas due to the strip artifacts mentioned above.

Two neighboring regions *j* and *k* are merged if the distance *d* of their mean densities

$$d_x = m_j - m_k \tag{9}$$

is smaller than some user supplied threshold *t* (Figure 5).

The merging procedure is implemented using a list of regions, where each element stores all necessary information (i.e. region size, mean and maximum value for each band and list of neighbor identifiers) about the region. Thus the merging procedure can be implemented efficiently, since access to the image data is not necessary.

In order not to prefer any direction, candidates for merging are selected randomly from the list. For each such region its nearest neighbor is found and if their distance is smaller than the selected threshold *t*, both regions are merged. It means that one of the regions is deleted from the list and the properties of the remaining one are recomputed (e.g. mean or list of neighbors).

The regions are not merged directly with the threshold *t*, but the procedure is realized in *s* steps (e.g., *s* =10), with thresholds $t_i = \frac{it}{s}$ at the step *i*. The reason for this iterative approach is the noticed fact that, at the threshold *t*, some regions can be randomly merged with one of their neighbors, depending only on the order of testing. E.g., if a region *A* has two neighbors *B* and *C* with region means $m_C < m_A < m_B$ and distances $d_{AB} < d_{AC} < t$, it may happen that *A* is merged with *C*, which modifies the mean of the new region AC in such a way, that it is not merged with B, in spite of the fact that the nearest neighbor to *A* was originally *B*. If the iterative merging starts from the most similar regions, pairs ambiguous at the threshold *t* are correctly merged with their nearest neighbor.

