Multi-Scale Gradient Magnitude Watershed Segmentation

Ole Fogh Olsen[1] and Mads Nielsen[2]

[1] DIKU, University of Copenhagen, Universitetsparken 1, DK-2100, Copenhagen E, Denmark, fogh@diku.dk
[2] 3D-Lab, School of Dentistry, University of Copenhagen, malte@lab3d.odont.ku.dk

Abstract. A partitioning of an nD image is defined as the watersheds of some locally computable inhomogeneity measure. Dependent on the scale of the inhomogeneity measure a coarse or fine partitioning is defined. By analysis of the structural changes (catastrophes) in the measure introduced when scale is increased, a multi-scale linking of segments can be defined. This paper describes the multi-scale linking based on recent results of the deep structure of the squared gradient field[1]. An interactive semi-automatic segmentation tool, and results on synthetic and real 3D medical images are presented.

1 Introduction

The goal of an image segmentation is a description of the shape of some image structure of predefined semantics. However, addressing the shape of an object is not simple since the shape is not intrinsically defined[2]; *shape is defined through an interpretation of measurements.* This introduces the measurements apparatus and its intrinsic resolution as an important part of a shape definition. This is well-known from the definition of coast-lines.

In this paper, we use a Gaussian probe as a linear measurement apparatus (i.e. Gaussian convolution) and thereby introduce the Gaussian scale-space formalism [3, 4]. We base the shape definition on the local scale-space n-jet. In general, the definition of shapes cannot be based solely on local information; global information may constrain local decisions. Following this line, segmentations have been defined as the minimum of a energy functional [5, 6]. This is computationally expensive and difficult to tune to prior information unless extensive statistical material is available [7, 8]. Also split and merge techniques[9] have been introduced. However, this strategy is captured more elegantly in multi-scale linking approaches [4, 10, 11, 12, 13].

Locally we compute a measure of dissimilarity of the image, at a certain scale. The watersheds of this measure defines the segmentation. Watershed cannot be identified locally, i.e. they capture global properties of the image.

A segment boundary is defined as a watershed of a dissimilarity measure in turn defined using a certain width (scale) of the Gaussian aperture function. When varying the scale parameter, the watersheds deform continuously until a transition point where a watershed appears/disappears. Analysis of such transitions in the multi-scale structure has been carried out for a number of local

image functionals (i.e. feature detectors) which may be used as dissimilarity measure. Damon established the catastrophe theory for diffused images [14] and also analysed ridge measures [15], Lindeberg analysed blob detectors [16], and Rieger analysed edge and corner detectors [17]. We analysed the gradient magnitude[1].

Multi-scale watershed segmentation has been carried out based on the intensity and ridges: Gauch[12] used the image intensity function directly as local measure of homogeneity. Eberly [13] defined a homogeneity measure based upon local "ridgeness". Griffin [18] used the image intensity or the image intensity gradient and based the segmentation on a multi-grid method. We use the recent results on the multi-scale structure of the gradient magnitude [1] to establish the multi-scale linking. The watersheds in the gradient magnitude intuitively partition the image where the gradient is large.

An object is defined through a root segment and its linking to a localization scale. To interactively select roots and scales, an interactive tool (serving same task as Pizer et al.'s [11]) has been constructed. Since the multi-scale structure can be pre-computed and hashed, interaction is fast.

The following section defines the scale space and the local dissimilarity measures. Then, watersheds, catchment basins, segments and multi-scale linking are defined in Section 3. Section 4 describes the interactive segmentation tool. Section 5 presents experimental results. Finally, in Section 6, we summarise.

2 Scale-space and local dissimilarity measures

Definition 1 Scale-space. The scale-space $L(\cdot, t)$ is generated from an image $I(\cdot) \equiv L(\cdot, 0)$ by Gaussian blurring $L(x,t) \equiv \int I(x')g(x-x',t)dx'$ where $g(\cdot,t)$ is a Gaussian and $t = \sigma^2/2$ it's spread.

Derivatives of the scale-space can be obtained robustly by differentiation of the Gaussian prior to convolution.

