Generic Image Structure

Ole Fogh Olsen

This thesis was set in LaTeX by the author.

Generic Image Structure

Ole Fogh Olsen

This is a Ph.D. thesis given in defence of the Danish Ph.D. degree in Science at the Department of Computer Science, Faculty of Science, University of Copenhagen, Denmark. Supervisors have been Ph.D. Mads Nielsen, IT University of Copenhagen and professor Dr.Phil. Peter Johansen, Department of Computer Science, University of Copenhagen. The advisory board consisted of Docent Tony Lindeberg, Computational Vision and Active Perception Laboratory, Department of Numerical Analysis and Computing Science, Royal Institute of Technology, Sweden and Professor Dr.Phil. Vagn Lundsgaard Hansen, Department of Mathematics, Technical University of Denmark.

Copenhagen, March, 2000

Acknowledgements

This thesis would not have been without the support from a lot of people. I will dare to give the forever incomplete list and in advance to the people not mentioned you were only momentarily forgotten. Thank you.

First of all I would like to thank my supervisors Peter Johansen and Mads Nielsen for their competent advice and our interesting discussions and collaborations.

I would like to thank the rest of the image group at DIKU, especially the numerous people including Kim Steenstrup Pedersen, Andreas Thomsen, Henrik Dohlmann who shared office with me and enduring among other things my so called singing. A special thanks to Jon Sporring and Niels Holm Olsen for many mind boggling discussions. Also thanks to Erik Dam for all your "stupid" questions (I could not answer).

Tony Lindeberg and Vagn Lundsgaard Hansen I thank for taking time in their schedules to participate in my advisory board, for their interest and valuable comments to my project.

Of the guests visiting the image group I especially thank Luc Florack, Lewis Griffin, Joachim Weickert and Robert Mass for sharing their insight into the field of image analysis and many inspiring discussions.

Professor Allen Tannenbaum I will thank for opening his lab at University of Minnesota for me, for his interest and many fruitful discussions. I would also like to acknowledge the rest of the people at the lab for making my stay a pleasant one.

Finally a special thanks to my wife Charlotte for your endless patience and support.

Ole Fogh Olsen

Contents

List of Publications **ix**

1 Generic image structure: Introduction **1**
1.1 Images are measured 2
1.2 Singularity theory in image analysis 4
1.2.1 Caustics and shocks 6
1.3 Outline of the rest of the thesis 6

I Genericity **7**

2 Genericity: Introduction **9**
2.1 Singularity theory I 10
2.2 Generic events for the gradient squared 11
2.3 Watershed junctions as feature detectors 11
2.4 Smoothing creates corners 11
2.5 Singularity theory II: Genericity via equivalence 12
2.6 The structure of the optic flow 13

3 Singularity theory I **15**
3.1 Preliminaries 16
3.1.1 Topological notions 16
3.1.2 Manifolds 17
3.2 Typical solutions 18
3.3 Transversality 20
3.4 Jet space 22
3.5 Transversality theorems 23
3.5.1 Transversality theorem for the heat equation 24
3.5.2 Transversality theorem for Nonlinear PDE's 25
3.6 Stratification 26

4 Generic events for the gradient squared **31**
4.1 Introduction 31
4.2 Scale-space 32
4.3 Generic events for the gradient squared 33
4.4 Watersheds and catchment basins 39
4.5 Linking 41
4.6 Segmentation Results 41
4.7 Conclusion 42

5 Watershed junctions as feature detectors **45**
5.1 Necessary conditions for watershed junctions 46

6 Smoothing images creates corners **51**
6.1 Introduction 51
6.2 Image structure 53
6.2.1 Geometric properties by transverse intersections in jet space 54
6.2.2 Transversal intersections are generic 55
6.2.3 Generic catastrophes for differential corner measures . 55
6.3 Experiments on two typical images 56
6.4 Summary 59

7 Singularity theory II: Genericity via equivalence **61**
7.1 Determinacy 63
7.2 How to read a list of normal forms 64
7.3 Genericity via equivalence for the heat equation 64

8 The structure of the optic flow **69**
8.1 Introduction 69
8.2 Optic flow: notation and definitions 71
8.3 Temporal gauge 71
8.4 Normal flow 72
8.5 Spatio-temporal iso-surfaces 73
8.6 Structure and genericity 74
8.7 Structure of the optic flow field 75
8.8 I-equivalent structure of the optic flow field 76
8.9 S-equivalent structure of the optic flow field 79
8.10 A comment on the multi-scale optic flow structure 82
8.11 Detection of structural changes 83
8.12 Temporal pole evolution 83
8.13 Fixed line scale-evolution 84

8.14 Summary 85

II Segmentation using deep structure 89

9 Segmentation using deep structure: Introduction 91
9.1 Automatic segmentation 92
9.2 Multi scale watershed segmentation 92
9.3 Multi-scale gradient magnitude watershed segmentation 93
9.4 The interactive segmentation tool 93
9.5 Segmentation by compression 94

10 Multi-scale gradient magnitude watershed segmentation 95
10.1 Introduction 95
10.2 Scale-space and local dissimilarity measures 96
10.3 Segments and linking 97
10.4 The interactive segmentation interface 99
10.5 Results and verification 99
10.6 Summary 101

11 The interactive segmentation tool 105

12 Segmentation by compression 109
12.1 Introduction 109
12.2 Least Committed Splitting 111
12.2.1 Olsen's Segmentation Tool 111
12.3 Consistent Merging 114
12.3.1 Specifying Semantics by Compression 114
12.3.2 A General Merging Algorithm 116
12.4 Corners Generate Outliers 119
12.5 Shapes in Data 120

A Smoothing images creates corners 123

B The structure of optic flow 127

Bibliography 130

Sammenfatning (in danish) 143
B.1 Den teoretiske del 143
B.2 Den mere praktiske del 145

Index **145**

List of Publications

In the case of several versions of a work only the latest version is included in this list. First listed are papers included in this thesis in order of appearance.

- "Generic events for the gradient squared with application to multi-scale segmentation" by Ole Fogh Olsen and Mads Nielsen[97] presented at First International Conference of Scale-Space Theory in Computer Vision, Utrecht, The Netherlands. 1997.
  This paper is included as chapter 4.

- "Smoothing Images Creates Corners" by Jon Sporring, Ole Fogh Olsen, Mads Nielsen and Joachim Weickert [125] published in Image and Vision Computing, 1999 .
  This paper is included as chapter 6.

- "The Structure of the Optic Flow Field" by Mads Nielsen and Ole Fogh Olsen[92] presented at European Conference on Computer Vision 1998.
  This paper is included as chapter 8.

- "Multi-Scale Gradient Magnitude Watershed Segmentation" by Ole Fogh Olsen and Mads Nielsen presented [98] at 9th International Conference on Image Analysis and Processing. Florence, Italy.1997
  This paper is included as chapter 10.

- "Segmenting by Compression using Linear Scale-Space and Watersheds" by Jon Sporring and Ole Fogh Olsen [124] presented at Second International Conference on Scale-Space Theory in Computer Vision, Corfu, Greece.1999.
  This paper is included as chapter 12.

- "Multi-Scale Watershed Segmentation" by Ole Fogh Olsen[94], chapter in "Gaussian Scale–Space" edited by Sporring et al.

- "Classification of toppoints for the gradient squared" by Ole Fogh Olsen and Peter Johansen[96] published as technical report, Department of Computer Science, University of Copenhagen, 1998.

- "Multi-Scale Segmentation of Volumes" by Ole Fogh Olsen[95] presented at Øresyn workshop, Lund, Sweden. 1999.
- "Scale-Space Theories in Computer Vision" edited by Mads Nielsen, Peter Johansen, Ole Fogh Olsen and Joachim Weickert[91]. 1999.
- "Interactive Multi-Scale Segmentation in Clinical Use" by Erik Dam, Peter Johansen, Ole Fogh Olsen, Andreas Thomsen, Tron Darvann, Andy B. Dobrzeniecki, Nuno V. Hermann, Noriyuki Kitai, Sven Kreiborg, Per Larsen, Martin Lillholm and Mads Nielsen[22] presented at European Congress of Radiology 2000.
- "Singular points in One-Dimensional Gaussian Scale Space" by Peter Johansen, Mads Nielsen and Ole Fogh Olsen.[57] is accepted for publication in Journal of Mathematical Imaging and Vision. 2000.

Chapter 1

Generic image structure: Introduction

The intended focus of this thesis is the role of genericity in image analysis.

One motivation for this quite theoretical angle has offset in our view on images as physical measurements. A physical measurement has to be stable in the sense that a small perturbation of the input only alters the measured quantity slightly. As we shall see later stability and genericity are closely related.

Another important motivation is to investigate the notion of structure. A common approach in singularity theory is to define structure via equivalence. More specifically one defines an allowed class of smooth deformations which locally can transform the image into a canonical appearance. The set of local forms which can be transformed into the same canonical form constitutes an equivalence class. Each equivalence class corresponds to one type of structure.

In image analysis a common approach is to define the structure of interest as the solution to an algebraic set of image derivatives. Each task, like edge detection or blob detection, has its own definition of structure. This view of structure can also be analysed using theorems from singularity theory. For instance, do we expect a smooth curve as output from the given edge detector? Do we expect any output at all? Do edges intersect themselves? Are corners on an edge?

The two approaches can both be analysed within the frame of singularity theory but there are differences in the strength of the results. We will return to this and elaborate.

We also believe that scale plays an important role in image analysis. Scale space theory supplies a way to simulate change of measuring scale in a continuous manner. The image is embedded in a family of images measured at

continuously coarser scales; in practice simulated with a partial differential equation. This embedding can not be handled by the classical singularity theory developed by Thom [128], Mather[77, 78, 76, 79, 80, 81] and others. The starting point for the classical theory is the full set of all smooth functions. Hence restricting the class of permitted images by the embedding does not fit into the premise for the classical approach. Fortunately Damon[25] has proved similar results for certain partial differential equations including the linear diffusion equation and some non linear diffusion equations.

To illustrate the practical use of genericity considerations the machinery has been used to compare properties of different corner detectors, and a multi scale segmentation tool has been developed for use in a clinical research lab.

The following sections will elaborate on the above themes and finally end the chapter with a guide to the rest the thesis.

1.1 Images are measured

We will consider images in a broad sense. An image is a collection of physical measurements acquired by an ensemble of detectors in a known spatial and temporal ordering. This includes a wide range of types of measurements like pictures and medical scannings.

- A picture could be acquired with a digital camera and the measured quantity would be the number of photons per area. An example of this modality can be seen in Figure 1.1
- A medical scanning could be a two dimensional x-ray picture where the density in a column of tissue is measured by its ability to absorb x-ray radiation.
- Another type of medical scanning is magnetic resonance (MR) in which the relaxation times T1 and T2 are measured. The relaxation times describe how fast the spin of nuclei return to their equilibrium after a magnetic pulse. These times differ between tissue types and can therefore be use to distinguish between different tissues.

Image features are per necessity something the observer defines. The observer has to supply the model in order to infer information from the images. In the mathematical sense the differential structure of the image has an existence of its own but in order to give operational access to these structures the finite limits of the measuring device have to be taken into account. The structure is assessed with an aperture of finite size and only a finite number of derivatives with finite precision can be used.

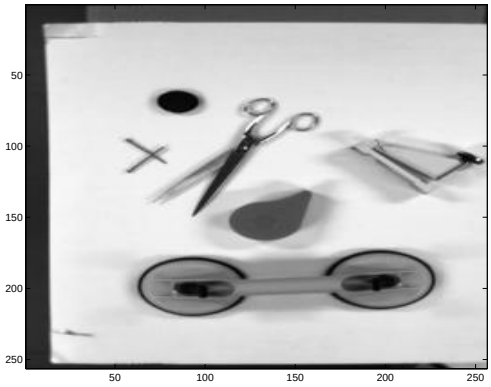


Figure 1.1: A picture of a desk scene.

