The Scale Structure of the Gradient Magnitude

This report is Deliverable 2 in the EU project IST-2001-35433

Ole Fogh Olsen

IT University Technical Report Series **TR-2003-29**

ISSN 1600–6100 **March 2003**

ISSN 1600–6100

ISBN 87–7949–039–5

Copies may be obtained by contacting:

IT University of Copenhagen
Glentevej 67
DK-2400 Copenhagen NV
Denmark

Telephone:	+45 38 16 88 88
Telefax:	+45 38 16 88 99
Web	`www.it-c.dk`

Contents

1 Abstract

This paper will describe the multi-scale structure of the singularities in gradient magnitude images including data structures for representation and algorithms for extraction. Especially the case of the minima of the gradient magnittude is considered as they constitute a dual representation of the gradient magnitude watersheds proven efficient for multi dimensional image segmentation.

2 Introduction

In this section we will outline the organisation of this paper. We will not give a thourough description of scale-space theory, differential geometry in image analysis and singularity theory. On the contrary we will assume the reader to be familiar with these subjects and otherwise refer the reader to the following literature as starting point ([21, 22, 20, 19, 13, 14, 26, 5, 3, 2])

First we will present the duality and relation between catchment basins and watersheds and minima of a landscape. Secondly watershed segmentation on the gradient magnitude image. Thirdly an intermezzo on maxima of the gradient magnitude is presented. Next step is to study the change of the singularities with changing scales. This analysis gives basis for discussing possible linking of the singularities across scales and we end the report by describing implementation of the linking and give illustrations of this on segmentation.

3 Duality between minima and watersheds

The idea of partitioning a landscape into hills and dales dates back at least as early as Maxwell [16]. There have been differences in the definition of watershed and watercourses [8, 25] although it seems the controversy depends on nongeneric cases (see Rieger [24] for a discussion).

Flowlines are the integrated curves along the gradient field. Per definition these flow lines start and end at singularities. They are the complement of isophotes. The common flow lines run from a maximum and end at minimum. The uncommon ones run from a maximum to a saddle or from a saddle to a minimum. The union of the points on all flowlines ending at the same minimum together with the minimum point is called a catchment basin.

In special but generic cases saddles, maxima and the flow line connecting them can be part of catchment basin. If the two downwards flow lines from a saddle end at the same minimum then the saddle point plus the two upwards flow lines coming from maxima are part of the catchment basin belonging to that particular minimum. If a maximum only has flowlines to saddle points and to the same minimum then the maximum point belongs to the catchment basin for the minimum. This is illustrated in figure 1 There is exactly one watershed following the circular ridge. There is *not* a watershed from the peak to the ridge structure since the water falling one either side of the peak end up in the same minimum. Hence, even though the water is parted by the line, it is not a watershed since two water-drops falling one on each side end up at the same minimum. This is a clear example of why watersheds can not be determined by local structure. The global topology is needed.

The saddles, maxima and the flow lines between them, which do not fall into the special case described in the previous paragraph, constitute the border between catchment basins and are referred to as watersheds.

Only certain configurations of separatrix and singularities are stable [17]. The stable ones are shown in Figure 2. One can build the whole graph of separatrices by combining these configurations.

A interesting point in watershed detection on general surfaces is that plataus are not generic. Hence for a sufficiently fine quantification in the intensity domain one can speed up the watershed detection with this knowledge. Of course if one works with for instance artificial images with constant areas then it is necessary to use implementation like the one presented by Vincent et al.[27].

3.1 Watersheds and catchment basins in 3D

The families of slope lines in 3D that run into a single extremum are separated by the three-dimensional analogous to watersheds which are singular surfaces. The surfaces divide space

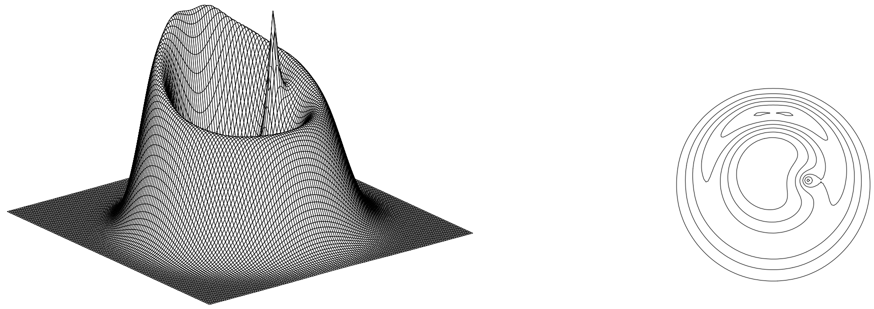


Figure 1: A two dimensional landscape. Only the ring structure produces a watershed, not the peak. Contours of the mesh plot.

into distinct cells labelled by the extremum. Koenderink [10, page 516] touch thesestructures briefly. Formally we can define watersheds for scalar function over a three dimensional domain exactly as in the 2D case.

Let us have a closer look at $P(f)$ the critical points above more than one minimum and at $S(f)$ the slope lines connecting the critical points in $P(f)$. The points in $P(f)$ are the maxima and hyper-saddles connected to two distinct minima. These points have in one of the three principal curvature directions slopelines descending to the distinct minima; the two slopelines run in opposite directions along the principal curvature direction. These points make the anchor points for a watershed surface defined by these points and the slope lines $S(f)$ connecting them. Let us now restrict ourselves to this surface manifold and consider the 3D scalar function restricted to this manifold. The critical points of this function are the points in $P(f)$. The maxima in 3D are maxima here as well. The two kinds of hyper-saddles are respectively saddles (C_{+--}) and minima (C_{++-}) on the manifold. The surface is spanned by slopelines connecting the critical points in $P(f)$. The common slopelines run between maxima C_{---} and minima on the manifold (hypersaddles of the type C_{++-}). These common slopelines may be regarded as the faces of cells. The faces are separated by the un-common slopelines from $S(f)$ running from maxima C_{---} to hypersaddles C_{--+} and from hypersaddles C_{-++} to minima C_{+++}. If the scale function on the two dimensional manifold is considered alone then the watersheds for this 2D function will be the latter described

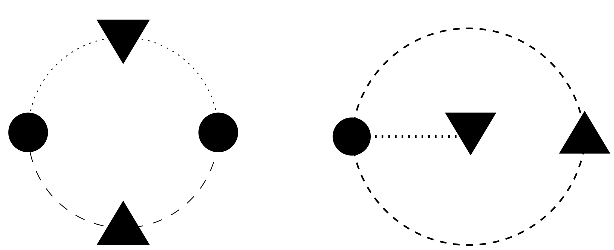


Figure 2: Stable configurations of separatrices and singularities.

uncommon slopelines in $S(f)$.

3.2 summary

An important point of this section is the one to one correspondence between minima and catchment basins. This duality allows transfering of results , for instnace the dynamic over scale of catchment basins can be deduced from the dynamic of the minima.

A catchment basin is the region which drains to exactly one minimum. The watersheds form the boundaries between catchment basins. Watersheds are a subset of the separatrices for the considered function and form closed curves in 2D (surfaces in 3D, etc. in nD). Catchment basins can be formally defined in arbitrary finite dimension by construction a partial ordering of the critical points. The partial ordering is based on slopelines (lines of steepest descent). Watersheds can not be locally detected, and consequently capture global topology.

4 Watershed segmentation

The notions of watersheds and catchment basins [1] come from the analysis of landscapes. The image is viewed as a landscape where the intensity is the elevation and the terminology is transferred from the topographical landscape to the image. Catchment basins are areas draining to the same minimum. The watersheds are the lines which separate different catchment basins. A well known watershed is the great dividing range in the US diving the Pacific and Atlantic. An 1D example is shown in Figure 3.