For images where segments are assumed to have homogeneous intensity we use the gradient magnitude $|\nabla L|^2 = L_x^2 + L_y^2 + L_z^2$ as dissimilarity measure. In images where only the texture differs from segment to segment a local texture measure is used: the local frequency contents of an image can be measured with a Fourier transform under a Gaussian window function:

$$\hat{L}(x,k,t) = \int I(x')g(x-x',t)e^{-ikx'}dx'$$

where k is the wave-vector and the total filter is an oriented Gabor function. When spatially differentiating this the local phase shift is taken into account so that $\tilde{\partial}_x \hat{L}(x,k,t) \equiv (\partial_x - ik \cdot e_x)\hat{L}(x,k,t)$ is assumed to be small in regions of same texture. e_x denotes a unit vector in the x direction. We define a local dissimilarity measure for texture segmentation (K is a subset of frequencies chosen to discriminate textures) as

$$m(\cdot,t) = \sum_{k \in K, x_i \in \{x,y,z\}} |\tilde{\partial}_{x_i} \hat{L}(\cdot,k,t)|^2 \quad . \tag{1}$$

3 Segments and linking

This section defines segments based on watersheds of an arbitrary local dissimilarity measure. The notion of *watersheds* and *catchment basins* arises when a function is viewed as a topographic relief with height identified with the image intensity. The watersheds are the boundaries between areas, the so-called catchment basins, which drain to one local minimum.

Definition 2 Catchment basin. A catchment basin of a local minimum is the inner points of the closure of the union of all steepest descent lines ending in the minimum.

Definition 3 Watersheds. The watersheds are the boundaries of the catchment basins.

A property especially interesting for segmentation is the fact that *watersheds form closed hyper-surfaces for Morse functions.* Hence, the watersheds of a function give a full partitioning of the multi-dimensional image domain; there is no need for closing or connecting edges to get a partition. This partition has a very flexible topology. As an example in 2D, any number of segments may generically meet in a point. On the contrary, a partition based upon zero-crossings of a feature detector will generically only exhibit 2- and 4-junctions.

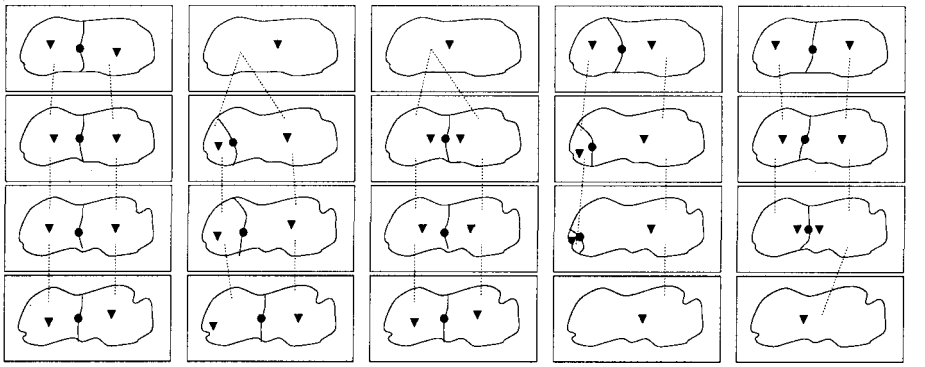
Each catchment basin contains exactly one local minimum, the seed of the basin. Instead of directly analysing the multi-scale structure of the watersheds, we can analyse the dual : the local minima. This makes analysis feasible in terms of catastrophes[19]. We suggest:

Definition 4 Segment. A segment is the catchment basin for a local minimum of a dissimilarity measure.

Often the image structure is probed on a much finer scale than the scale of the structures of interest, giving rise to over-segmentation. A common solution [20] is to "flood" the image. Maes et al.[21] post-processed the segmentation by merging neighbouring regions using a MDL Principle. Griffin et al. [18] simplified the image stepwise by treating districts (bounded by maximum gradient paths) as one point and recalculating the slopelines. We suggest to detect objects at coarse scale and localise them at finer scale. In order to do this, the structures must be linked across scale.

Single scale watershed segmentation on the gradient is well known [22, 23]. The singularities of the gradient magnitude and with them the seeds of segments occur in the critical points of the image but also in the points where the second order structure of the image vanishes in one direction. These points evolve when scale is changed, and at certain catastrophe points in scale-space, they interact: appear or annihilates.

The only generic events in scale-space of the gradient magnitude image is a fold catastrophe and a cusp catastrophe involving a minimum[1]. The duality between segments and the minima of the gradient magnitude suggests the linking



(a) No change (b) Annihilation (c) Merging (d) Creation (e) Splitting

Fig. 1. Multi-scale linking of generic events in watersheds of the gradient magnitude. The events (annihilation, merging, creation, splitting) are named after the interaction between the saddle and the minimum (or minima). Minima and saddles are symbolised with triangles and circles, respectively. A line from a segment to a segment indicates the linking.