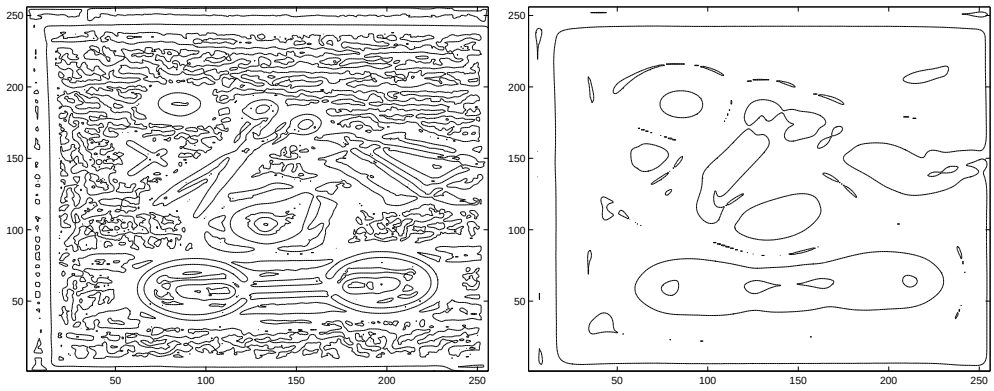


Figure 1.2: Edge maps at two different scales for the image in Figure 1.1. The standard deviation of the Gaussian for the left and right subfigure is respectively 1.5 pixels and 8 pixels. The zero crossing of $L_{ww} = 0$ is used as edge detector (The second order derivative in the direction of the gradient).

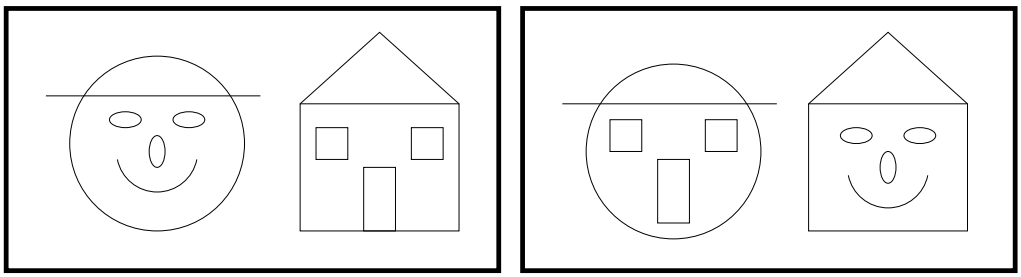


Figure 1.3: Illustration of the importance of connecting structure across scales. Both scenes have two blobs at coarse scale and face details together with house details at fine scale. Without a description of the connection across scales the description of the two scenes are the same.

The inner resolution of the image is given by area and denseness of the detectors. The outer resolution is given by the dimension of the image. In between a continuum of resolutions exists.

Structure change caused by scale variation is central for image understanding. Details disappear with increasing scale but the same process dislocates the remaining structure. A visual example of edges at two diffused scales is given in Figure 1.2. Structure simplifies with increasing scale but structure can also locally appear with coarser scale (depending on the definition of structure).

How does structure connect in the scale direction? With a connection, structure at different scales can be related and the family of images over scale can be regarded as a family and not an unrelated collection of images. In Figure 1.3 is illustrated two scenes with the same features at coarse scale and fine scale; without a description of the connection the two scenes have identical descriptions!

1.2 Singularity theory in image analysis

An image is often perceived as the graph of an intensity function over the spatial domain. However, not every characteristic of this graph is of interest to the observer. An observer only interested in image edges, will consider two images having the same edges as being per definition identical, since every observable (edge) is identical. This leads to the conclusion that the most appropriate operational definition of image structure is the collection of image features. The set of image features is not fixed for all observers/tasks, thus two images of identical structure for one task may have deviating structure for another task (set of features). Since the number of possible operationally

defined features is infinite, we may attribute an infinite number of different structures to a given image.

The structure of a given image is easily assessed in practice since we demand the features to be operationally defined. This is solely a problem of implementation and representation. However, it is also possible to determine how we expect the image structure to express itself even in the case when the image is yet unknown. That is, given a feature or a set of features some structures are impossible, and other structures are unlikely. One well-known example is given in the classical paper by Koenderink [66]: The multi-scale iso-intensity surfaces will never close downwards in scale. Another example is a specific iso-intensity curve in an image. In general one would expect it to be a simple closed curve and it is unlikely that it will intersect itself.

The emphasis is on the likely structure to arise from an arbitrary image. More precisely, we will analyse for the local *generic* image structure. That is, we will analyse which structures will occur in an open and dense set of images. In this sense we will attribute to a given feature detector the typical structure it implies.

Quite some work has been done in this area of attributing generic properties to locally defined feature detectors. We mention here, iso-intensity curves [66, 23], singularities [66, 58, 34, 60], singularities of gradient magnitude [97], edges [106, 107, 108], ridges [26, 27], cores [26, 27], corners [125, 106], and parabolic lines [108]. However, these analyses have been performed "feature by feature". The central focus of this section is to provide a general scheme, which may to a high degree be automated. The spirit of this methodology follows the lines first outlined in image processing by Rieger[106, 107, 108], and later successfully applied to other features by Olsen and Nielsen[94, 97], and Sporring et al.[125] .

The fundamental work of finding generic properties of solutions to partial differential equations and especially to the Heat Equation (the Gaussian scale-space) was performed by Damon [23]. However the spirit in his work is quite different from this one. Damon defines a class of admissible deformations of images, and then in turn finds events or structures that may not be removed by the admissible deformations. Finally he attributes normal forms to these events. In this way, he defines equivalence classes of images. These classes in turn define what is the image structure. Our approach is different, since we define what is image structure directly in terms of feature detectors. We believe that observables (features) constitute the key to image analysis. The image analyst defines the feature *a priori* and may then in turn ask the question: "Which structure may I expect that the features exhibit?". This paradigm corresponds to what is normally adopted in analysis of shape singularities. Here the features such as parabolic lines, umbilici, etc.

are pre-defined and their generic relations and evolutions may then in turn be derived [102, 17].

1.2.1 Caustics and shocks

Another branch in image analysis using results from singularity theory is focused on caustics and shock formation. This kind of partial differential equation was the first in mathematics to be analysed for generic behaviour [116, 42, 4, 5, 8, 7]. In image analysis these kind of evolutions are often used in conjunction with shape representation, for instance curvature evolution (Euclidean shortening flow), grass fire evolution, medial axis and skeletons.

One of the latest efforts to bring this into image analysis is by Giblin and Kimia [39] based on previous work by Bruce and Giblin [16, 15].

This thesis will not explore this direction in the field. But the reader is referred to the above mentioned literature and references therein.

1.3 Outline of the rest of the thesis

This thesis consists in part of published articles written during my Ph.D. study. The articles are written in collaboration with others therefore I have chosen to include the articles in the original form and only altered the layout for a more homogeneous presentation including collecting the references in one common list plus correcting some errors according to the defence committee. This implies that most chapters can be read without reading other chapters first. The drawback is repeated information in different chapters for instance several similar short introductions to scale space theory.

This thesis is organised in two major parts introduced by this general introduction. Each part has its own introduction where a short summary of each chapter is given.

The first part discusses what generic structure is and how to establish it. Examples are given for the generic structure of the gradient squared and corner measures in Gaussian scale space as well as the generic structure of the optic flow.

The second part presents a multi scale segmentation tool based on knowledge of the deep structure of the image feature, the gradient squared. The application is presented all the way from theory to a semi-automatic segmentation tool in clinical use as well as to a prototype for automatic segmentation.

Two appendices elaborate on derivations in the chapters 6 and 8, respectively.

Part I

Genericity

Chapter 2

Genericity: Introduction

This part will give a framework for and examples on establishing generic properties for image features expressed by local differential expressions. A vast amount of methods exist for extracting image features, here we will mainly develop a framework for analysing features defined by semi algebraic equations of image derivatives.

First an introduction to singularity theory is presented followed by a framework for using this on image structure. Examples on the use of the framework on the gradient magnitude and corner detectors are then presented. The end of this part is devoted to presenting theory and an example on generic structure via equivalence. The major example used for illustration is the structure of the optic flow.

The first category of examples applies this to geometric properties represented by varieties in jet space (loosely speaking the space of Taylor coefficients). Since the library of image features is indefinite even in our restricted view, we aim for constructing an algorithm for analysing the generic structure imposed by any new differential feature detector. The illustrating features are the singularities for the gradient magnitude usually applied in segmentation and in junction detection, and the isophote curvature times the gradient magnitude often applied as corner measures.

Furthermore we focus on the Gaussian scale-space, so that multi-local features may be expressed as features at a finite scale. The features of an image change as a function of the free scale parameter, and this total scale-space behaviour of features in general describes the image graph to a very high degree. For instance, it has been proven that one may reconstruct the total image from the multi-scale behaviour alone from the multi-scale zero-crossings of the Laplacian (in the 2D case) [51] or alone from the scale-space top-points in the band-limited one dimensional case [58, 57].

Changing the focus from the general class of infinitely differentiable func-

tions to a subset with members which are solutions to a specific partial differential equation changes the premises for the general results of Thom [127]. Fortunately Damon [23, 25] has proved similar results for the solution space to certain partial differential equations including the heat equation.

The last two chapters 7 and 8 in this part take a different approach by defining structure via equivalence under deformation. It is applied to the structure of the optic flow.

In the first approach the task is to establish whether or not a variety in jet space can be stratified into smooth submanifolds and then apply transversality arguments. In the second approach the task is to establish that equivalence classes of functions map to stratifiable subsets of jet space. The latter task demands a lot more mathematics than the former.

A very short and straightforward introduction text to catastrophe theory is written by Saunders [115] including illustrations of the geometry. A short introduction and an array of nice applications is presented by Arnold [7]. For a more and more thorough introduction the reader is referred to Gilmore [41] and Poston & Stewart [103]. The book by Bruce and Giblin [17] is also highly recommended for thorough presentation of singularity theory relevant for the analysis of curves. If the focus is on generic properties for partial differential equations one could start with papers by Damon and Rieger [24, 106, 107, 108]. For the mathematically inclined the following literature will be worth a look: Mather [77, 78, 76, 79, 80, 81] has made a very important series of papers , also a book by Gibson, Wirthmüller, du Plessis and Looijenga [40] gives valuable insight to the subject. In the PDE direction one starting point are the papers by Damon [23, 25]. The above is by no means a complete list but will provide a good starting point.

A summary of the chapters in this part are presented in the following sections.

2.1 Singularity theory I

The first chapter on singularity theory (chapter 3) presents properties of functions as smooth manifolds in the jet space. For a given coordinate system, jet space is the space of Taylor coefficients equivalent with Euclidean space. The "property manifold" can be represented as a variety which can be split into disjoint smooth manifolds (this splitting process is called stratification). A function is represented by a map (jet extension map) from its domain to jet space. The graph of the jet extension map is a manifold in jet space. Transversal intersection between the property manifold and the function graph manifold is generic. Transversal in the previous sentence implies

that the property/feature will be a smooth manifold in the function domain with the same codimension as the "property manifold"; if the codimension is larger than the dimension of the domain the feature will not be present. Codimension is the dimension of the embedding space minus the dimension of the manifold. The generic part of the statement implies (on a compact domain) that the well behaved behaviour will occur for an open and dense subset in the function space of interest.

2.2 Generic events for the gradient squared

Chapter 4 focuses on minima for the gradient squared for an image embedded in a linear scale space. The main focus on minima is driven by the direct application to segmentation; the knowledge derived in this chapter is utilised in part two of this thesis. The main results are the presence of two generic events. The first one is the expected one where two singularities appear or disappear as a pair. If one counts the number of positive principal curvatures at the two singularities respectively before/after the disappearing/appearing then the count will differ by one. The less obvious event is the event of two global minima joining a saddle point and becoming one minimum. The reverse process is also generic going from one minimum into two global minima and a saddle.

2.3 Watershed junctions as feature detectors

Chapter 5 gives an analysis of the singularities of the gradient magnitude where the focus is on the maxima not the minima. The correspondence with junctions of watershed is established and their occurrence is connected to geometric properties of the images. It is shown that the watershed junction generically speaking is always part of the classical edges.

2.4 Smoothing creates corners

Chapter 6 explores corner measures of the following type: singularities of the isophote curvature times some power of the gradient magnitude where some power equals zero, one, two or three. The results with respect to generic structure of scale space images show that it is generic for each of the corner measures to exhibit both creation and annihilation events. Another direction in the comparison is changing the exponent on the gradient magnitude. The interesting point here for feature extraction is how the corners are selected as

the best ones (or if one prefers which corners are disregarded). One can order the corners for a given measure simply by the absolute value. The question is: will the different corner measures give the same ordering? The answer is no. One can get a quick feeling for this by considering the isophote curvature times some power of the gradient magnitude as members of one family continuously indexed by the exponent as parameter. Changing the parameter will cause a continuous change from one member to another including the singularities; maxima will appear and disappear together with saddles. Not only the singularity set (sometimes called the bifurcation set) but also the Maxwell set is of interest here. On the Maxwell set two or more critical values are equal. Subsets of the Maxwell set correspond to two maxima having equal value and different orderings exist on each side of the Maxwell set. Crossing this set is generic for one parameter families. Literature pointers on the Maxwell set is Gilmore [41] and Bruce et al. [16, 15].