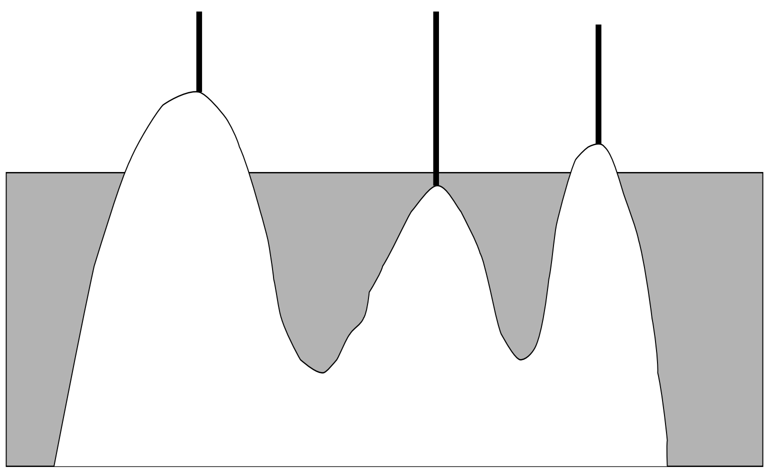


Figure 3: Flooding of landscape. Black borders correspond to watersheds

In the one dimensional case the watersheds are simply all the maxima of the image and the catchment basins are the regions in between, see Figure 3. It gets a bit more complicated in higher dimensions since the water can run separate ways and still end up in the same lake.

In Figure 4 is presented a checker board image with noise at scale $\sigma = 1.43$ pixels. The surface plot and the watersheds of this function is presented. The watersheds separate each

[1] Catchment basins are also known as regions of influence

of the minima in the function and run from maximum to saddle to maximum, etc. The border of the image is treated as an infinitely high. This example is intended to demonstrate what watersheds are, it does not illustrate the segmentation technique itself.

The watersheds of the image itself do not in general provide a good segmentation. This is clearly seen in Figure 4. What gives a good segmentation are watersheds of the *gradient magnitude*, a non-linear differential of the image. An example of this method on the checker board in Figure 4 is given in Figure 5. The saddle regions in the image L give rise to regions but the extend of these regions are for the presented scale below the resolution of the segmentation. The watersheds of the image (Figure 4) and of the gradient squared (Figure 5) are presented for the coarse scale structures.

The use of the global structure make it possible to determine wether or not an edge separate two distinct regions. The use of global topology also helps when detecting edge junctions. The watersheds are basicly a limiting set of the slopelines, where two limiting sets meet there is a junction. In edge detection with zero–crossings, the edges so to speak fight for the right to the junction and the strongest wins. The need for the global information is on the other hand a disadvantage for the method simply because the feature is not locally detectable. This means that critical points are needed as an anchor points when constructing the watersheds.

Hence the watersheds are not locally detectable. Watersheds are only detectable if the global topology is probed. This latter property is one of the reasons that watersheds of the gradient magnitude handle junctions fairly gracefully. Another important property especially interesting for segmentation is the fact that: Watersheds form closed curves for Morse functions in 2D. In 3D watersheds form closed surfaces.

Finding the watersheds of a function gives a full partitioning of the domain; there is no need for closing or connecting edges to get a partition. The closing of edges in a consistent way is one of the problems which has to be solved when segmenting based on edge detection [1].

4.1 Watersheds for the gradient magnitude

The idea of segmenting by watersheds of the gradient magnitude image is well known from the field of morphological segmentation [7]. There is a segment for each minimum in the gradient magnitude. The gradient magnitude is not differentiable in critical points of the image. To ease the analysis we will studied the square of the gradient magnitude which of course has the same extrema and saddles the same place as the gradient magnitude. All analyses and all experiments are carried out on the gradient magnitude squared which has the same steepest descent lines and thereby the the same watersheds as the gradient magnitude itself.

The singularities (and hence the minima) of the gradient squared come in two flavours formally stated in 2D in the Equations 1 and 2:

$$L_x = 0 \wedge L_y = 0 \tag{1}$$

$$L_x^2 + L_y^2 \neq 0 \wedge L_{ww} = 0 \wedge L_{wv} = 0. \tag{2}$$

Where the gauge coordinate system (w,v) is used, w is in the direction of the gradient, v is perpendicular to w. The latter conditions imply that the slope line curvature $\mu = \frac{-L_{wv}}{L_w}$ and the relative variation of the scalar field L_w along slope lines $\delta = \frac{-L_{ww}}{L_w}$ must both equal zero. In other words either there is a singularity, or the slope lines are locally straight and the isophotes are locally equally spaced along the slope lines.

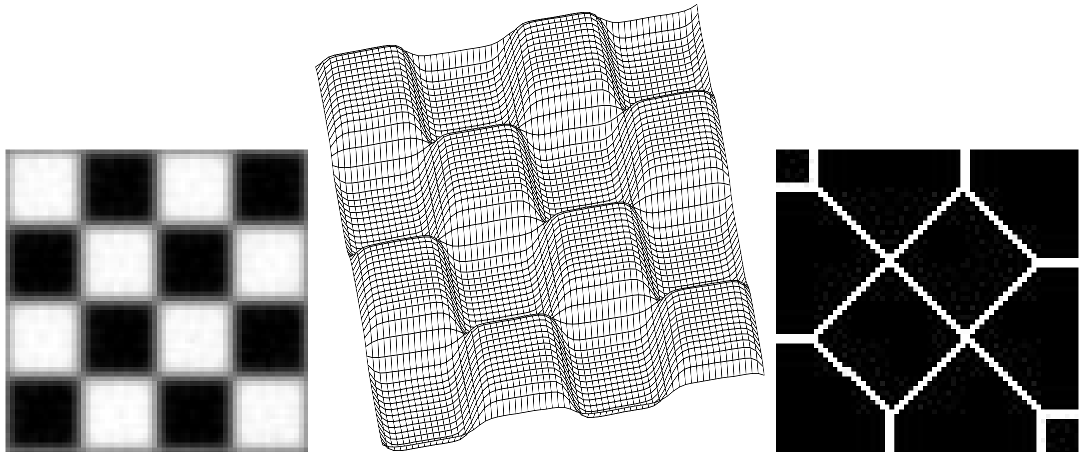


Figure 4: Checker board with noise at scale 1.43 pixel. Secondly a mesh plot. Thirdly Watersheds of the image is presented. This example is intended to demonstrate what watersheds are, it doesnot illustrate the segmentation technique itself.

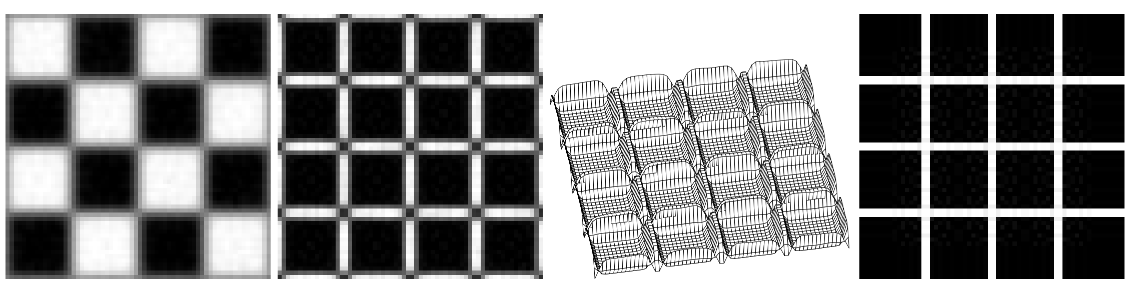


Figure 5: Checker board with noise. The gradient squared of the checker board and the surface plot of the gradient squared . Last the watersheds of the gradient squared is shown, this subset of the separatrices yields in general a good segmentation.