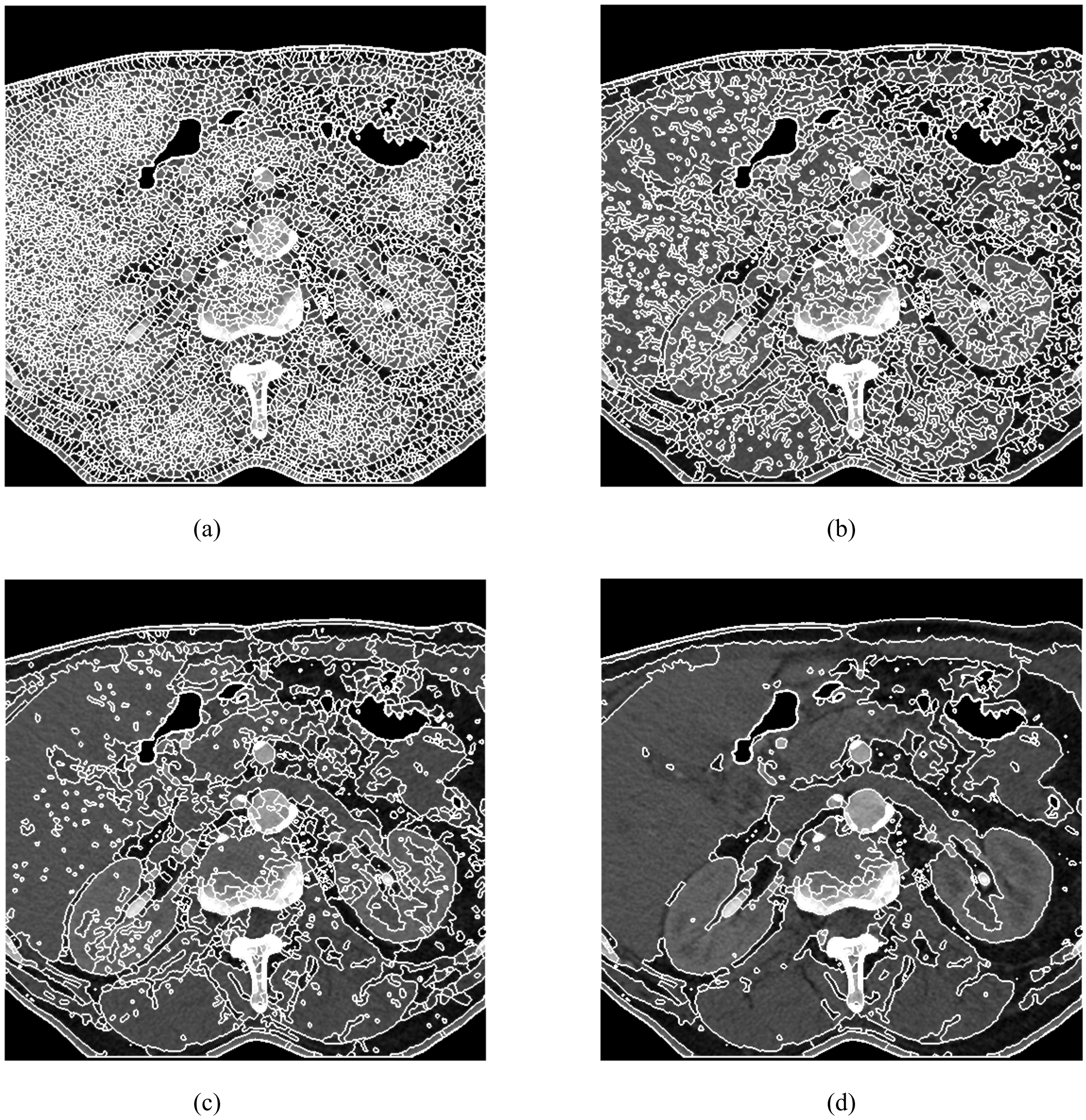

Fig.5: Region merging based on difference between region mean values: (a) Original segmentation with smoothing by σ=2.0, (b) merged regions with threshold 5 and (c) threshold 10 and (d) threshold 20.

3. IMPLEMENTATION, RESULTS AND FUTURE WORK

The proposed segmentation procedure was implemented in the C language in the framework of the Xite[1] image processing library and tools on a 1200MHz Athlon PC under the GNU/Linux OS. Building a 3 level hierarchy of a 512×512 image with subsequent region merging takes about 2 seconds.

[1] http://www.ifi.uio.no/~blab/Software/Xite/

In the near future, we would like to extend the technique to the processing of 3D data sets. We expect this extension to be more or less straightforward for the first step, building and processing of the watershed hierarchy. However, the region merging step, since it has to cope with 3D neighborhood relations, will be much more complicated. Finally, in order to identify the objects of interest, we intend to employ either knowledge based techniques (e.g., for similar, often repeated tasks) and/or interactive techniques with hardware supported volume visualization.

BIBLIOGRAPHY

[1] M. E. BRUMMER, R. M. MERSEREAU, R.L.EISNER, and R.J.R.LEWINE. Automatic detection of brain contours in {MRI} data sets. *IEEE Transactions on Medical Imaging*, 12(2):153–166, June 1993.

[2] J. CANNY. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8(6):679–698, November 1986.

[3] J. J. KOENDERINK. The structure of images. *Biolg. Cybern.*, 50:364–370, 1988.

[4] T. SCHIEMANN, U. TIEDE, M. BOMANS, and K. H. HÖHNE. Interactive 3D-segmentation. In R. A. Robb, editor, *Visualization in Biomedical Computing 1992, Proc. SPIE 1808*, pages 376--383. SPIE, Chapel Hill, NC, 1992.

[5] M. ŠRÁMEK. Interactive segmentation of tissues for medical imaging. In Václav Hlaváč and Tomáš Pajdla, editors, *Czech Pattern Recognition Workshop '93*, pages 164–171, Temešvár u Písku, Czech Republic, November 4^{th}-6^{th}, 1993.

[6] L. VINCENT and P. SOILLE. Watersheds in digital spaces: An efficient algorithm based on immersion simulation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(6):583–598, June 1991.

[7] X. ZENG, L. H. STAIB, R. T. SCHULTZ, and J. S. DUNCAN. Segmentation and measurement of the cortex from 3{D} {MR} using coupled surfaces propagation. *IEEE Trans. Med. Imaging*, 18(10), 1999.