2.5 Singularity theory II: Genericity via equivalence

The second chapter on singularity theory (chapter 7) focuses on genericity via equivalence. The usual way is to define allowed warpings of the function domain. If two functions can be put on the same form by warping one of the domains then the two functions are equivalent. Equivalent functions define an equivalence class where all members have the same qualitative properties. The restriction on the warping determines how coarse the final classification will be. The most commonly used warping is a diffeomorphism; that is a one-to-one mapping where both directions of the map are smooth. A diffeomorphism can stretch the level lines to the extreme in all directions but cannot introduce singularities. That is new sharp corners can not be removed or introduced. For instance a convex shape can be shaped to a circle but not to a point or a square.

It can be proved that these equivalence classes are mapped to smooth manifolds in jet space by the jet extension map. Next step is to use transversality arguments for proving that the equivalence classes are generic. From each equivalence class a representative can be picked usually a simple polynomial and denoted normal form. The list of normal forms gives all generic types of local appearance induced by the chosen warping.

2.6 The structure of the optic flow

In chapter 8, the optic flow field is defined as preserving the intensity along flow-lines. Due to singularities in the image at fixed time, poles are created in the optic flow field. In this chapter we describe the generic types of flow singularities and their generic interaction over time. In a general analytic flow field, normally the topology is characterised by the points where the flow vanishes again subdivided into repellers, attractors, whirls, and combinations hereof. We point out the resemblance, but also the important differences in the structure of a general analytic flow field, and the structure of the optic flow field expressed through its normal flow. Finally, we show examples of detection of these singularities and events detected from non-linear combinations of linear filter outputs. This relates to the previous chapter by defining structure by equivalence classes.

Chapter 3

Singularity theory I

We would like to know which specific properties to expect for members of a function set (images). More specific we will investigate the set of infinitely differentiable functions C^∞ and the solution set for certain partial differential equations, in particular the heat equation. Here we define a property by a set of algebraic equations of derivatives.

In general to list all expected properties for all members is an impossible mission since anything can occur if the right member in the set is picked. Instead the aim is set for the feasible goal of determining expected properties for most members.

Section 3.1 introduces a few definitions and notions within topology and differential geometry. This is followed by section on typicality of regular values k and with an explanation on why $f^{-1}(k)$ is a smooth manifold. Section 3.2 is meant as an example on typicality warming up to generic properties for function classes. The next section 3.3 introduces transversality , a natural extension of regular values. In section 3.4 the notion of jet space (with fixed coordinates it is the space of Taylor coefficients) is presented. In this space, properties and functions can be represented as submanifolds. Transversal intersection of these manifolds implies that the properties express themselves as smooth manifolds in the domain for the function. This is followed by section 3.5 which presents theorems showing transversal intersection to be generic. We define properties/features by an algebraic set. Section 3.6 describes how an algebraic set can be broken into smooth manifolds (stratified) and subsequently each submanifold fits into the notions and theorems presented in the previous sections.

3.1 Preliminaries

This section is meant as a reminder and not as a complete introduction to the subjects presented.

3.1.1 Topological notions

In the following we assume given a topological space which is a set with an assigned topology. The subset A in the following is part of a topological space.

A topology on a non-empty set M is a system G of subsets of M with the following properties:

- $\emptyset, M \in G$
- If $g_1, \ldots, g_n$ are finitely many sets from G then their intersection $g_1 \cap \ldots \cap g_n$ is in G.
- If $(g_i)_{i \in I}$ is an arbitrary family of sets from G then their union $\cup_{i \in I} g_i$ is in G

Given a metric the neighbourhood $N_\epsilon(a)$ of a point a is defined as all member within the distance ϵ. We will define a metric in definition 13 to be used on the function space. Using a metric, the topological notion open can be defined as below. The system of open sets defines a topology.

Definition 1 (Open set)
A subset A is open if and only if $\forall a \in A \;\; \exists \epsilon > 0 \;\; N_\epsilon(a) \subseteq A$

Open implies that a sufficiently small continuous perturbation of a member will yield a new member also in the subset. If members with a specific property form an open set, then we call the property stable because a sufficiently small perturbation yields a new member with the same property.

For example, the interior of a simple closed surface in Euclidean space will be an open subset. For instance, the interior of a circle or a square in the plane will be an open subset.

Definition 2 (Dense set)
A subset A is dense in a base set B if and only if $\forall b \in B \;\; \forall \epsilon > 0 \;\; N_\epsilon(b) \cap A \neq \emptyset$

Dense implies that any member in the base set can be transformed by an arbitrary small transformation into a member of the dense subset. Note that denseness always refers to a base set. A set has to be dense compared to some other set. The closure of a dense set is the base set. For example, the rational

numbers is a dense but not an open subset of the reals. A dense subset has a member almost every where or at least in the immediate neighbourhood. A dense and open subset implies that almost all members are in the subset.

When combining several subsets it is interesting to know if the result of the operation possesses the same properties as the operands. It is called closed under this or that operation. A set of open and dense sets is closed under the intersection of a finite number of sets but not under the intersection of a countable infinite collection of open and dense sets [24]. The latter set is called residual.

Definition 3 (Residual set[24])
The intersection of a countable infinite collection of open and dense sets is called residual.

The intersection of a countable infinite collection of residual sets is obviously residual. Hence the set of residual sets is closed under this operation. A residual set is dense but not necessarily open [24].

Open and dense are the properties we hope for when considering typicality. Openness insures stability of the property and denseness gives that arbitrarily close is a member with the property, in this sense the property is typical. Often openness and denseness can be achieved but in the most general cases the given property only holds for a residual set.

The solution space to some particular partial differential equation (for instance the heat equation) is denoted H.

Definition 4 (Genericity [24])
A local property of functions in $C^\infty(U)$ (respectively $H(U)$) is generic if:

- *the set of functions having the property at each point of U is residual; and*
- *for any compact subset $C \subset U$, there is an open and dense set of functions in $C^\infty(U)$ (respectively $H(U)$) which have the property at each point of C.*

The assigned topology is the regular C^∞-topology. See definition 14

3.1.2 Manifolds

A manifold is a set which locally "looks" like the Euclidean space.

Definition 5 (Differentiable Manifold)
A differentiable manifold M is a set M together with a differentiable manifold structure on it.

A differentiable manifold structure of class C^k is given on the set M if a family A (often called an atlas) of charts f_i is prescribed. $A = \{(f_i, O_i)\}_{i \in I}$ where I is an index set. The following must be true:

1. *$\forall i \in I$: O_i is an open subset of $\mathbb{R}^n$ and $f_i(O_i)$ is an open subset of M.*
2. *$\forall i \in I$: f_i is a homeomorphism on O_i.*
3. *$M = \cup_{i \in I} f_i(O_i)$*
4. *$\forall i, j \in I : \; f_i^{-1} f_j \; : \; f_j^{-1}(f_i(O_i) \cap f_j(O_j)) \to f_i^{-1}(f_i(O_i) \cap f_j(O_j))$ is a C^k mapping*

If $0 < k \leq \infty$ in definition 5 then we call the manifold differentiable. If $k = 0$ then the manifold is called a topological manifold. If $k = \infty$ we will refer to the manifold as a smooth manifold. If the mapping is not only smooth but also analytic[1] the manifold is called analytic. We will almost always consider smooth manifolds.

3.2 Typical solutions

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a smooth function and study the solution set for $f(x) = 0$. In general one would expect the set $f^{-1}(0)$ to be a smooth manifold of dimension $n - 1$ but theorem 1 by Whitney [14] tells what happens if all members are considered:

Theorem 1 (H.Whitney)
Any closed subset in $\mathbb{R}^n$ occurs as the solution set $f^{-1}(0) = \{x \in \mathbb{R}^n : f(x) = 0\}$ for some smooth function $f : \mathbb{R}^n \to \mathbb{R}$ [14]

Hence even an awful set like the Cantor set is the solution set to $f(x) = 0$ for some smooth f. But in general one can expect well behaved structure. Let us introduce some notions in order to present this result and results to come.

Definition 6 (Submersion)
Let $f : \mathbb{R}^n \mapsto \mathbb{R}^m$ be a smooth function where $\mapsto$ means the domain is an open subset of the left argument. The mapping f is a submersion at x if $Df(x)$ is surjective (the Jacobian of f has rank m which implies that $m \leq n$). $Df(x)$ is the first order derivative of f represented by a matrix.

[1] A function is analytic if its Taylor series converges to it in a neighbourhood of each point

Definition 7 (Regular Point and Value, Critical Point)
Let $f : \mathbb{R}^n \mapsto \mathbb{R}^m$ be a smooth function. If f is a submersion at x then x is a regular point. A regular value of f is a point $a \in \mathbb{R}^m$ such that every x in the domain of f with $f(x) = a$ is a regular point.

A point is called critical if the rank of the Jacobian matrix is less than $min(n, m)$. If $m \leq n$ then non-regular and critical is the same.

The following theorem by Sard describes that regular values are the common kind of values.

Theorem 2 (Sard)
Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a smooth map. Then the set of non regular values of f in $\mathbb{R}^m$ has measure zero in the sense of the Lebesgue measure. [17]

The complement of a set of measure zero is dense and open (note denseness alone or openness alone does not imply measure zero for the complement). In the following we will see that for regular values the solution set will have the well behaved structure one would expect.

Theorem 3 (The inverse function theorem.)
Let $f : \mathbb{R}^m \mapsto \mathbb{R}^m$ and $u \in$ domain(f). Assume the Jacobian matrix of f at u is nonsingular then there exists an open set $U \subset$ domain(f), such that $f : U \to f(U)$ is a diffeomorphism. f is called a local diffeomorphism at u

The implicit function theorem (a consequence of the inverse function theorem) states

Theorem 4 (The implicit function theorem)
Let $f : \mathbb{R}^{n+k} \to \mathbb{R}^k$ be a smooth map, defined on a neighbourhood of $(x, y) \in \mathbb{R}^n \times \mathbb{R}^k$, with $f(x_0, y_0) = a$. The coordinates in $\mathbb{R}^{n+k}$ are denoted x_i and y_i respectively. Consider the $k \times k$ submatrix $\frac{\partial f_j}{\partial y_i}$ of the Jacobian matrix for f. If the submatrix is nonsingular at (x_0, y_0) then there exist neighbourhoods A of x_0 and B of y_0 such that $\forall x \in A$ there is an unique point $g(x)$ in B with $f(x, g(x)) = a$ and the map $x \mapsto g(x)$ is smooth.

In other words the implicit function theorem states: If the submatrix of the Jacobian matrix is non-singular then $f^{-1}(a)$ can be parametrised using the n x's.

Now if $f : \mathbb{R}^n \to \mathbb{R}^m$ is a submersion at x with $f(x) = a$ then the Jacobian has rank m and according to the implicit function theorem the set $f^{-1}(a)$ is a parametrised $n - m$ manifold. If a is a regular value then the f is submersion for all $x \in f^{-1}(a)$. Consequently $f^{-1}(a)$ is a smooth $n - m$ manifold for all regular values a which according to Sard's theorem are the common ones.

3.3 Transversality

Transversality is an essential notion in singularity theory. It is based on the intuitively appealing fact that two randomly picked embedded manifolds only intersect if the sum of their dimensions is equal to or greater than the dimension of the embedding space. If they do intersect their tangent spaces will span the embedding space. This is true with probability one.

In a sense transversality is the natural generalisation of regular values[17]. First of all the results in section 3.2 have straightforward generalisation to manifolds by replacing the Euclidean space $\mathbb{R}^n$ and $\mathbb{R}^m$ in definition 7 and Sard's theorem 2 with smooth manifolds N and M. The next step is to go from a regular value to a "regular" manifold $f^{-1}(X), X \subset M$. The question is which conditions on the manifold X and the mapping must hold for $f^{-1}(Y)$ to be a smooth manifold. The answer is transversality and in section 3.5 we will see that transversality is generic.