4.2 Singularities for the gradient squared in 3D

The singularities of the gradient squared and with them the basis of segments occur in the critical points of the image (see Equation 3) and in the points where the second order structure of the image vanishes in one direction (see Equation 4), just as in 2D:

$$L_x^2 + L_y^2 + L_z^2 = 0 \tag{3}$$

$$L_x^2 + L_y^2 + L_z^2 \neq 0 \wedge L_{ww} = 0 \wedge L_{wv} = 0 \wedge L_{wu} = 0 \tag{4}$$

where $(u, v.w)$ forms are Cartesian co-ordinant system with w in the gradient direction and (u, v) in the perpendicular plane to w (the tangent plane to the isophote).

5 Maxima of the gradient magnitude/Watershed junctions

Junctions in the watershed on the gradient image have been suggested as junction detectors. Although watersheds and their junctions can not be expressed with a local differential expression one can give necessary but not sufficient local conditions and their generic behavior can be determined. The watershed junctions are located at maxima of the flooded image. Watersheds are a subset of the separatrices which are the special flowlines connecting critical points. In saddle points four separatrices meet: two going up to a maximum and two going down to a minimum. In a minimum several separatrices can meet but they do not separate different catchment basins. In a maximum several separatrices can meet which separate several minima. Hence these are candidates for junctions.

The first and second order structure of the gradient magnitude squared G is given in tensor notation by:

$$G_j = L_{ij} L_i \tag{5}$$

$$G_{jk} = L_{ij} L_{ik} + L_{ijk} L_i \tag{6}$$

For a two-dimensional domain, the second order invariants, the determinant and trace of the hessian matrix, are given by :

$$Det(G_{jk}) = G_{jj} G_{kk} - G_{ij} G_{ji} \tag{7}$$

$$= (L_{ij} L_{ij} + L_{ijj} L_i)(L_{ij} L_{ij} + L_{ijj} L_i) - (L_{ij} L_{ik} + L_{ijk} L_i)(L_{ik} L_{ij} + L_{ikj} L_i) \tag{8}$$

$$G_{jj} = L_{ij} L_{ij} + L_{ijj} L_i \tag{9}$$

Equation 5 is the condition for a critical point in G. Outside global minima $\nabla L \neq 0$ the condition consequently describes a degeneracy in the second order structure for L in the direction of the gradient for L. In other words points on the parabolic curves for L (parabolic hyper surfaces) where the degenerated second order structure is in the gradient direction.

The condition for a critical point for G expressed in the gauge coordinate system (v, w) where w is the gradient direction is:

$$0 \neq L_w \tag{10}$$

$$0 = L_{wj} \tag{11}$$

In a maximum we have

$$\begin{aligned} 0 &< L_w(L_{wvv}L_{www} - L_{wwv}^2) + L_{vv}^2 L_{www} && (12) \\ 0 &> L_{vv}^2 + (L_{wvv} + L_{www})L_w && (13) \end{aligned}$$

Figure 6 (bottom right) illustrates the conditions.

From equation 13 we deduce that $0 > L_{wvv} + L_{www}$ since both L_{vv}^2 and L_w are positive. We want to show that $0 < L_{www}$ in maximum. We do it by contradiction. Assume $L_{www} > 0$, this implies that $L_{www}L_{wvv} = -|L_{www}||L_{wvv}| < -L_{www}^2$. From this and equation 12 we deduce that

$$\begin{aligned} 0 &< L_w(L_{wvv}L_{www} - L_{wwv}^2) + L_{vv}^2 L_{www} \\ &< L_w(-L_{www}L_{www} - L_{wwv}^2) + L_{vv}^2 L_{www} && (14) \\ &\Rightarrow \\ 0 &< L_{www} && (15) \end{aligned}$$

This is a contradiction and we conclude that $L_{www} < 0$ is always true for a maximum. So the maxima will always lie on the classical edges $L_{ww} = 0 \wedge L_{www} < 0$. Consequently watershed junctions will also always lie on the classical edges. Note that these watershed junctions are not the same as intersection points for the classical edge detector. One major difference is that intersection points for the classical edge detector are not generic.

Minima and saddle points for G can both lie on edges and not. Figure 6 (bottom row) illustrates the singularities for the gradient magnitude expect global minima.

In a minimum we have:

$$\begin{aligned} 0 &< L_w((L_{wvv}L_{www} - L_{wwv}^2) + L_{vv}^2 L_{www} && (16) \\ 0 &< L_{vv}^2 + (L_{wvv} + L_{www})L_w && (17) \end{aligned}$$

In a saddle we have

$$0 > L_w((L_{wvv}L_{www} - L_{wwv}^2) + L_{vv}^2 L_{www} \qquad (18)$$

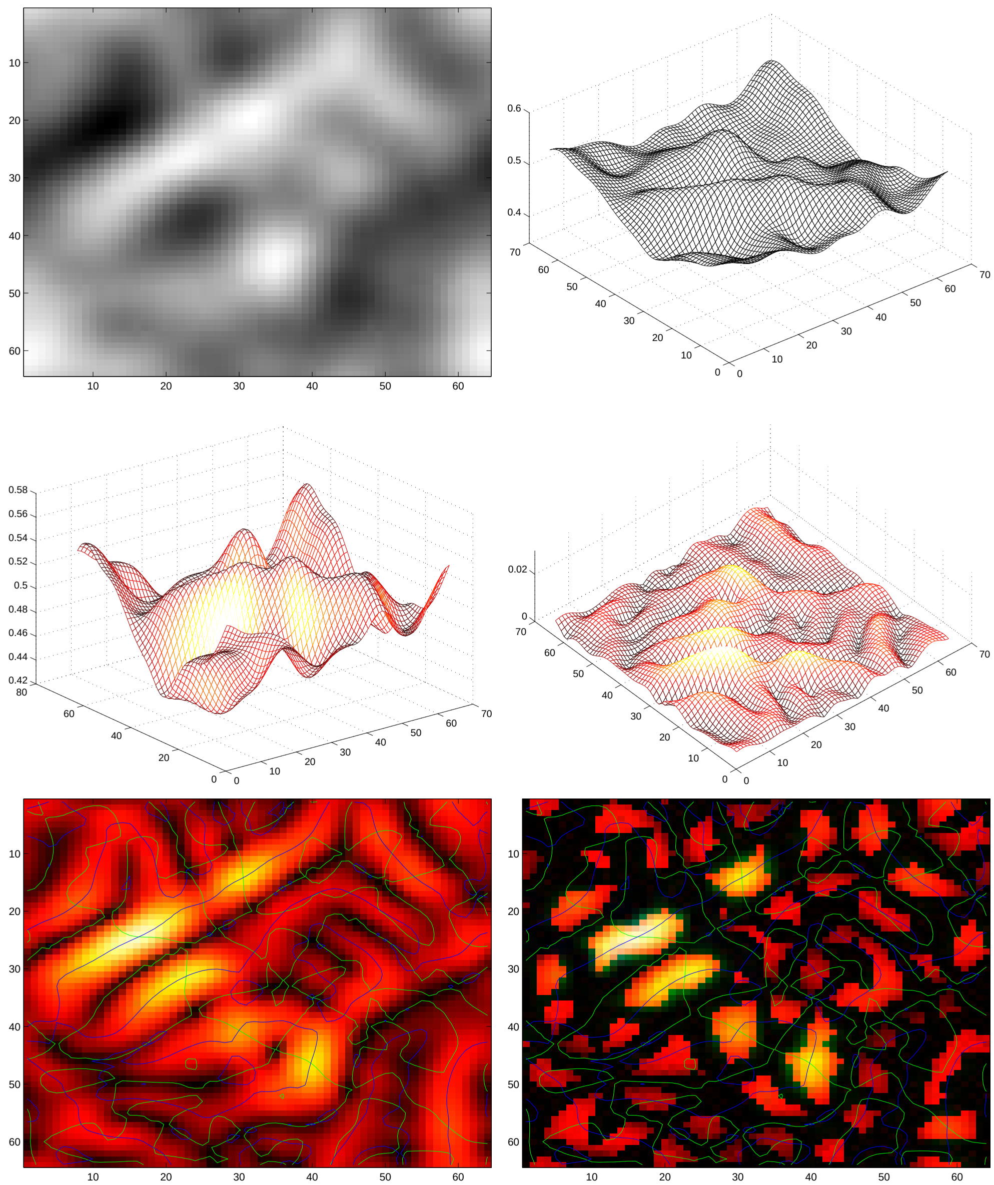


Figure 6: A blurred version of an image generated by random values. Top row, left: the image is depicted with intensity proportional to value. Top, right: the image is plotted as a graph over the image domain. Middle, left: the image plotted as a graph with a shading proportional to the magnitude of the gradient. Middle, Right: The magnitude of the gradient plotted as a graph. Bottom left. The gradient magnitude with zero-crossings for L_{ww} and L_{vw}. Bottom, right: The zerocrossing for L_{ww} and L_{vw} plus the gradient magnitude with the black areas where the conditions for a maximum is not fulfilled.