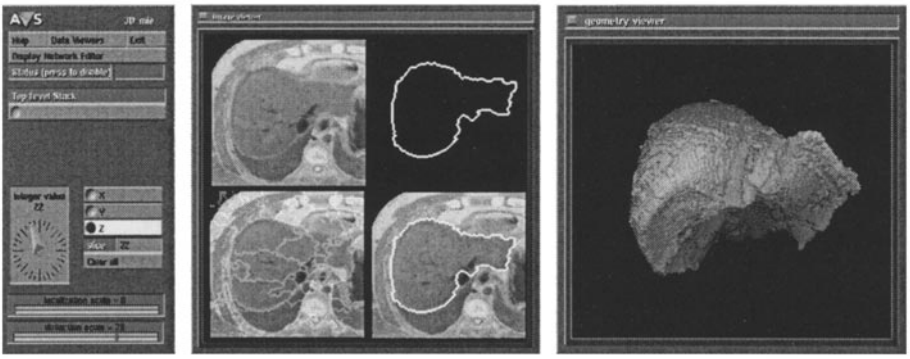
scheme. A cusp is the interaction between three singularities, in the present case two minima and a saddle. The two minima and the saddle either meet and become one minimum or the reverse event. A fold is the interaction between two singularities, in the present case one minimum and a saddle. The two singularities meet and annihilate or the reverse creation event.

Figure 1 illustrates the idea in 2D with scale increasing upwards. In the cases of annihilation (b) and merging (c) two minima and a saddle are reduced to one minimum, corresponding to the disappearing of a border between the two segments. The cases of splitting (d) and creation (e) are the reverse events where the emerging saddle corresponds to the appearing of a border between the segments (dual to the two minima). Hence, the linking is in all cases given by the saddle connecting the involved minima.

The implementation of the linking exploits the fact that image structure changes smoothly with scale, therefore a spatial maximum correlation between segments at neighbouring scales can be used as linking criterion. Lindeberg [16] used a similar idea for blob linking.

4 The interactive segmentation interface

A user-interface has been constructed for accessing the multi-scale segment structure (Figure 2). Raising the detection scale gives generally fewer segments and vice versa. Raising the localisation scale results in more smooth boundaries and vice versa. The user gets interactive 3D feedback on the selections limited in speed mainly by the computers rendering capabilities.



(a) (b) (c)

Fig. 2. User interface windows. In window (*a*), the localization and detection scale is selected as well as a slice in one of the three Cartesian directions. This gives a partition of the domain. Window (b) displays the image slice (top left), the partition superimposed on the image slice (bottom left), the union of the selected segments (top right) and the selected segments superimposed on the image slice (bottom right). The object is defined by selecting/deselecting volume segments in one of images in window (*b*). The third window (*c*) continuously renders the union of the selected volume segments.

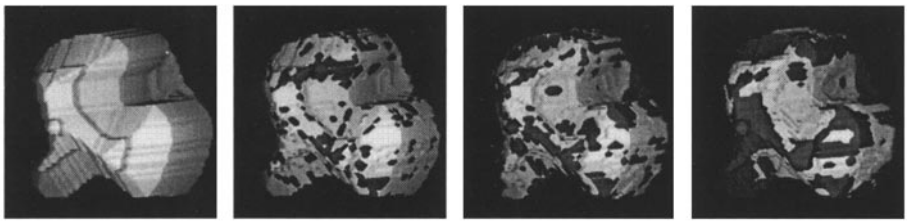
5 Results and verification

This section presents results on three types of images using the gradient magnitude squared as the dissimilarity measure. The images are a software simulated liver phantom (Figure 3), a CT head scan of a patient with abnormal growth (Subfigures 5.c,d) and digital photos (red channel) from the visible human project (Subfigures 5.a,b). The tasks are to segment the phantom, jaw muscles and the liver, respectively. Furthermore results of texture segmentation on a toy image is shown in Figure 6.

Figure 3 (a) displays a rendering of the true phantom. Different levels of noise has been added: respectively 0, 25, 50 and 75 percent of the voxels have been modified with Gaussian additive noise with zero mean and standard deviation of 80% of the phantom to background contrast. We shall refer to the different noise level as the 0, 25, 50 and 75 percent case. Segmentation has been performed using an appropriate high detection scale in order to define the object as one single segment, which has automatically been tracked to a lower localisation scale (see Figure 3).