Here follow a couple of examples on transversal and not transversal intersection. In figure 3.1, are plotted two surfaces which intersect transversely. At each point of the intersection set the two tangent planes have to span the embedding space to form a transversal intersection. We will check this. The two surfaces are graphs over the same domain and we use the two obvious maps:$f(x,y) = \{x, y, -x^2 - y^2\}$ and $g(x,y) = \{x, y, x^2 + y^2 - 4\}$. The intersection set is given by $2x^2 + 2y^2 - 4 = 0$. The tangents to the surfaces are:

$$[T1_{s1}T2_{s1}T1_{s2}T2_{s2}] = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ -2x & -2y & 2x & 2y \end{bmatrix} \tag{3.1}$$

This matrix has rank 3 unless $x = y = 0$. Since $x = y = 0$ is not part of the intersection set the tangents span the embedding space at all points of the intersection set. In other words a transversal intersection.

In figure 3.2 is an example on a non transversal intersection. Checking the algebra establishes this. The intersection set in the domain of the maps is $x^2 + y^3 = 0$ (a cusp shaped curve). The tangents are given by:

$$[T1_{s1}T2_{s1}T1_{s2}T2_{s2}] = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 2x & 3y^2 & -2x & 3y^2 \end{bmatrix} \tag{3.2}$$

The rank of the matrix is less than 3 if $x = 0$. Since $x = 0 = y$ is part of the intersection set we have an intersection point where the tangents for the two surfaces only spans a plane and not the entire 3-space. Consequently it is not a transversal intersection. The two surfaces have a higher order contact.

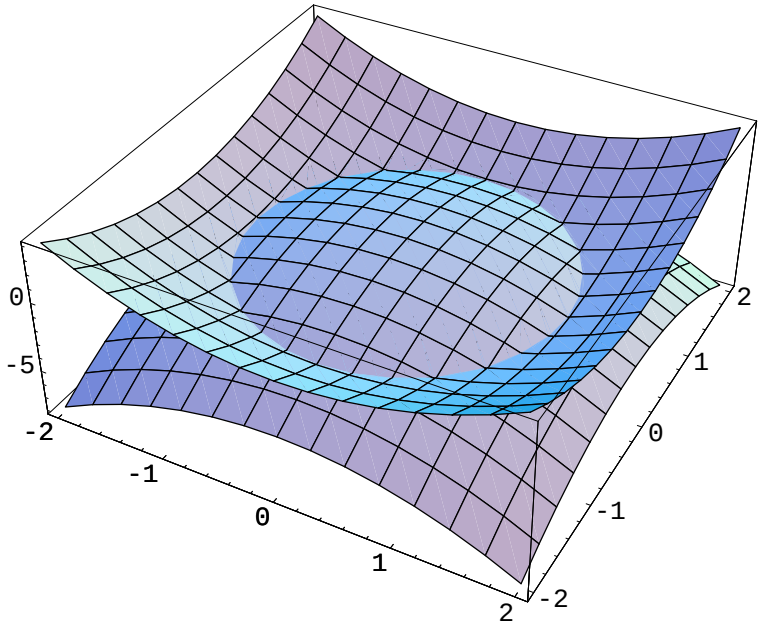


Figure 3.1: Transversal intersection of two surfaces. The defining equations are $z + x^2 + y^2 = 0$ and $z - x^2 - y^2 + 4 = 0$

Definition 8 (Transversality)

- *Transversal intersection in a point*
  Two manifolds intersect transversely or are transverse at an intersection point x_0 if the tangent spaces to the two manifolds span the embedding space.

- *Transversal intersection of manifolds*
  Two manifolds intersect transversely or are transverse if

  - *They do not intersect at all or*
  - *They intersect transversely at all points of intersection.*

- *Transversal intersection of mapping(s)*
  A mapping is said to be transverse to a manifold if the image of the mapping is transverse to the manifold. Two mappings are transverse if their images meet transversely.

Proposition 1 (Inverse image of transversal intersection)
If $f : N \to \mathbb{R}^m$ is transverse to M then $f^{-1}(M)$ is a smooth $m - dim M$ manifold.

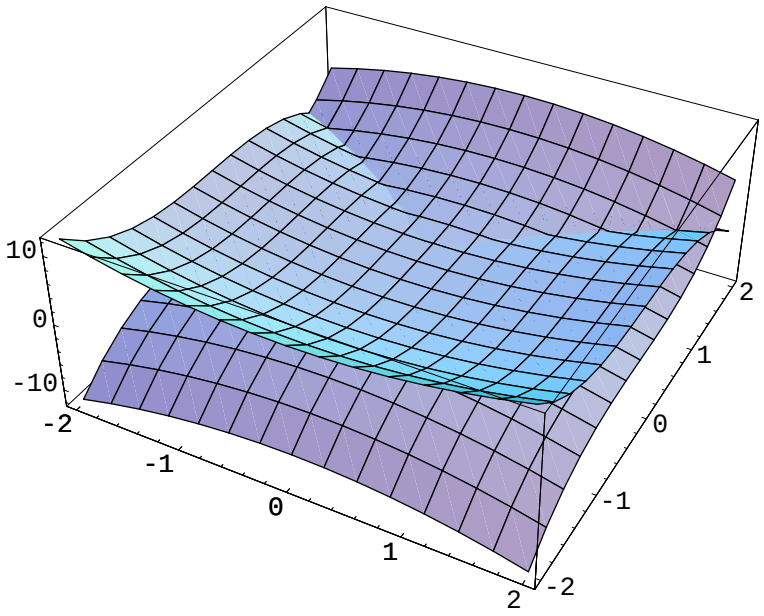


Figure 3.2: Non transversal intersection of two surfaces. The defining equations are $z - x^2 + y^3 = 0$ and $z + x^2 - y^3 = 0$

The number $(m - dimM)$ is called the codimension of the manifold M. It counts the number of constraining equations needed locally to describe the manifold. The dimension of the manifold $dimM$ describes the number of parameters required for a local parametrisation. For instance the codimension of a hyperplane is always one. The dimension is $(n - 1)$ where n is the dimension of the embedding space. Another example is a curve: The dimension is always one and the codimension is $(n - 1)$.

3.4 Jet space

In order to relate functions and properties the two types of entities are mapped into the same space referred to as a jet space. In this space, both functions and properties can be represented as smooth manifolds. Transversal intersection of the two types of manifolds insures that the property expresses itself in the function domain as smooth manifolds with the same codimension as the manifold representing the property in jet space. In the next section 3.5 we shall see that it is generic for mappings and manifolds to be transverse.

Definition 9 (Contact)
A pair of smooth maps f and g between smooth manifolds. $f, g : N \rightarrow$

$M, x \to y = f(x) = g(x)$ has k order contact at x if $|f(x) - g(x)| = O(|x|^k)$

The choice of coordinates $(x_1, \ldots, x_n)$ on N and $(y_1, \ldots, y_m)$ on M does not affect the order of contact. The contact order k at x defines an equivalence relation $\sim_k$.

Definition 10 (k-jet)
The k-jet of a C^∞ map f at x is the equivalence class:

$$j^k f(x) = \{g \in C^\infty(N, M) : g \sim_k f \quad at \quad x\} \tag{3.3}$$

Definition 11 (Jet space $J^k(N, M)$)
The jet space

$$J^k(N, M)_{x,y}$$

is the set of all k-jets for smooth $f : N \to M, y = f(x)$ and let

$$J^k(N, M) = \cup_{(x,y)\in N\times M} J^k(N, M)_{x,y}$$

If coordinates are given on N and M then the jet space is the space of Taylor polynomials of degree $\leq k$. We can now establish a k-jet extension map which maps from the function domain to jet space; simply by mapping each point to the corresponding k-jet at that point.

Definition 12 (k-jet extension map)

$$j^k f : N \to J^k(N, M), \quad x \to j^k f(x) \tag{3.4}$$

The k-jet extension maps to a "function" manifold in jet space. Properties can be represented by conditions given in terms of smooth manifolds in jet space. Now we can bring transversality to play between these manifolds. According to proposition 1 the following holds: if $j^k(f)$ is transverse to W then $j^k(f)^{-1}(W)$ is a smooth manifold with the same codimension as W.

3.5 Transversality theorems

In this section we will see that the transversal intersection between a given smooth manifold and the jet extension mapping is a generic property.

Results on genericity can only be stated with reference to a topology on the function space of interest because the definition 4 refers to topological notions such as open and dense subsets of the function space.

A distance measure on the space of C^k differentiable functions is easily obtained via the well known measures on Euclidean spaces. Here we use the max norm:

Definition 13 (The max norm)
Let $f, g \in C^k(C)$ be defined on a compact set C. A distance measure is given by:

$$d_k(f,g) = max|D^\alpha(f)(x) - D^\alpha(g)|\forall|\alpha| \leq k, \forall x \in C$$

where α is multi index and $|\alpha| = \sum_{i=1}^{n} \alpha_i$.
For example if $\alpha = \{1,0,3\}$ then $D^\alpha(f) = \frac{\partial^4 f}{\partial x_1 \partial x_3{}^3}$.

The max norm on C^k for a fixed k and one particular compact set does not give a distance measure on C^∞ but it can be use to define a topology on C^∞.

Definition 14 (Regular C^∞ topology)
A topology is made from the open sets for all possible C and k in definition 13

Using this topology the residual sets are dense [23]. The essence of Thom's lemma on transversality is that it is a generic property for a mapping to have a jet extension map which is transverse to a given submanifold:

Theorem 5 (Thom's transversality lemma)
Let $P \subset J^k(N,M)$ be a closed smooth submanifold then

$$S = \{f \in C^\infty(N,M) : j^k f \text{ is transverse to } P\} \tag{3.5}$$

is open and dense. [127]

This lemma imply that the strong statement of genericity can be established merely by stating a property in terms of smooth submanifolds and determine their codimensions.

3.5.1 Transversality theorem for the heat equation

When embedding the image in a family of images via an evolution equation one would like statements of genericity based on this new base set. In other words, one is more interested in generic phenomena within a subspace of solutions to the PDE of interest rather than the space of all smooth functions.

One cannot use the theorems and results presented by Thom. The reason for this is that the open and dense subsets of the space of all smooth functions g, for which the generic features occur, might fail to intersect the subspace H of smooth solutions g_t of some diffusion equation in dense open sets.

Damon's transversality theorem for the heat equation makes it possible to establish genericity results in the standard way but the generic properties are generic in the subspace of solutions to the heat equation. The regular C^∞ topology insures that residual subsets of $H(U)$ are dense.

Definition 15 ($H(U)$ and $H^k(U)$)

$$H(U) = \{f \in C^\infty(U) : f \text{ is a solution to the heat equation on } U \subset \mathbb{R}^n \times \mathbb{R}_{++}\} \tag{3.6}$$

$$H^k(U) = \{k\text{-jets of solution to the heat equation}\} \tag{3.7}$$

On an open subset $U \subset \mathbb{R}^n \times \mathbb{R}_{++}$ the transverse mappings form a residual set in $H(U)$. Here $\mathbb{R}_{++}$ is the strictly positive reals.

Theorem 6 (Damon's Transversality Theorem[23, 107])
Suppose W is a smooth submanifold of $H^l(U)$, then the set of mappings $f \in H$ that are transverse to W (i.e. for which $j^l(f)$ is transverse to W) form a residual set for the regular C^k-topology $(l+1 \leq k \leq \infty)$.

If the domain is compact the result can be strengthened such that the mappings form an open and dense set in $H(U)$.

Theorem 7 (Damon's Transversality Theorem[24, 25])
Suppose W is a smooth submanifold of $H^l(U)$, then the set of mappings $f \in H(U)$ for which $j^l(f)$ is complete transverse to W on a compact subset $C \subset U$ is an open and dense set for the regular C^∞- topology.

If a mapping is transverse to the closure of a manifold we say that the mapping is completely transverse to the manifold. See Damon [25] for further details.

Using these theorems the normal procedure for establishing genericity can be followed by representing properties in terms of smooth manifolds and determine their codimension.

3.5.2 Transversality theorem for Nonlinear PDE's

Damon's result also applies to a certain class of non-linear differential operators which satisfy certain "filtration conditions" and whose initial value problem has at least one local C^∞-solution [108, 25].

In theorem 6 the space $H(U)$ is simply replaced with the solution space to a specific partial differential equation.

Damon [25] and Rieger [107] have given two examples on these non linear cases:

$$P_1(g) = \frac{\partial g}{\partial t} - div(r(|\nabla g|^2)\nabla g) = 0, g_0 \in C^\infty(U, \mathbb{R}) \tag{3.8}$$

where r is a globally positive analytic function. This type of PDE was first suggested by Perona and Malik [101].