6 Deep structure of the gradient magnitude

The deep structure denotes the structure between scales of the studied entities. In this section we will describe how the singularities can interact generically when the scale parameter is tranversed.

It has been shown that when a catastrophe occur then it occur in one and only one direction. Consequently the non-degenerate directions can be ignored in the further analysis according to the splitting lemma [6, 23, 12, 18].

6.1 Local structure around catastrophes

We start this section by presenting in figure 7 the result of the following analysis. The result is that the fold and cusp catastrophes are the only generic events.

In the following we assume without loss of generality that the degenerate direction is the x direction (or index 1). We are only interested in the immediate neighbourhood of ($0;0$), and disregard terms of $O(t^2)$. Since the value of $\breve{G}$ (defined below) is unimportant the zeroth order term in x is disregarded.

$$\begin{aligned}\breve{L}_m(x,t) &\equiv \tilde{L}_m(x,0,\ldots,0;t)\\ &= L_m + L_{mjj}t + (L_{mx} + L_{m1kk}t)x\\ &\quad +\frac{1}{2}L_{m11}x^2 + \frac{1}{6}L_{m111}x^3\end{aligned} \tag{19}$$

$$\begin{aligned}\breve{G} &= \breve{L}_m\breve{L}_m\\ &= (2L_mL_{m1} + 2L_{mjj}L_{m1}t + 2L_mL_{m1kk}t)x\\ &+ (L_{m1}^2 + L_mL_{m11} + (2L_{m1}L_{m1kk} + L_{mjj}L_{m11})t)x^2\\ &+ (L_{m1}L_{m11} + \frac{1}{3}L_mL_{m111} + L_{m1kk}L_{m11}t\\ &\quad +\frac{1}{3}L_{mjj}L_{m111}t)x^3\\ &+ (\frac{L_{m11}^2}{4} + \frac{L_{m1}L_{m111}}{3} + \frac{L_{m1kk}L_{m111}}{3}t)x^4\\ &\quad +\frac{L_{m11}L_{m111}}{6}x^5 + O(x^6)\end{aligned} \tag{20}$$

A singularity of type $L_i = 0$
The Hessian for G is diagonalised by choosing the coordinate system where mixed second order image derivatives are zero, hence $L_{kl} = 0, k \neq l$. So far we have change the coordinate system to simplify the terms and we are analysing an event of codimension n due to the constraint $L_i = 0$. Next we increase the codimension with one by introducing an extra constraint and analyse the outcome of this: The constraint $L_{11} = 0$ is imposed in this system, the geometrical meaning is that the degenerate direction is aligned with one of the main curvature directions for the image. The local structure is given by:

$$\breve{G} = L_{mjj}L_{m11}tx^2 + (L_{m1kk}L_{m11} + \frac{1}{3}L_{mjj}L_{m111})tx^3$$

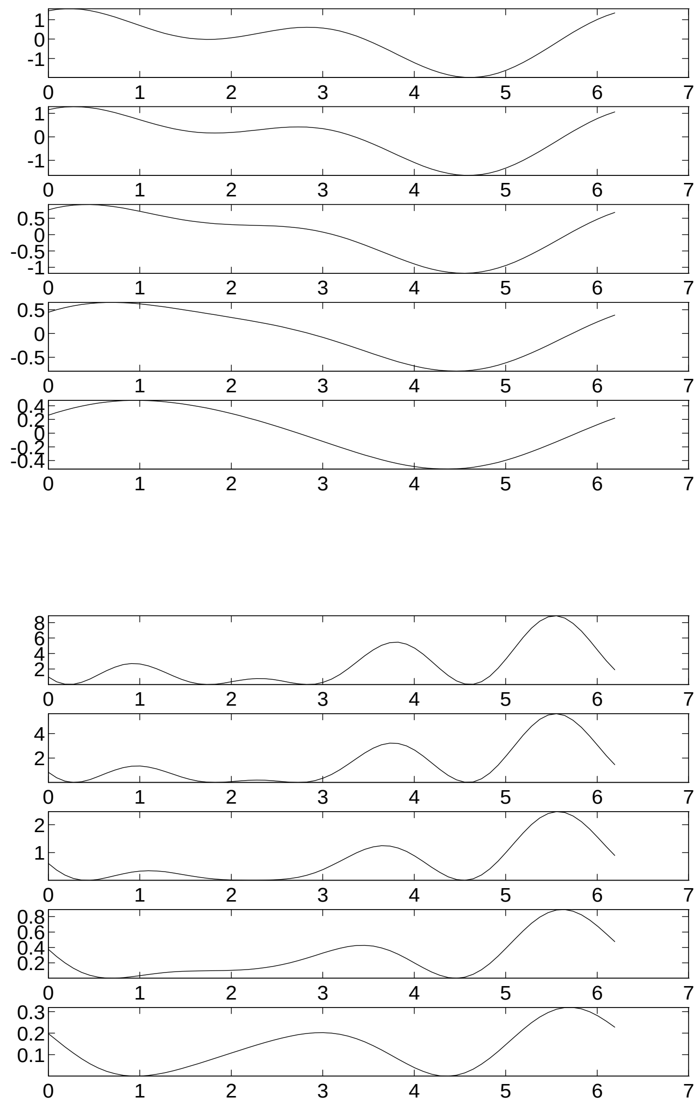


Figure 7: The top subfigure displays a 1D signal embedded in a family of Gaussian blurred versions of the signal. The top frame is the original signal and the degree of blurring increases downwards from frame to frame with the most blurred version presented at the bottom. From the top frame to the middle frame a fold catastrophe takes place approximately at $x = 2.3$, the minimum at $x = 1.7$ and the maximum at $x = 2.7$ in the top frame annihilates in the middle frame. From middle frame to the last frame the concavity at approximately $x = 2.3$ disappears. The bottom subfigure displays the squared first order derivative of the signal in the top subfigure. From the top frame to the middle the minima at $x = 1.7$ and $x = 2.7$ and the maximum at $x = 2.3$ meet in a cusp catastrophe. In the middle frame the singularities have merged into one minimum at $x = 2.3$. From the middle frame to the bottom frame the fold catastrophe happens. The minimum at $x = 2.3$ annihilates with the maximum at $x = 1.4$.

$$+ \quad (\frac{L_{m11}^2}{4} + \frac{L_{m1kk}L_{m111}}{3}t)x^4 + O(x^5) \tag{21}$$

For $t = 0$, all structures up to fourth order vanish for the gradient squared. This event described algebraically by Equation 21 is known as a cusp catastrophe. It is observed that the algebraic constraints $L_i = 0$ and $L_{11} = 0$ for a cusp catastrophe in the gradient squared coincide with a fold catastrophe in the image. In 1D, the annihilation of a minimum and a maximum in the image L corresponds to the merging of two minima and one maximum into one minimum for L_iL_i.