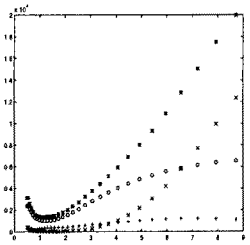
Segmentation has been performed using an appropriate high detection scale in order to define the object as one single segment, which has automatically been tracked to a lower localisation scale (see figure 3).

The errors in localisation of the boundary are due to two different sources. The noise pixels influences the multi-scale linking so that a noisy boundary is created at low scales. By increasing the localization scale, a smooth boundary can be constructed. However, at this scale, the Gaussian blurring has deformed the



(a) (b) (c) (d)

Fig. 3. The true object is presented in (a) as a bright surface rendering. In (b), (c) and (d) is a bright surface rendering of the segmentation for noise level 25, 50 and 75, resp. . The true object (a) is for comparison superimposed as the dark surface in (b), (c), (d). The phantom consists of 57708 voxels in a 64^3 volume.



(a)

Noise level:	0 %	25 %	50 %	75 %
Best localization scale (σ)	0.605	1.18	1.73	2.78
Number of wrong voxels	79	1342	2830	7121
Wrong voxels / size of phantom	0.0014	0.0233	0.0490	0.1234

(b)

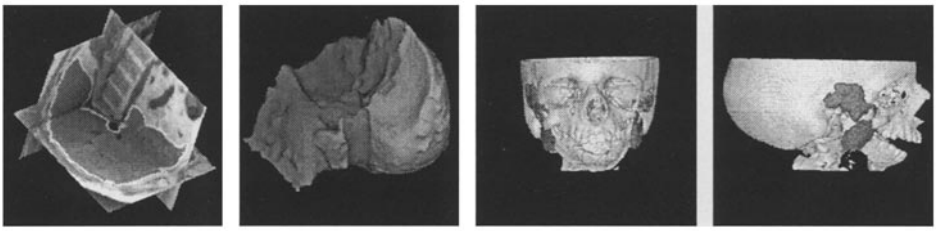
Fig. 4. Number of erroneous voxels as a function of localization scale for noise levels 25. Bold crosses indicate total number of voxels, circles indicate missing voxels on surface, crosses indicate missing interior voxels, and pluses indicate additional voxels. The qualitative shape of the curves are similar for the other noise levels. The statistics of phantom segmentation is summarised in (b)

object deterministically so that sharp corners (convex or concave) are rounded. An optimal scale may be established from prior information on noise level, object size, etc. This in done empirically in Figure 4. In Figure 3 (d) the phantom is generally exposed in convex patches while the segmentation is exposed in concave patches due to the deterministic shape distortion at higher scales.

6 Summary

A framework for multi–scale segmentation has been presented. The partition by the watersheds of the gradient magnitude has been analysed and implemented in the case of Gaussian scale space. The multi-scale linking has been defined on the basis of results from catastrophe theory.

The selection mechanism of interactively picking objects at an appropriate scale and combining the result at localization scale provides a fast way of doing semi–automatic segmentation.



(a) (b) (c) (d)

Fig. 5. Liver segmentation (a) and (b) from a cube of size 128x128x112 voxels each 1.758 mm^3. In (a), the liver boundary is superimposed on three orthogonal slices of the subject cube. The same liver segment is visualised in (b) as a surface rendering. The view from the spine (b) clearly reveals the imprint from other internal organs. The segmentation is difficult for mainly two reasons: The high similarity between liver tissue and the neighbouring muscle tissue; and the inhomogeneity of the liver tissue itself. **Jaw muscles**. The segmentation was verified by Professor Sven Kreiborg. The subject is a 256x256x64 cube of size 1x1x2 mm^3. The muscular structures are located next to bone (high value), skin (low value) and salivary glands (approximately same value) which makes the task difficult for standard techniques. A fine detection scale must be used because the muscles are flat structures, that is fine scale structure in one direction. The coarse structures of a muscle was selected with a few (< 5) mouse clicks using a coarse detection scale ($\sigma \approx 3.06$ pixels) , and the segmentation was then refined with a few (< 10) clicks using a finer detection scale ($\sigma \approx 0.805$ pixels). Localisation scale is 0.5 pixels.

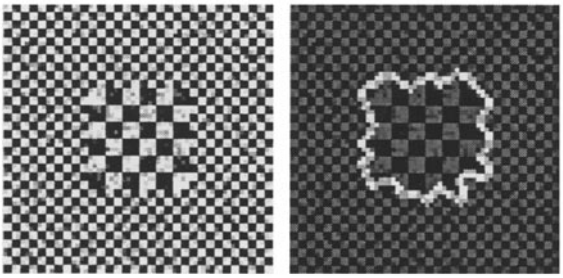


Fig. 6. Multi-scale texture segmentation based on local frequency differences defined by Gabor functions. Two distinctive textures with Gaussian noise added is segmented using the dissimilarity measure defined in Eq. 1. These are preliminary results serving only as indication of the generality of the multi-scale watershed segmentation approach.