The new partial differential equation investigated by Rieger was :

$$P_2(g) = \frac{\partial g}{\partial t} - r(|\nabla G_s * g|^2) \times (\triangle g - \frac{h(|\nabla g|^2)}{|\nabla g|^2} < H_g(\nabla g), \nabla g >) = 0 \quad (3.9)$$

where $g_0 \in W_2^2(U, \mathbb{R})$ and $W_2^2 \in L_2$ having generalised derivatives up to order 2, r is a global positive C^∞ function and $h(v) = \frac{(1-\epsilon)(v-c)}{l(v-c)+l(2c-v)}, c > 0, \epsilon \geq 0$ where for $v > 0 : l(v) = e^{-\frac{1}{v}}$ and for $v \leq 0 : l(v) = 0$.

A special case of P_2 corresponds to the modification of the Perona-Malik scheme made by Alvarez et al. [2]. See Rieger [107] for further details. Rieger [107] suggested further that the result will hold for most types of autonomous evolution equations of the type in Equation 3.10 whose initial value problem has smooth local solutions. The main problem according to Rieger is proving smooth local solutions to the initial value problem.

$$\frac{\partial g}{\partial t} = Q(x, y, D^\alpha g; |\alpha| \leq d) \quad (3.10)$$

where Q is a non-linear function of the image coordinates and the spatial derivatives of the image g.

This thesis will not go further in this direction. We will just state here that the generic properties for the heat equation examined in the rest of the thesis also hold for at least the two above mentioned partial differential equations P_1 and P_2. It is an open question to what extent it holds for equation 3.10.

3.6 Stratification

A set of algebraic equations on derivatives defines a variety [2] in jet space. Since we only have results on how to handle properties given as smooth manifolds we need a way to break a variety into smooth manifolds. This task is called stratification.

A stratification of a subset V of a smooth manifold M is a partition P of V into disjoint smooth manifolds of M, each one called a stratum. A good example to keep in mind is an algebraic set $V \subset \mathbb{R}^m$ defined by finitely many polynomial equalities. That is a common way to define a feature detector.

[2] A Variety is the geometric object defined by the set of all solutions of a system of equations $f_i(x_1, \ldots, x_n) = 0$. See Cox et al.[20] for further details.

Example 1 (Stratification of $ab - c^2 = 0$)
Let a variety be defined by $ab - c^2 = 0$ in a three dimensional Euclidean space. The variety is plotted in Figure 3.3. One stratum (the surface) has codimension one and the second stratum (the intersection point at origin) has codimension 3. Two smooth maps for the surface could be:

$$\begin{aligned} f(x,y) &= \{x - (x^2+y^2)^{1/2}, -x - (x^2+y^2)^{1/2}, y\}, (x,y) \in \mathbb{R}^2 \setminus (0,0) \\ f(x,y) &= \{x + (x^2+y^2)^{1/2}, -x + (x^2+y^2)^{1/2}, y\}, (x,y) \in \mathbb{R}^2 \setminus (0,0) \end{aligned}$$

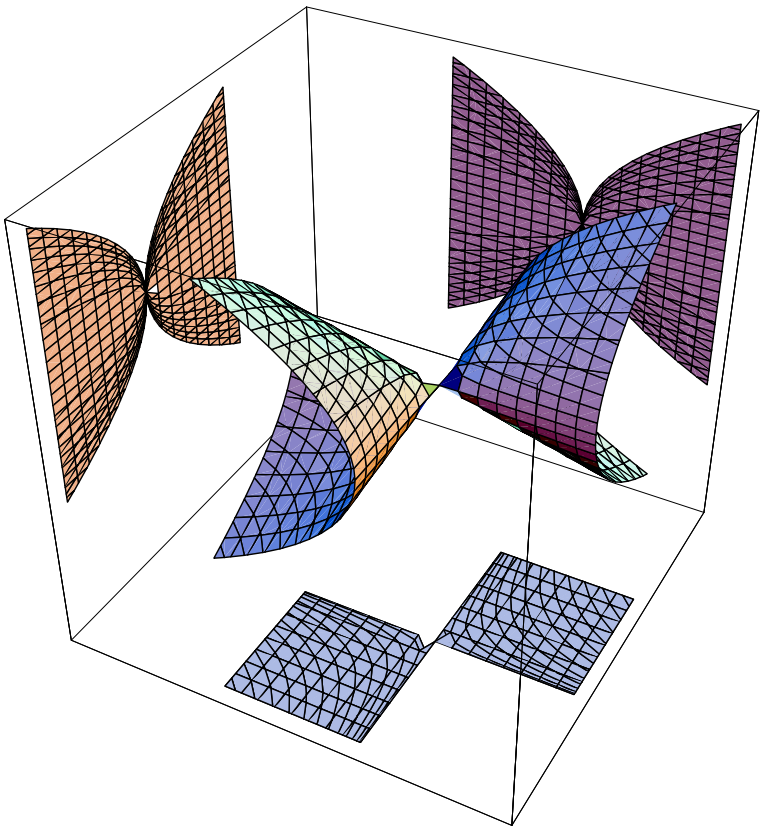


Figure 3.3: The surface given implicit by $ab - c^2 = 0$. A stratification is given by the origin being one stratum and the rest of the surface another stratum

Definition 16 (Stratification[40])
Let X be a closed subset of a smooth manifold M. A Stratification of X is defined as a partition P of X into subsets called strata and denoted X_α satisfying the following conditions:

1. *Each stratum X_α of X is a smooth submanifold of M*

2. *The family P is locally finite*

Here locally finite means that for each point in X there exists a sufficiently small neighbourhood such that the number of subsets in P in the neighbourhood is finite.

With the right kind of stratification it is possible to determine the codimension for each stratum separately and draw conclusions concerning genericity etc. Let us take a simple example and see what can go wrong.

Example 2 (Stratification of $xy = 0$)
One stratum could be $x = 0$, that is the y axis. The second stratum could be $x \neq 0, y = 0$ which consists of the x axis except $x = 0$. Of course this covers all points of the variety and the two strata are disjoint smooth manifolds. Each of the strata has codimension 1. This line of reasoning could give the misleading impression that the whole variety would give structures of codimension 1 but the conditions on the stratification are not strict enough to draw that conclusion.

One improvement could be to avoid having singular and non-singular point on the same stratum. This would handle the problem for the variety defined by $xy = 0$. Let $V \subset \mathbb{R}^m$ be an algebraic set. The set SV of singularities for V is also an algebraic set of strictly lower dimension than V. The difference $V \backslash SV$ is a smooth manifold[40]. Next one can find the difference set between SV and the singularity set SSV for SV. The construction of difference sets will terminate since the dimension strictly decreases for each new difference set and the set of difference set will be a stratification of V.

Example 3 (Stratification of Whitney Umbrella)
The Whitney Umbrella is defined by $x^2 = zy^2$. Using the mentioned partition method (where a set of smooth manifolds is constructed by difference between an algebraic set and its singular set) the surface will be stratified into two strata. One stratum being the z-axis and the other stratum being the surface with the following map $z = x^2y^{-2}, y \neq 0$. A visualisation is given in Figure 3.4.

Even with this construction there is still an unpleasant property for such a stratification namely on the same stratum the variety can have changes in topology. In the above stratification of the Whitney umbrella the variety has three different topologies on the stratum consisting of the z-axis. For $z < 0$ the topology is a line , for $z > 0$ the topology is two surfaces intersecting transversely, and for $z = 0$ the local topological type changes.

The solution is to strengthen the demands on the stratification with conditions first given by Whitney [138, 139]

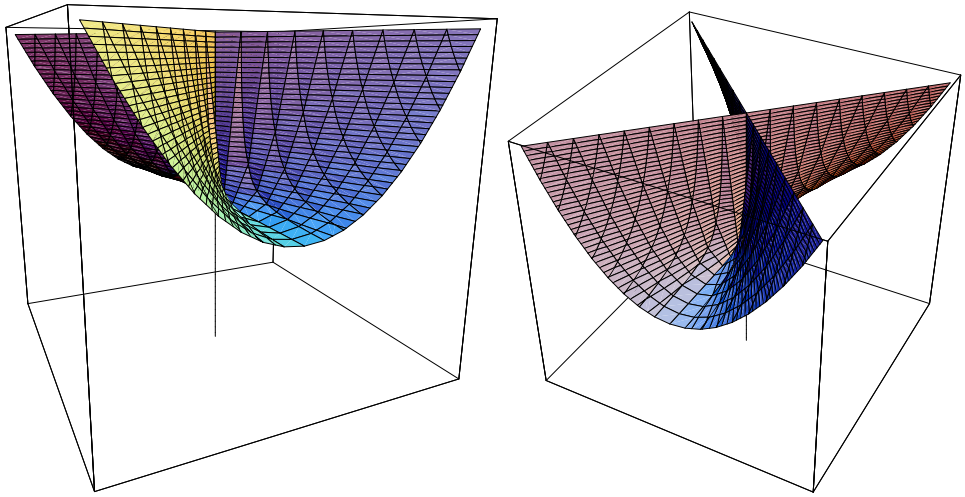


Figure 3.4: Two views of the variety defined implicit by $x^2 = zy^2$ usually referred to as Whitney's Umbrella.

Definition 17 (Whitney stratification[40, 82])
A Whitney Stratification is a stratification as in definition 16 which fulfils the following extra conditions:

1. The condition of the frontier*: if* $X_\beta \cup \bar{X_\alpha} \neq \emptyset$*, then* $X_\beta \subset \bar{X_\alpha}$ *where* $\bar{X_\alpha}$ *denotes the closure of* X_α*.*

2. Regularity condition*: If* X_α *and* X_β *are strata and* $X_\beta \subset \bar{X_\alpha}$*, then* X_α *is Whitney regular over* X_β

A Whitney stratification of Whitney's Umbrella is given by the following four strata: the surface consisting of two connected components and the z axis split into three namely $z > 0$, $z < 0$ and $z = 0$. This Whitney stratification has the property that all points on the same stratum have the same local topological type.

The regularity condition gives relations for $X_\beta \subset \bar{X_\alpha}$ between limiting tangent planes from X_α, limiting secant lines, and tangent planes of X_β:

Definition 18 (Whitney regular[40])
Let X and Y be smooth submanifolds of a smooth manifold M and let $x \in X$. Y is Whitney regular over X at x if the following holds.

If $(x_i), (y_i)$ are sequences in X, Y $(x_i \neq y_i)$ both converging to x,

and if the sequence of tangent space $(T_{y_i}Y)$ converges to a subspace T

and if the sequence $(x_i\vec{y}_i)$ of lines containing $x_i - y_i$ converges to a line L

then $L \subset T$.

Chapter 4

Generic events for the gradient squared [1]

4.1 Introduction

The very introduction of the scale-space formalism [140, 66] emphasised the importance of relating structures at different scales, the deep structure[2]. We can quote Koenderink [66]:

> Study the family as a family, i.e. define deep structure, the relations between structural features of different derived images.

When an image is embedded into a family governed by the standard linear diffusion equation, transitions between objects of qualitative different appearance will take place. It has been argued that the understanding of these changes is an essential key to the understanding of shape [67].
The deep structure for several image descriptors like image extrema, blobs, edges and ridges has been analysed[72, 55, 46, 24, 107]. This article follows this line of work by revealing the generic deep structure of singularities for the gradient magnitude squared $L_i L_i$ in Gaussian scale-space.
Due to a duality between catchment basins (basins of attraction) and the minima of the function of interest, it is possible to transfer the derived deep structure to the catchment basins. Since segments can be defined as catchment basins (for which watersheds are boundaries), the deep structure of a multi-scale watershed segmentation is established with this analysis. The deep structure of blobs[72] reveals a similar property where the deep structure of singularities can be transfered to the structural changes of blobs with

[1] This chapter has been published in a conference proceedings [97]

[2] In this chapter, we used the term "deep structure" as first introduced by Koenderink. It is equivalent to the term "multi-scale structure"

varying scale. The multi-scale structure of alternative ridge definitions has been analysed [107, 27]. These results are however not directly applicable to ridges in a non-linear expression (such as the gradient magnitude).
The choice of watersheds/catchment basins for segmentation is somewhat arbitrary, but it facilitates a mathematical clean formulation by the above mentioned analysis, and single scale watershed segmentation has been reported [90] to give good intuitive results although with a problem of over-segmentation. The problem of over-segmentation has been approached in a number of ways. A common solution in mathematical morphology[90] is flooding from special interest points. Another approach is fine scale smoothing followed by grouping of neighbouring regions according to a Minimum Description Length principle [75]. In the multi-scale formulation, the over-segmentation is greatly reduced merely by detecting the interesting objects at a coarse scale. The price is a dislocation of structure caused by the blurring. This can be counteracted by detection at coarse scale and localisation at fine scale. In order to do this, the structures must be linked across scale. This is a well-known approach for feature detection in scale-space, and in particularly for multi-scale segmentation [66, 133, 71, 72]. The region-based definition of segments allows a robust region-based detection and linking, and avoids the difficult problem[71] of linking point entities.
The diffusion scheme and notation is presented in the next section. In Section 4.3 the generic catastrophe events for the gradient squared have been derived. Watersheds, catchment basins and segmentation based on this are defined in Section 4.4. Based on the catastrophe analysis and the segment definition, Section 4.5 presents the segment linking scheme. Verified segmentation results based on the defined multi-scale segmentation are put forward in Section 4.6. Finally, in Section 4.7 we summarise and suggest further studies.