Note that the sign of the x^2 coefficient $L_{mjj}L_{m11}t$ determines whether singularities emerge or disappear. In the 1D case ($m = j = 1$), the sign depends solely on t and since t is increasing (and never decreasing) only disappearing is possible. For higher dimensionality, appearing singularities are possible.

The type of the singularity, a minimum, can be read from the fourth order coefficient evaluated at $t = 0$ equal to $\frac{L_{m11}^2}{4}$.
A singularity of type $L_{wj} = 0$.
The singularity condition ($L_mL_{mj} = 0$) and the degeneracy condition ($L_{i1}^2 + L_iL_{i11} = 0$) reduce Equation 20 to :

$$\begin{aligned} \breve{G} &= (2L_{mjj}L_{m1} + 2L_mL_{m1kk})tx + (2L_{m1}L_{m1kk} \\ &\quad +L_{mjj}L_{m11})tx^2 + (L_{m1}L_{m11} + \frac{1}{3}L_mL_{m111} + \\ &\quad (L_{m1kk}L_{m11} + \frac{1}{3}L_{mjj}L_{m111})t)x^3 + O(x^4) \end{aligned} \tag{22}$$

For $t = 0$ all structure up to third order disappear, this event is referred to as the fold catastrophe. In the 1D case, Equation 23 makes it clear that only annihilation of a maximum and a minimum is possible:

$$\breve{G} = L_1L_{1111}(2tx + \frac{1}{3}x^3) + O(x^4) \tag{23}$$

In higher dimensions, say 3D, the reverse event is possible. The number and position of the singularities can be determined from equation 22. The number and positions will be a function of t, and this dependency tells whether it is an annihilation or creation catastrophe.
Summary
The types of catastrophes generically occurring for L_iL_i are determined by the dimensionality of the domain, and the type of singularities involved. This is summarised in table 1.

In the event of the cusp catastrophe, terms of order three and less disappear. This event coincides with the fold catastrophe in the image structure following the linear diffusion equation. The fold catastrophe for L_iL_i corresponds to a degeneracy in the second and third order in the underlying image structure. Singularities can generically disappear with increasing scale regardless of the dimension of the domain. Appearing singularities can and will occur for domains of dimension two or higher.

Condition Catastrophe	Singularity	Type	Creating events
Fold	$L_m L_{mj} = 0$	$L_i L_i \neq 0$	$N > 1$
Cusp	$L_m L_{mj} = 0$	$L_i L_i = 0$	$N > 1$

Table 1: All generic catastrophes for $L_i L_i$ are listed with the necessary conditions for their presence. The dimensionality of the domain is denoted N. The singularity condition has to be fulfilled for a catastrophe to occur. The additional condition is required for creating events to be possible.

7 Different possible choices of graph representation

Multi–resolution segmentation techniques attempt to gain a global view of the image structure by examining it at many different resolutions. The coarse scale view gives the global structure, the fine scale view provides the details. For the ideas of coarse and fine scale structures to make sense it is expected that the number of segments decreases with increasing scale. Multi–scale techniques use a fine sampling in the scale direction to make it practical to link structures across scale. The term "tracking of structures" is also used since a motion of the image structure takes place over scale. The design of a multi–scale segmentation method must consider which image structures are of interest, analyse these structures and their expected transformations with scale and determine how to link them across scale.

The terms cause and effects will be used. A cause is an entity at some fine scale causing an entity, an effect, at a coarser scale.

Multi-scale algorithms produce segments extending over scale. In other words, the linked segments will form a scale–space segment extending both in the spatial as well as in the scale direction. Taking a slice of the scale–space segment corresponding to a particular scale produces the segmentation at that scale. The linking scheme determines the possible topological form of the scale–space segments. A multi–scale segmentation algorithm allows:

- **disappearing, (annihilation)**
  if a scale–space segment may exist on a fine scale but not on a coarser scale.
- **creation,**
  if a scale–space segment may exist on a coarse scale but not on a finer scale.
- **merging,**
  if a scale–space segment is disconnected at a fine scale and connected at a coarse scale.
- **splitting,**
  if a scale–space segment is connected at a fine scale and disconnected at a coarse scale.

A slice of a scale–space object/segment for a particular scale will give the object/segment at that particular scale. For a 3D scale–space segmentation the slice will be a segmentation of the domain for a 2D blurred version of the original image where the degree of blurring corresponds to the chosen scale.

7.1 Creating the graph structure with linking

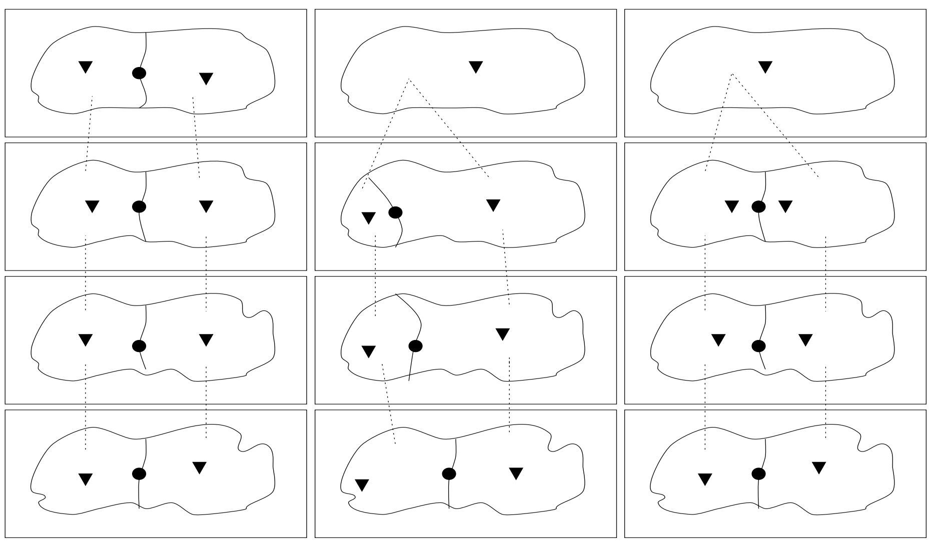
In this section are established possible linking schemes for segments defined by local minima for the gradient squared. The linking will give us the graph or tree presentation. The linking

of catchment basin segments can be based on the analysis of the singularities, due to the duality between segments and the minima of the gradient squared. There are five generic events for minima of the gradient squared:

- No interaction with other singularities
- Creation in a pair with a saddle
- Annihilation with a saddle
- Merging with a saddle and another minimum into one minimum
- Splitting into saddle and two minimum

The above reference to a saddle hold for the 2D case. In the 1D case all occurrences of saddles should be replaced with the word maximum and in 3D the replacement should be hypersaddle.

A linking establishes the structural relation between segments at different scales. The cause and effect relationship is needed for projecting the coarse scale structure to a sufficiently fine localization scale. The term projection will be used in the sense that a coarse scale structure can be tracked via the linking to a finer scale. If one coarse scale segment links to several fine scale segments then the union of these segments are the projection of the coarse scale segment.



(a), No change (b), An annihilation (c), A merge

Figure 8: Scale increases from bottom to top. No change in topological structure is presented in (a). In (b) is a segment annihilated. The corresponding minimum annihilates with the saddle on the boundary to the neighbouring segment. In (c) two segments merge.

It will be instructive to see the motion of some segments and minima together. In Figure 8 and 9 is presented the relation between segments and singularities based on the gradient squared. The chosen linking scheme is indicated by broken lines; the possible linking schemes will be discussed in section 7.2. In the figures the following symbols are used: A triangle pointing downwards indicates a minimum, a filled circle indicates a saddle, a unbroken line indicates a border between segments and broken lines symbolises a link between segments. The relation between segments and singularities will now be discussed. Scale always increases upwards in the figures, and the description of the events will be from fine to coarse, so bottom up. When there is no interaction between the singularities, it simply implies that the topology for the segments stay the same. This situation is displayed in Figure 8 (a): the minima (the triangles) move about but do not interact. In (b) an annihilation takes place. The left minimum and the saddle annihilate, in the segmentation this implies that the neighbouring minimum (the right one) takes control of the region previously belonging to the annihilated minimum. In (c) the two minima and the saddle move towards each other and interact in a cusp catastrophe this implies for the segments that the dividing border disappears and one segment remains; a merge has taken place. In Figure 9 (left) A minimum and a saddle is created via a fold catastrophe. The created segment conquers more and more of the image domain previously owned by the stable minimum to the right. A new segment has emerged. In the right subimage two minima and a saddle split from one minimum causing the creation of a new segment.