The definition of segments can be changed by using another measure of dissimilarity instead of the gradient magnitude. This is possible within the same general framework although different structural changes might occur generically for other measures and diffusion schemes.

Acknowledgements We thank S. Kreiborg, P. Larsen, and A. B. Dobrzeniecki, 3D-Lab, School of dentistry, University of Copenhagen, for providing the CT data and phantom image and supervising the interactive segmentation.

References

1. O.F.Olsen and M.Nielsen, "Generic events for the gradient squared with application to multi-scale segmentation." SS97, Utrecth, 1997.
2. J. J. Koenderink, *Solid Shape.* Cambridge, Mass.: MIT Press, 1990.
3. A. P. Witkin, "Scale space filtering," in *Proc. International Joint Conference on Artificial Intelligence*, (Karlsruhe, Germany), pp. 1019–1023, 1983.
4. J. J. Koenderink, "The structure of images," *Biol. Cybern.*, vol. 50, pp. 363–370, 1984.
5. D. Mumford and J. Shah, "Boundary detection by minimizing functionals," in *Proc. IEEE Conf. on CVPR*, (San Francisco), 1985.
6. Y. G. Leclerc, "Constructing simple stable descriptions for image partitioning," *IJCV*, vol. 3, pp. 73–102, 1989.
7. U. Grenander, Y. Chow, and D. Keenan, *Hands. A Pattern Theoretic Study of Biological Shapes.* Springer Verlag, 1991.
8. T. Cootes, J. Taylor, D. Cooper, and J. Graham, "Active shape models–their training and application," *CVIU*, vol. 61, pp. 38–59, January 1995.
9. R. C. Gonzales and R. E. Woods, *Digital Image Processing.* Addison Wesley, 1993.
10. K. L. Vincken, C. N. de Graaf, A. S. E. Koster, M. A. Viergever, F. J. R. Appelman, and G. R. Timmens, "Multiresolution segmentation of 3D images by the hyperstack," in *VBC*, pp. 115–122, Los Alamitos, CA: IEEE CS Press, 1990.
11. L. M. Lifshitz and S. M. Pizer, "A multiresolution hierarchical approach to image segmentation based on intensity extrema," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 12, no. 6, pp. 529–541, 1990.
12. J. M. Gauch, W. R. Oliver, and S. M. Pizer, "Multiresolution shape descriptions and their applications in medical imaging," in *IPMI 10*, 1988.
13. D. Eberly and S. M. Pizer, "Ridge flow models for image segmentation," Tech. Rep. TR93-056, University of North Carolina, Dept. of Computer Science, 1993.
14. J. Damon, "Local morse theory for solutions to the heat equation and gaussian blurring," *Journal of Differential Equations*, vol. 115, January 1995.
15. J. N. Damon, "Properties of ridges and cores for twodimensional images." Unpub.
16. T. Lindeberg, *Scale–Space Theory in Computer Vision.* Kluwer Academic Publishers, 1994. ISBN 0-7923-9418-6.
17. J. H. Rieger, "Generic evolutions of edges on families of diffused greyvalue surfaces," *JMIV*, vol. 5, pp. 207–217, 1995.
18. L. D. Griffin, A. C. F. Colchester, and G. P. Robinson, "Scale and segmentation of grey-level images using maximum gradient paths," *Image and Vision Computing*, vol. 10, pp. 389–402, July/August 1992.
19. R. Gilmore, *Catastrophe Theory for Scientists and Engineers.* Dover, 1981. ISBN 0-486-67539-4.
20. L. Najman and M. Schmitt, "Watershed of a continuos function," *Signal Processing*, vol. 38, pp. 99–112, July 1994.
21. F. Maes, D. Vandermeulen, P. Suetens, and G. Marchal, "Computer–aided interactive object delineation using an intelligent paintbrush technique," in *CVRMed95* (N. Ayache, ed.), pp. 77–83, Springer–Verlag, 1995. Lecture Notes 905.
22. F. Meyer, "Topographic distance and watershed lines," *Signal Processing*, vol. 38, pp. 99–112, July 1994.
23. L. D. Griffin, *Descriptions of Image Structure.* PhD thesis, Uni. of London, 1995.