4.2 Scale-space

Definition 19 (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 $t = \sigma^2/2$ its spread.

The scale-space formalism facilitates well-posed differentiation because the Gaussian can be differentiated prior to convolution. We use $L_{i_1 \ldots i_k}(x,t)$ for spatial derivatives of L with respect to the variables $i_1 \ldots i_k$. The scale space

L follows the diffusion equation $\partial_t L = \triangle L$, as does any *linear* combination of image derivatives, e.g. the Laplacian. The scale-space image $L(\cdot, t)$ is infinitely differentiable ($\in C^\infty$). Morse functions form an open and dense set in the C^∞ function space, and in this sense Morse functions are typical (formally generic) and highly suitable as image models[72, 46, 24].

4.3 Generic events for the gradient squared

In this section, we determine how interacting critical points for the gradient squared, $L_i L_i$, create an abrupt change of structure when the scale parameter varies smoothly; this change is referred to as a catastrophe. Algebraically a catastrophe corresponds to a degeneracy in the Hessian matrix ($H_{ij} = L_{ij}$) of the studied functional. The degree of degeneracy in the Hessian determines the number of degenerate directions. The central *splitting lemma* [41, 103] states that non-degenerate directions can be safely ignored. In the following we first establish a local model for $L_i L_i$; secondly the Hessian for this local functional is analysed in order to determine the generic number of degenerate directions and finally we end the section by analysing possible generic catastrophes.

The local model and degenerate directions

For simplicity we study the square of the gradient magnitude $L_i L_i$, but the derived results are applicable to any strict monotonic transformation of the gradient magnitude as well. That is, the local topology is preserved by a strict monotonic transformation. The result holds in general for the gradient squared of an entity following the standard linear diffusion equation. In 1D, this is every image derivative.

Our local image model is an expansion at $(\bar{x}; t) = (0; 0)$ to fourth order in heat polynomials[3] (polynomials obeying the diffusion equation $\triangle f = \partial_t f$ [57, 33, 31]) using a Cartesian coordinate system:

$$\begin{aligned} \tilde{L}(\bar{x}; t) &= L + L_i x_i + \frac{1}{2} L_{ij} x_i x_j + L_{ii} t + \frac{1}{6} L_{ijk} x_i x_j x_k \\ &\quad + L_{ijj} x_i t + \frac{1}{24} L_{ijkl} x_i x_j x_k x_l + \frac{1}{2} L_{ijkk} x_i x_j t \\ &\quad + \frac{1}{2} L_{iikk} t^2 \end{aligned} \tag{4.1}$$

[3] A recursion formula for the heat polynomials in 1D is given by: $v_n(x,t) = x v_{n-1}(x,t) + 2t(n-1)v_{n-2}(x,t), v_1(x,t) = x, v_0(x,t) = 1$

Where $L_{i...l}$ are image derivatives evaluated at $(\bar{x};t) = (0;0)$, i, j, k and l are indices running over the N spatial variables and Einstein summation convention[4] applies. First order derivatives can easily be calculated:

$$\begin{aligned}\tilde{L}_m(\bar{x};t) &= L_m + L_{mj}x_j + \frac{1}{2}L_{mjk}x_jx_k + L_{mjj}t \\ &\quad + \frac{1}{6}L_{mjkl}x_jx_kx_l + L_{mjkk}x_jt \end{aligned} \tag{4.2}$$

Hence the functional of interest G, the gradient squared, becomes

$$G(\bar{x};t) = \tilde{L}_m(\bar{x};t)\tilde{L}_m(\bar{x};t) \tag{4.3}$$

In the following we assume without lack of generality that a singularity for G and a possible degeneracy of the Hessian for G is located at $(\bar{x};t) = (0;0)$. This approach of expanding at the catastrophe is general applicable [41, 103] and usually result in a tremendous simplification. A singularity for G at $(\bar{x};t) = (0;0)$ is given by

$$G_{x_j}(0;0) = 2L_mL_{mj} = 0, \forall j \tag{4.4}$$

For each spatial index j there is an equation. This corresponds to an N dimensional sub-manifold of the D dimensional jet-space. The singularities can be split in two distinct types: global minima for G (i.e. critical points for L) and the rest:

$$L_m = 0, \forall m \quad \text{global minima} \tag{4.5}$$

$$L_{wj} = 0 \quad \text{the rest} \tag{4.6}$$

In both cases, there are N equations given by the free indices m or j. In the latter equation, the notation has been abused slightly. The equation is expressed in "gauge coordinates" [32, 31] , where one direction denoted w is fixed to the direction of the image gradient. Hence, Equation 4.6 expresses that all second order derivatives involving the w direction must equal zero.

The Hessian H_G for G evaluated at $(\bar{x};t) = (0;0)$ is given by

$$H_{Gjk} = 2L_{ij}L_{ik} + 2L_iL_{ijk} \tag{4.7}$$

We are interested in the generic values for the co-rank of H_G. A diagonalised Hessian H_G will have off diagonal elements equal zero:

$$2L_{ij}L_{ik} + 2L_iL_{ijk} = 0, \text{if } j \neq k \tag{4.8}$$

[4]Implicit summation over identical indices

In the coordinate system implied by Equation 4.8 the co-rank is equal to the number of eigenvalues equalling zero. The eigenvalues are given by:

$$2L_{ij}L_{ik} + 2L_iL_{ijk}, \quad \text{where } j = k \tag{4.9}$$

We can split the analysis according to the two types of singularities of G:

Assume a singularity of the type $L_i = 0$
The Hessian is diagonalised by turning to the Cartesian coordinate system where $L_{ij} = 0, i \neq j$. The eigenvalues of H_G are $2L_{ij}L_{kl}, i = j = k = l$. Hence only constraints imposed on the second order derivatives, like $L_{xx} = 0$, can degenerate the Hessian. A degenerated singularity is a generic event for a scale-space image[55], since this event coincide with this type of degenerated singularity for G, we conclude that it is a generic event for G.

Assume a singularity of the type $L_{wj} = 0$, where w is in the gradient direction and j is a free index. The remaining directions $(u, \ldots, v)$ can fixed by changing the coordinate system in such a way that mixed second order terms are zero: $L_{ij} = 0, i \neq j$. In terms of algebra, we have the following setup: a real symmetric $n \times n$ matrix A as in Equation 4.10 with a $(n-1) \times (n-1)$ submatrix B. There exists a matrix X such that $X^{-1}BX$ is a diagonal matrix D.

$$A = \left(\begin{array}{c|c} L_{ww} & 0 \\ \hline 0 & B \end{array}\right) \qquad \left(\begin{array}{c|c} L_{ww} & 0 \\ \hline 0 & D \end{array}\right) = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & X^{-1} \end{array}\right) A \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & X \end{array}\right) \tag{4.10}$$

In the constructed $(u, \ldots, v, w)$ coordinate system the Hessian is

$$H_{Gjk} = 2L_wL_{wjk} + \begin{cases} 2L_{jk}^2 & \text{if } j = k \wedge k \neq w, \text{ No summation over j and k} \\ 0 & \text{if } j = k = w \\ 0 & \text{if } j \neq k \end{cases} \tag{4.11}$$

Equation 4.11 is derived from equation 4.7 by using $L_w \neq 0 \wedge L_i = 0, i \neq w$ and $L_{ij} = 0, i \neq j$ and $L_{ww} = 0$.

Since no special constraints connect the terms of $Det(H_G)$ the determinant will not by coincidence equal zero. Only a constraint like $Det(H_G) = 0$ can raise the co-rank. This corresponds to a manifold of codimension 1 in the jet space. A manifold of higher codimension than 1 will not intersect the manifold corresponding to the function family of interest (1-parameterised jet) which has codimension $D - 1$ (a curve). Hence, generically the co-rank for the Hessian H_G is 1 for specific values of the parameter t, and never higher. The direction of degeneracy is determined by Equation 4.8. In this

coordinate system a degeneracy, say in the x direction, is given by:

$$2L_{i1}^2 + 2L_i L_{i11} = 0 \tag{4.12}$$

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 [41, 103, 72, 93].

Local structure around catastrophes

We start this section by presenting in figure 4.1 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.

$$\begin{aligned}
\breve{L}_m(x,t) &\equiv \tilde{L}_m(x,0,\ldots,0;t) \\
&= L_m + L_{mjj}t + (L_{mx} + L_{m1kk}t)x \\
&\quad + \frac{1}{2}L_{m11}x^2 + \frac{1}{6}L_{m111}x^3
\end{aligned} \tag{4.13}$$

$$\begin{aligned}
\breve{G} &= \breve{L}_m\breve{L}_m \\
&= (2L_mL_{m1} + 2L_{mjj}L_{m1}t + 2L_mL_{m1kk}t)x \\
&+ (L_{m1}^2 + L_mL_{m11} + (2L_{m1}L_{m1kk} + L_{mjj}L_{m11})t)x^2 \\
&+ (L_{m1}L_{m11} + \frac{1}{3}L_mL_{m111} + L_{m1kk}L_{m11}t \\
&\quad + \frac{1}{3}L_{mjj}L_{m111}t)x^3 \\
&+ (\frac{L_{m11}^2}{4} + \frac{L_{m1}L_{m111}}{3} + \frac{L_{m1kk}L_{m111}}{3}t)x^4 \\
&\quad + \frac{L_{m11}L_{m111}}{6}x^5 + O(x^6)
\end{aligned} \tag{4.14}$$

Again we split the analysis according to the type of the assumed singularity:

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

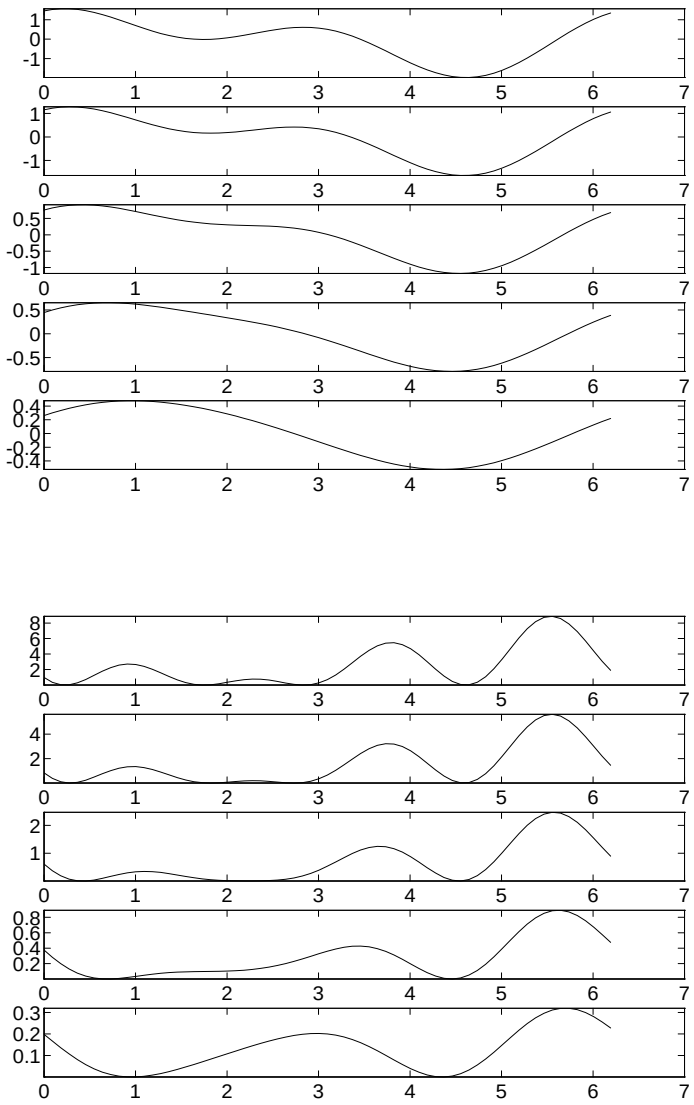


Figure 4.1: The top subfigure displays a 1D signal embedded in a family of Gaussian blurred versions of the signal. The top frame is the original signal and the degree of blurring increases downwards from frame to frame with the most blurred version presented at the bottom. From the top frame to the 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 $x = 1.4$.