The illustration of the creation and the split are schematic to emphasize the change of structure. In practice emerging segments seem to disappear fairly quickly after appearing. They have a short lifetime over scale in scale–space. The range of scale over which an object exists has been observed by Witkin [28] to be a possible measure for the importance of the object.

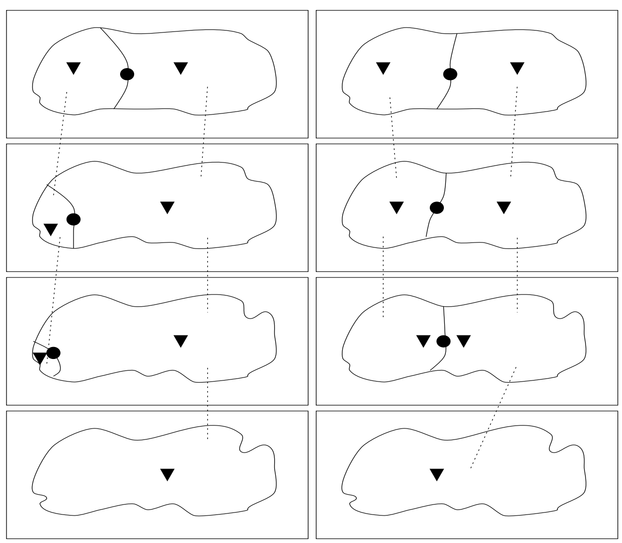


Figure 9: Scale increases from bottom to top. Left to right: a segment is created, a segment splits into two segments

7.2 Possible linking schemes

We shall present three linking schemes based on the derived generic events for the singularities. Several considerations have to been taken: How can it be implemented. What will be the projection of a single segment. What will be the projection of a coarse scale segmentation. An end user of the systemwill typically only be interested in the final result, the projection of a coarse scale segmentation, that is the well localised coarse scale structures.

The three linking schemes are itemised according to the resulting structure of links. The last scheme is the one implemented and the one displayed in Figure 8 and 9. The differences and similarities are discussed with this linking scheme as reference. The differences will arise for the annihilation, creation and split events. The events of no change and merging do not leave much choice for a reasonable linking scheme and are found to be similar for all schemes. Projecting the coarse scale segment after a merge results in a supersegment consisting of the two merged segments. Projecting the two coarse scale segments after no change of topology simply localises these segments at the finer scales. Projections for the remaining types of events will be discussed for each linking scheme.

A set of undirected graphs.

A theoretical possibility for links are the paths of the minima. This linking requires a robust way of tracking singularities even through catastrophes. Point tracking through catastrophes is not feasible in practice. The tracking through catastrophes is necessary to differentiate between annihilation and merge, and between creation and split. Tracking of singularities will be further discussed in Section 8.1.

The linking in Figure 8 and 9 has to be altered to reflect such a linking. For the annihilation in Figure 8 (b) the left link from the top box has to be removed and for the split in Figure 9 (b) a link should be added from the bottom segment to the left segment above. The remaining links stay. This linking structure will be a set of undirected graphs.

After an annihilation event, the projecting of the coarsest scale segment results in one segment at the finest scale (in the Figure 8 it is the rightmost). After a creation event a projection of the created segment will result in nothing at the finest scale. The neighbouring coarse scale segment projects to the one segment at finest scale. After a split event either coarse scale segments will project to the one segment at finest scale.

A projection of an entire segmentation at a coarse scale will correspond to a segmentation at fine scale: Some regions are not segments because they do not cause any segment at the coarse scale, this is the consequence of an annihilation somewhere between finest and coarsest scale. Some regions will correspond to two or several different segments, this is the consequence of splittings.

Two things suggest that the similar catastrophes should be treated in the same manner, namely (1) the problem of distinguishing between the catastrophes and (2) the events for the segments are similar. Let us elaborate point (2) a bit: an annihilation of singularities does not imply that a segment simply disappears out in the blue, as previously noted it implies that the neighbouring segment moves in and take over the disappearing segment. Hence the same kind of linking should be applied in the event of an annihilation and a merge. Similar should the split and creation be treated equally. The linking of a merge event as presented in Figure 8 seems to be *the* right way to link otherwise the number of segments will increase instead of decrease with increasing scale. Consequently the linking for the annihilation event must follow the merge and not vice versa. The two remaining schemes will use this approach, the linking is presented in Figure 8. For the creation and splitting events there are two possibilities either the creation adapts to the splitting event or vice versa. These two possibilities correspond to the last two linking schemes presented below. The latter scheme is presented in Figure 9.

A set of undirected graphs without ends at in between scales.
The linking in the case of creation and splitting is unified by linking the segments according to the involved saddle. For the creation it means that the created segment is linked to the same segment as its neighbour. The neighbourhood relation is given by the created saddle. For the splitting event it implies that the splitting segment links to both subsegments. In Figure9 two links should be added to illustrate this scheme. One link in (a) between the created segment (the smallest one) and the large segment at the finest scale and in (b) one link between the unlinked, split segment and the finest scale segment.

A projection of one segment will always be at least one segment at the finest scale.

A projection of an entire coarse scale partition will result in a fine scale partition. Because of the linking of splitting and creation event one fine scale segments can corresponds to several coarse scale segments. So one could say that in this partition several coarse segments lay on top of each other or co–exist at the same location.

The resulting linking structure is a set of undirected graphs where all loose ends are located at finest scale and the coarsest scale.

A set of trees, (hierarchies).
The last linking is the one presented in the Figures 8 and 9. Here merge and annihilation are also treated equally; the involved segments are linked to the resulting coarse scale segment. Split and creation is treated in the following way. The split is adjust to the creation. A creation is a dangling end in the linking tree. The two split links are pruned, so only one link remain between the fine scale segment and one of the segments above. A simple pruning method is to keep the link between the segments which have highest spatial correlation.

The linking structure is a set of trees. The size of set reduces when coarser and coarser scales are included. Each segment in the structure is an internal node or an external node (a leave) in the tree. Each internal node has links downwards in scale and these links form a tree structure, a hierarchy.

For this link scheme the projection of one segment results either in non, one or several fine scale segments. The number of segments corresponds to the number of leaves located at the finest scale. No segments results from projecting a segment where the downwards links end before the finest scale is reached.

The projection of an entire partition will result in a partition. One realises this since each segment at the finest scale has a path of links up to a segment at the coarse scale. So a projection from any coarse scale will result in a partition. Since each segment only have one upward link, a segment will always be the projection exactly of one segment.

The last linking scheme introduces an interesting and pleasant property: The projections to the same fine scale but from increasing coarser scales form by definition a object hierarchy where objects are formed by the merging of sub objects. The sub objects are merged if they are causing the same coarse scale object. This effect can be observed in Figure 10 top row, left to right.

8 Detection methods

8.1 Tracking singularities

The problem of point tracking singularities through catastrophes has not been solved. The main problem is that features can move arbitrarily fast. Let a feature be defined by:

$$g(x;t) = 0, \;\; x \in \mathbb{R}^d, t \in \mathbb{R}_+ \tag{24}$$

The drift velocity of this feature can be estimated [15]:

$$\partial_t x = -(\partial_x^T g)^{-1} \partial_t g \tag{25}$$

This means the drift velocity does not have an upper bound. This should be understood in the following sense: given a scale interval and spatial distance then there exists a signal for which the considered feature moves further than the given spatial distance within the allowed scale interval. In other words, let a fixed cylinder in scale–space be defined by a spatial radius and a scale interval. Let some feature be located in the middle of the cylinder, then there exists a signal such that the path for the feature in scale–space will cross the side of the cylinder.