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:

$$\begin{aligned} \breve{G} &= L_{mjj}L_{m11}tx^2 + (L_{m1kk}L_{m11} + \frac{1}{3}L_{mjj}L_{m111})tx^3 \\ &+ (\frac{L_{m11}^2}{4} + \frac{L_{m1kk}L_{m111}}{3}t)x^4 + O(x^5) \end{aligned} \tag{4.15}$$

For $t = 0$, all structures up to fourth order vanish for the gradient squared. This event described algebraically by Equation 4.15 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 in Equation 4.4 ($L_mL_{mj} = 0$) and the degeneracy condition in Equation 4.12 ($L_{i1}^2 + L_iL_{i11} = 0$) reduce Equation 4.14 to :

$$\begin{aligned} \breve{G} &= (2L_{mjj}L_{m1} + 2L_mL_{m1kk})tx + (2L_{m1}L_{m1kk} \\ &\quad + L_{mjj}L_{m11})tx^2 + (L_{m1}L_{m11} + \frac{1}{3}L_mL_{m111} + \\ &\quad (L_{m1kk}L_{m11} + \frac{1}{3}L_{mjj}L_{m111})t)x^3 + O(x^4) \end{aligned} \tag{4.16}$$

For $t = 0$ all structure up to third order disappear, this event is referred to as the fold catastrophe. In the 1D case, Equation 4.17 makes it clear that only annihilation of a maximum and a minimum is possible:

$$\breve{G} = L_1L_{1111}(2tx + \frac{1}{3}x^3) + O(x^4) \tag{4.17}$$

In higher dimensions, say 3D, the reverse event is possible. The number and position of the singularities can be determined from equation 4.16. 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 4.1.

Condition Catastrophe	Singularity	Type	Creating events
Fold	$L_mL_{mj} = 0$	$L_iL_i \neq 0$	$N > 1$
Cusp	$L_mL_{mj} = 0$	$L_iL_i = 0$	$N > 1$

Table 4.1: All generic catastrophes for L_iL_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.

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.

4.4 Watersheds and catchment basins

In this section we review definitions of segments based on the watersheds of a differentiable function. View the function as a topographic relief with height identified with the image intensity. Slope lines on this follow the gradient field and begin/end in critical points. Most slope lines begin in a minimum and end in a maximum, however the atypical slope lines begin or end in saddle points. In each saddle two slope lines end and two slope lines begin. These slope-lines are denoted the separatrices of the graph of the function. A subset of these form the watersheds: if the slope line begins in a saddle which is connected by slope-lines in descending direction to more than one minimum, then it is part of the watershed structure. The watersheds surround the catchment basins of the minima. Intuitively, a rain drop falling

on the topographic relief will drain to a minimum if it falls in catchment basin of the minimum. For proper mathematical definitions of the above mentioned topological entities see e.g. [93] and for illustration see Figure 4.2.

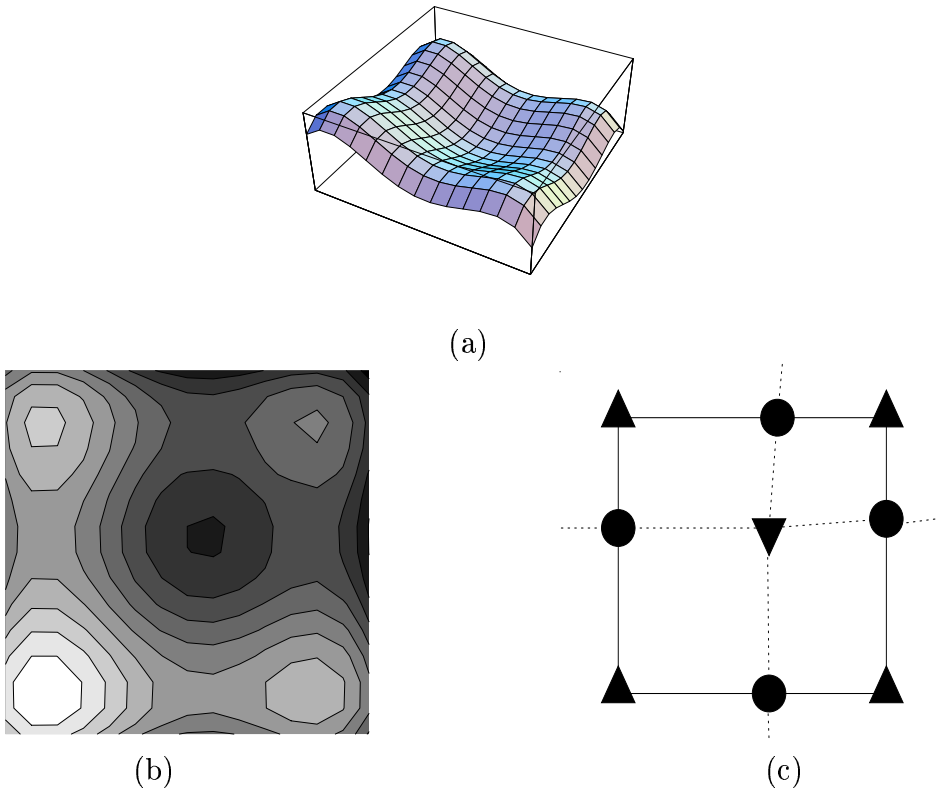


Figure 4.2: A mesh plot (a) and a contour plot (b) for a simple surface patch is illustrated. The positions of the critical points and the separatrices are projected to the domain in subfigure (c). Maxima, the minimum and saddles are indicated with respectively upward triangles, downwards triangle, and circles. The watersheds are obtained from the separatrices by removing the dotted lines.

Watersheds capture the global image structure with the implication that watersheds are not in general locally detectable. Formally stated (following Najman [90]);

Proposition 2
Let b be a point of the domain of f such that $\nabla f(b) \neq 0$. Let N_b be a neighbourhood of b which does not contain any critical point. Let γ be a path containing b and parallel to the gradient of f on N_b. Then there exists a function f_0, equal to f on N_b such that γ is in the watershed of f_0.

Hence, a gradient field patch not containing critical points does not contain

enough information for determining whether or not a watershed is present. This global probing of structure is one of reasons that watersheds handle junctions more flexible than zero-crossings. Generically any number of watersheds may meet in a point. The possible topology for zero-crossings is that $2n, n \in N$ meet in a point, and generically this will only happen for $n < k$, where k depends on the invariant expression at hand.

Another important property especially interesting for segmentation is the fact that *watersheds form closed hypersurfaces for Morse functions.* Hence, the watersheds of a function give a full partitioning of the multi-dimensional image domain, and consequently the problem of closing or connecting edges to get a partition is overcome. We suggests the general view of segmentation by watersheds on a dissimilarity measure:

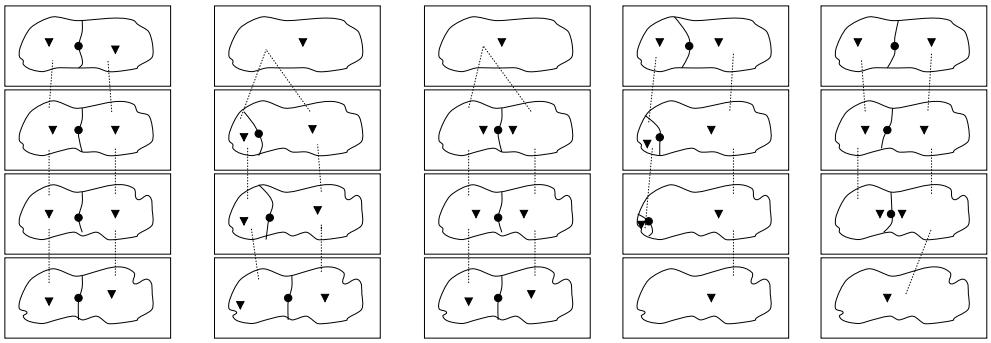
Definition Segment: A segment is the catchment basin for a minimum of a dissimilarity measure.

4.5 Linking

We have established the generic superficial and deep structure of the gradient magnitude squared. The only generic events in scale-space of the gradient magnitude image is a fold catastrophe and a cusp catastrophe involving a minimum. The duality between segments on the one hand and the minima of the gradient magnitude on the other hand suggests the linking scheme for the segments. Figure 4.3 illustrates the idea in 2D.

4.6 Segmentation Results

Without a prior knowledge of which objects are interesting (and equally important: their range of scale), it is not possible to make an automatic segmentation. Based on this observation, we have made a segmentation tool where an user interacts with the multi-scale segmentation by selecting a detection scale and a segment from the partitioning on this scale. The selected object is then localised at a lower scale (the localisation scale); the localisation scale is also selected by the user. Complex or multiple objects can be segmented by selecting or de-selecting segment roots possibly on different detection scales. The idea applies for any dimensionality and has been implemented in the two and three dimensional case. See [93] for further details. The figures 4.4, 4.5 display results on two types of 3D images: digital photos from the visible human project and a CT head scan. The tasks are to segment the liver and jaw muscles, respectively.



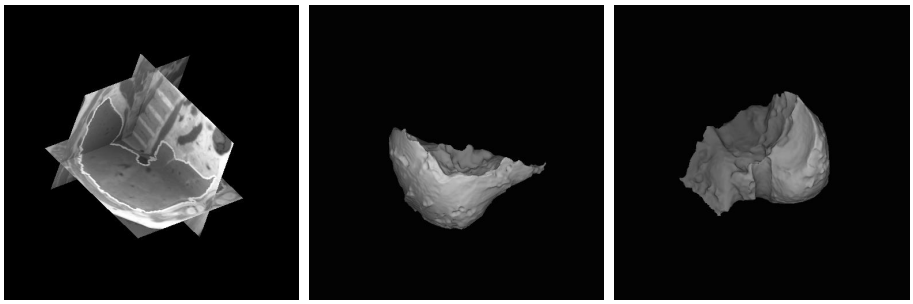
(a) No change (b) Annihilation fold (c) Merging cusp (d) Creation fold (e) Splitting cusp

Figure 4.3: Multi-scale linking of generic events in watersheds of the gradient magnitude. Scale increases upwards in the figure. Minima and saddles are symbolised with triangles and circles, respectively. The events (annihilation, merging, creation, splitting) are named after the interaction between the saddle and the minimum (or minima). 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 creation (d) and splitting (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. A line from a segment to a segment indicates a link.

4.7 Conclusion

For families of functions obtained by convolution with Gaussians of increasing width, the generic catastrophe events for the gradient magnitude squared have been established to be the fold catastrophe and the cusp catastrophe. The results hold for domains of arbitrary finite dimension. The cusp catastrophe in L_iL_i corresponds to a fold catastrophe in the image L, and consequently only occurs when global minima for L_iL_i are involved. The fold catastrophe for L_iL_i corresponds to simultaneous degeneracy in second and third order structure in the image. For one dimensional signals, the number of singularities for L_iL_i decreases strictly monotonically. For higher dimensional cases, singularities can also be created when scale increases. The result carries directly to any linear combinations of image derivatives.

The duality between minima and catchment basins has been pointed out. This duality plus the established generic events form the rationale for a multi-scale linking scheme for catchment basins on L_iL_i. Using the linking scheme, the well known single scale watershed segmentation has been embedded in a multi-scale framework. The embedding has greatly reduced the over-segmentation problem often encountered with single scale water-



(a) (b) (c)

Figure 4.4: Segmentation of the liver is challenged by two problems: similarity between liver tissue and the neighbouring muscle tissue; and the inhomogeneity of the liver tissue itself. Subfigure (a) shows the boundary of the liver segment superimposed on three orthogonal slices of the subject cube. (b) and (c) visualise surface renderings of the liver segment. The views are (relative to a standing patient): above from the right side ((a) and (b)) and above from the back (c). The view from the back (c) clearly reveals the imprint from other internal organs.