In practice this means that is hard to distinguish between annihilation and merge, and between creation and split. As an example consider figure 8 (b) and (c): if only the top and bottom slice are available then it is impossible to distinguish the two events. A finer sampling in the scale direction can lighten the problem, but does not solve it. The core of the problem is the quantification of the scale–space. The quantification determines when an event can be classified as one or the other. More concretely, if a toppoint for a fold catastrophe is located in a quantification cube of space through which another singularity path passes then it is not possible to say if the event was a fold catastrophe or a cusp catastrophe. This is possible no matter how small the quantification cube is. Hence, a discrimination between the catastrophes are inherently connected to the quantification.

Stilian Kalitzin [9] have suggest a promising way of detecting singularities using results from integral and differential geometry. The concept of winding number is used. The winding number can be defined and calculated in each point of image domain by integrating the vector field along a closed curve, the curve have to be without singularities. The integral equals an integer called the winding number for the region, in the limit of an infinitely short curve the winding number is for the surrounded point. Each type of singularity has a specific number. The integral over a closed curve give the sum of the winding numbers for the enclosed region. Flanders [4] has an introduction to winding numbers and further references. The method seems robust and it seems possible to make the linking of slow singularities merely by matching equal winding number across scale. In practice the detection scheme is basicly region–based since the curve will have a finite length. The problem of the catastrophes has not been solved by the method. The detection scheme is not mathematically defined for a catastrophe point since it is not possible to enclose the point by a curve without singularities. But the approach seems promising.

Kuijper [11] has reported interesting and good results on tracking of singularities using a point based approach although a quite fine sampling in the scale direction is inherently needed.

Lifshitz and Pizer tracked extrema and regular points in their segmentation scheme. They reported problems with the robustness of this scheme. Especially the serious problem of non-containment: a point which at one scale has been classified to be within a region associated to a local maximum, can wander out of that region when scale increases and the linking paths can be intertwined in a rather complicated way.

Lindeberg [12, 13] suggested a region–based tracking scheme of extrema using the blobs as described in section, and reported good results using this scheme.

8.2 The segment linking scheme

The first presented linking scheme for the gradient squared has to rely on a method for tracking singularities and connecting them in catastrophes. The two other schemes are based on the following observations: an annihilation and a merge correspond to the merging of two segments; a creation and a split correspond to the appearance of a new segment or the splitting of an existing segment.

The focus on segments in the linking scheme allows the implementation to be based on segments. Structures move smoothly with scale, hence if the scale sampling is sufficiently fine then segments can be matched by measuring the spatial overlap with the segments at neighbouring scales. So even though the boundaries might move too fast to be tracked, the entire segment can still be tracked. Another important observation is that the appearing or disappearing border is limited in its motion by the outer borders of the segments. So even when a border is about to disappear and moves fast, the remaining outer borders are fixed relative to the motion of the disappearing border.

Hence, in the limiting case of extremely fine scale sampling: segments are linked correctly if the segments are linked according to spatial overlap with segments at neighbouring scales. The observation on the relative stability of the outer boundaries implies that the subsegments do not move significantly outside the border of the resulting segment during the catastrophe event. This latter observation justifies the method for a fine scale sampling.

The latter of three linking schemes have been implemented by a pair–wise scale linking of the segments using a linking criteria of maximum spatial correlation. A fine scale segment is linked to the one coarse scale segment with which it overlap the most. A similar idea was used for linking of blobs by Lindeberg [13].

The second of three schemes could have be implemented by regarding maximum overlap both in the bottom–up scale direction and in the up–down scale direction. The last presented and simplest linking scheme has been implemented.

A short summary including tree comments

The linking criteria is in practice implemented by maximum spatial overlap for cause segment to the effect segments. In the limit of very fine scale sampling this corresponds to linking the minima. The linking for minima is based on the appearing or disappearing saddle connect the involved minima. When a saddle appear a new segment is created, when a saddle disappear two segments are merged into one.

The resulting representation is a forest of trees. The forest can be artificially connect to a single tree by introducing a super root node this might introduce problems in matching problems so care shold be taken. All trees have their node at the top level. Leaves are not

restricted to any particular level but can occur on all levels. In the generic case a node has zero, one or two children but faster implemtations of the linking will introduce several children.

8.3 Linking in action

In Figure 10 we try to visualize the ideas and their connection. An artificial made image of a check board plus some low scale noise will be segmented as an example.

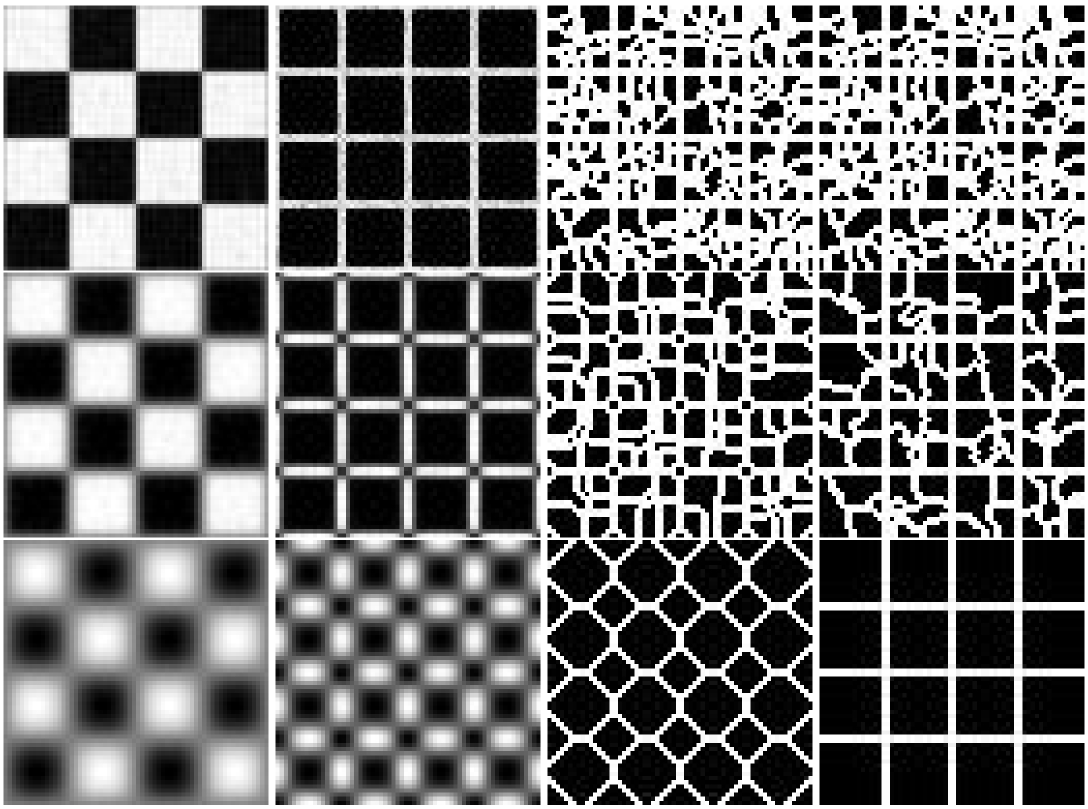


Figure 10: Each row corresponds to the same scale. The sigma values left to right ($\sigma = 0.50, 1.43, 3.70$ pixels) The left column are slices in the scale–space image, the original image is 64x64 pixels and is an artificial check board with fine scale noise added. Column number two is the gradient magnitude image. Column number three is watersheds of the gradient magnitude. Column number four shows the result of tracking the segments in row three down to the scale in the first column.

In the left columnare slices from the scale–space image where scale increases top down. The other columns describe the structure in the left column in different ways. The second column shows the squared of the gradient magnitude. The third column shows the watersheds of the squared of the gradient magnitude. The boundaries are marked by colouring the pixel on each side of the boundary. Note how the structures at low scale include many boundaries which normally are consider to be unwanted. What is usually meant by over-segmentation is that the detected structure is detected at a too fine scale. At the highest scale (row three) the low scale structure is not present. The price paid for this result is a delocaliza-

tion of the boundaries. The answer to this is to probe the deep structure of the segmentation and determine the fine structure causing these interesting coarse scale structure. In practice this means tracking the structures down to a sufficiently fine scale. The result of this is shown in the fourth column. The segments in column three is tracked down to the scale presented in row one. Hence the sub-pictures in row one, column three and four is identical. In row three is the coarse scale structure of interest and this is tracked to scale of localization which provide the wanted segmentation. The gradient squared at coarsest scale (column two, row three) have some large black regions and some small black regions. Black is low value. The large black regions corresponds to black (minima) and white (maxima) squares in the checker board. The small black regions corresponds to the grey (saddles) regions in the checker board. The watersheds of the gradient squared (column three, row three) divide all regions both the large and small ones. It looks like large and small octagons where the large ones correspond to the extrema in the image and the small ones correspond to the saddles in the image. These watersheds are track down to finer scale and the result is displayed in column three. All the large octagons have been tracked down to the well localised squares (note the number of large octagons equal the number of squares). The small octagons are tracked down to regions which are smaller than the resolution in the segmentation. The reason is that the saddle regions in the image are very small at fine scale since the corners of the squares are sharp here. When the image is blurred the extension of the saddle regions grow and at the coarsest scale they have grown to a size given be the small octagons.

9 Summary

We have describe the superficial and deep structure of the gradient magnitude derived from an image. That is, the appearence of the differential when the spatial and scale parameters are changing. The local structure around the spatial singularities has been described and their relation to catchment basins and watershed junctions have been summarised. The interaction of the spatial singularities with changing scale has been described. The generic interaction events have been the basis for discussing possible linking of singularities in the scale parameter directions. The generic events are a fold and cusp catastrophe. In dimensions higher than one both creation and annihilation versions of the two types of catastrophes can occur.

The relation to a graph representation has been emphasied. Each minima on any level corresponds to exactly one node in a forest of trees. The parent-children relation in the trees is dictated by the choice of linking scheme. The linking scheme is based on the analysis of the interaction of the singularities.

References

[1] Nicholas Ayache. Medical computer vision, virtual reality and robotics. *Image and Vision Computing*, 13(4), May 1995.

[2] J.W. Bruce and P.J. Giblin. *Curves and singularities*. Cambridge University Press, 1984. ISBN 0-521-42999-4.

[3] James Damon. Local morse theory for solutions to the heat equation and gaussian blurring. *Journal of Differential Equations*, 115(2), January 1995.

[4] Harley Flanders. *Differential Forms with Applications to the Physical Sciences*. Dover Publications, Inc., New York, 1989. ISBN 0-486-66169-5, Published by Academic Press in 1963.

[5] Luc M J Florack. *Image Structure*. Kluwer, 1997.

[6] Robert Gilmore. *Catastrophe Theory for Scientists and Engineers*. Dover Publications, Inc. 1993; Originally by John Wiley and Sons, 1981. ISBN 0-486-67539-4.

[7] Lewis D. Griffin. *Descriptions of Image Structure*. PhD thesis, Uni. of London, 1995.

[8] C. Jordan. Sur les lignes de faite et de thalweg. *C.R.Acad.Sc. Paris*, 74:1457, 1872.

[9] S. Kalitzin. Topological numbers and singularities in scalar images. scale space evolutio n properties. In Jon Sporring, Mads Nielsen, Luc Florack, and Peter Johansen, editors, *Gaussian Scale–Space*, chapter by Kalitzin. Kluwer Academic Publishers, 1996.

[10] J. J. Koenderink. *Solid Shape*. MIT Press, Cambridge, Mass., 1990. ISBN 0-262-11139-X.

[11] A. Kuijper. *The Deep Structure of Gaussian Scale Space Images*. PhD thesis, Utrecht University, 2002.

[12] Tony Lindeberg. Scale–space behaviour of local extrema and blobs. *Journal of Mathematical Imaging and Vision*, 1:65–99, 1992.

[13] Tony Lindeberg. *Scale–Space Theory in Computer Vision*. The Kluwer International Series in Engineering and Computer Science. Kluwer Academic Publishers, Boston, USA, 1994.

[14] Tony Lindeberg. Scale-space: A framework for handling image structures at multiple scales. Technical Report CVAP-TN15, Department of Numerical Analysis and Computing Science, Stockholm University, 1996. Technical Report CVAP-TN15.

[15] Tony Lindeberg and Bart ter Haar Romeny. Linear scale–space. In ter Haar Romeny [26], chapter 2. ISBN 0-7923-3087-0.

[16] J.C. Maxwell. On hills and dales. *The London, Edinburgh, and Dublin Philosophical Mag. and J. Sci, 4th Series*, 40:421–425, dec. 1870. reprinted in The Scientific Papers of James Clerk Maxwell, W.D.Niven, Ed. Vol.II. New York: Dover, 1965.

[17] Lee R. Nackman and Stephen M. Pizer. Three-dimensional shape description using symmetric axis transform i: Theory. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 7(2):187–202, 1985.

[18] Ole Fogh Olsen. Multi-scale segmentation of grey-scale images. Technical report, Department of Computer Science, University of Copenhagen, Universitetsparken 1, DK-2200 Copenhagen East, Denmark, August 1996. DIKU-rapport 96/30.

[19] Ole Fogh Olsen. Generic image structure. Technical report, Department of Computer Science, University of Copenhagen, Universitetsparken 1, DK-2200 Copenhagen East, Denmark, March 2000. Technical Report DIKU-2000/04.

[20] Ole Fogh Olsen and Peter Johansen. Classification of toppoints for the gradient squared. Technical report, Department of Computer Science, University of Copenhagen, Universitetsparken 1, DK-2200 Copenhagen East, Denmark, Oct 1998. Technical Report DIKU-98/24.

[21] Ole Fogh Olsen and Mads Nielsen. Generic events for the gradient squared with application to multi-scale segmentation. In *Scale-Space Theory in Computer Vision, Proc. 1st International Conference*, volume 1252 of *Lecture Notes in Computer Science*, pages 101–112, Utrecht, The Netherlands, July 1997.

[22] Ole Fogh Olsen and Mads Nielsen. Multi-scale gradient magnitude watershed segmentation. In *ICIAP'97 - 9th International Conference on Image Analysis and Processing*, volume 1310 of *Lecture Notes in Computer Science*, pages 6–13, Florence, Italy, September 1997.

[23] T Poston and I N Stewart. *Taylor Expansions and Catastrophes*. Pitman, 1976. ISBN 0-273-009-64-8.

[24] J.H. Rieger. Topographical properties of generic images. *IJCV*, 23(1):79–92, May 1997.

[25] R. Rothe. Zum problem des talwegs. *Sitzungsber.Berliner Mathem. Gesellschaft*, 14:51–68, 1915.

[26] Bart M. ter Haar Romeny, editor. *Geometry-Driven Diffusion in Computer Vision*, volume 1. Kluwer Academic Publishers, 1994. ISBN 0-7923-3087-0.

[27] Luc Vincent and Pierre Soille. Watersheds in digital spaces: An efficient algorithm based on immersion simulations. *Pattern Analysis and Machine Intelligence*, 13(6), June 1991.

[28] Andrew P. Witkin. Scale space filtering. In *Proc. of International Joint Conference on Artificial Intelligence (IJCAI)*, Karlsruhe, Germany, 1983.