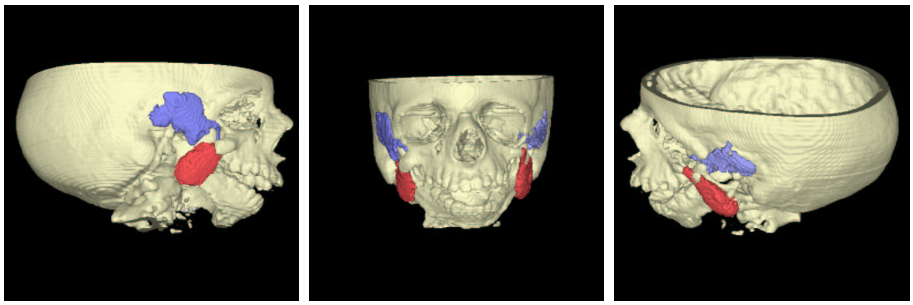


Figure 4.5: Jaw muscles segments and iso-surface cranium visualised for a patient with abnormal growth. The head scan has been provided and segmentation verified by Professor Sven Kreiborg, School of Dentistry, University of Copenhagen, Denmark. The muscular structures are located next to bone (high value), skin (low value) and salivary glands (approximately same value) which makes the task difficult for standard techniques.

shed segmentation. The interactive selection of root segments at detection scales and automatically combining the selections at a lower localisation scale provides a fast way of doing semi–automatic segmentation.

The idea can be extended to a general framework for multi–scale segmentation. The definition of segments can be changed by using another measure of dissimilarity instead of the gradient magnitude, for instance a texture measure [93]. This is possible within the same general framework although different structural changes might occur generically for other measures and diffusion schemes.

Future studies are in establishing a thorough theoretical basis for alternative dissimilarity measures and extensions to other blurring schemes that, for example, enhance elongated structures. Furthermore alternative segment definitions based on ridges can be analysed. Using locally defined boundaries sub-pixel segmentations may be easily computed.

Acknowledgements We would like to thank Peter Johansen, Lewis Griffin, Tony Lindeberg, Luc Florack, Bart ter Haar Romeny and Steve Pizer for fruitful comments and discussions. Furthermore we would like to thank Sven Kreiborg and Per Larsen in 3D-Lab, School of dentistry, University of Copenhagen, for providing the CT data and supervising the interactive segmentation.

Chapter 5

Watershed junctions as feature detectors

The idea of partitioning a landscape into hills and dales dates back at least as early as Maxwell [83]. There has been differences in the definition of watershed and watercourses [59, 113] although it seems the controversy depends on nongeneric cases (see Rieger [108] 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 a 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.

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 [88]. The stable ones are shown in Figure 5.1. One can build the whole graph of separatrices by combining these configurations.

A interesting point in watershed detection on general surfaces is that plateaus are not generic. Hence for a sufficiently fine quantification in the

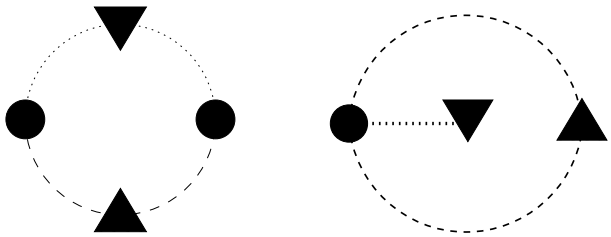


Figure 5.1: Stable configurations of separatrices and singularities.

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 implementations like the one presented by Vincent et al.[131].

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 behaviour can be determined. The watershed junctions are located at maxima of the gradient magnitude 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.

5.1 Necessary conditions for watershed junctions

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

$$G_j = L_{ij}L_i \tag{5.1}$$

$$G_{jk} = L_{ij}L_{ik} + L_{ijk}L_i \tag{5.2}$$

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

$$Det(G_{jk}) = G_{jj}G_{kk} - G_{ij}G_{ji} \tag{5.3}$$

$$= (L_{ij}L_{ij} + L_{ijj}L_i)(L_{ij}L_{ij} + L_{ijj}L_i)$$

$$- (L_{ij}L_{ik} + L_{ijk}L_i)(L_{ik}L_{ij} + L_{ikj}L_i) \tag{5.4}$$

$$G_{jj} = L_{ij}L_{ij} + L_{ijj}L_i \tag{5.5}$$

Equation 5.1 equal zero is the condition for a critical point in G. A global minimum for G coincides with a critical point for L ($\nabla L = 0$). Hereafter in this chapter we will only consider the points where $\nabla L \neq 0$. The conditions for a critical point for G describe a degeneracy in the second order structure for L in the direction of the gradient for L. Points fulfilling this condition lie on the parabolic curves for L (parabolic hypersurfaces) where the direction for the degenerated second order structure coincide with the gradient direction for L.

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

$$0 \neq L_w \tag{5.6}$$
$$0 = L_{wj} \tag{5.7}$$

Using the equations 5.7 and the equations 5.4 and 5.5 for the determinant and the trace we have in a maximum:

$$0 < L_w(L_{wvv}L_{www} - L_{wwv}^2) + L_{vv}^2 L_{www} \tag{5.8}$$
$$0 > L_{vv}^2 + (L_{wvv} + L_{www})L_w \tag{5.9}$$

Figure 5.2 (bottom right) illustrates the conditions.

Expressed in terms of G the conditions for a maximum are $G_{vw}^2 < G_{ww}G_{vv}$ and $0 > G_{ww}$. The last condition plus $L_{wj} = 0$ give us :

$$0 > G_{ww} = L_{www}L_w \tag{5.10}$$

We conclude that $0 > L_{www}$ is always true for a maximum. So the maxima will always lie on the classical edges $L_{ww} = 0 \wedge 0 > L_{www}$. Consequently watershed junctions will also always lie on the classical edges. Note that these watershed junctions are not the same as the intersection points for $L_{ww} = 0$. Rieger [108] considers the closure of $\{p \in \Omega | \nabla L \neq 0 \wedge L_{ww} = 0\}$. He shows that this set has intersection points in saddle points for L and that the set never has intersection points when $0 > L_{www}$.

In the previous we showed the connection between watershed junctions and the classical edge measure. For completeness, we will in the following describe the conditions for local minima and saddles for the gradient squared. Minima and saddle points for G can both lie on edges and not. Figure 5.2 (bottom row) illustrates the singularities for the gradient magnitude except global minima.

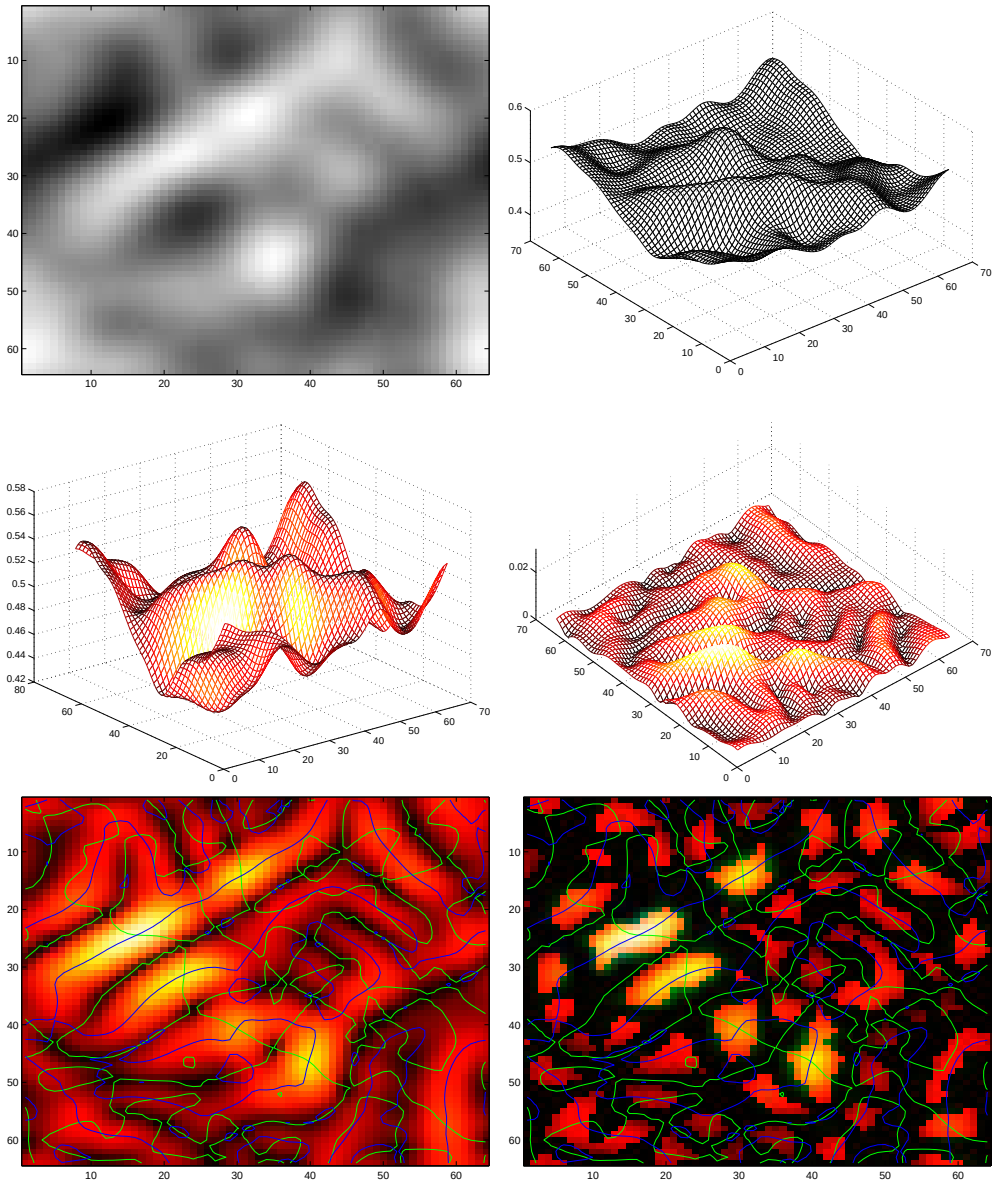


Figure 5.2: A blurred version of an image generated by random values. Top row, left: the image is depicted with intensity proportional to value. Top, right: the image is plotted as a graph over the image domain. Middle, left: the image plotted as a graph with a shading proportional to the magnitude of 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_{vv}. Bottom, right: The zerocrossing for L_{ww} and L_{vv} on the gradient magnitude with the black areas where the conditions for a maximum are not fulfilled.

In a minimum we have:

$$0 < L_w(L_{wvv}L_{www} - L_{wwv}^2) + L_{vv}^2 L_{www} \quad (5.11)$$

$$0 < L_{vv}^2 + (L_{wvv} + L_{www})L_w \quad (5.12)$$

In a saddle we have

$$0 > L_w(L_{wvv}L_{www} - L_{wwv}^2) + L_{vv}^2 L_{www} \quad (5.13)$$

Chapter 6

Smoothing images creates corners [1]

6.1 Introduction

Corner detection plays a central role in many image analysis applications ranging from character recognition to landmark identification. The literature on corner detection roughly divides into two classes. Some use explicit models [12, 105, 111, 119], while others use derivative expressions like the Gaussian curvature, the structure tensor (interest operator, second moment matrix), expressions involving the isophote curvature, and the curvature of Canny edges [85, 89, 112].

One subclass of the latter defines corners as singularities of the following expression: the isophote curvature times the absolute gradient magnitude to some power a.

$$C(x, y, a) = |\nabla L(x, y)|^a \kappa(x, y) \tag{6.1}$$

$$= L_w(w, v)^{a-1} L_{vv}(w, v), \tag{6.2}$$

where L is an image, $|\nabla L|$ is its gradient magnitude, and κ is its isophote curvature image. Further, (w, v) is a local right hand coordinate system, where w is the gradient direction of L and v the (perpendicular) tangent direction of the isophote. Subscript denotes differentiation. We only consider the measure in point $|\nabla L| \neq 0$. Kitchen and Rosenfeld [63] suggested to use $a = 1$, Zuniga and Haralick [141] proposed $a = 0$, and Blom [13], Lindeberg [73], and Alvarez and Morales [3] investigated $a = 3$.

[1] This chapter has been accepted for publication in Image and Vision Computing as "Smoothing Images Creates Corners" by Jon Sporring, Mads Nielsen, Joachim Weickert, Ole Fogh Olsen

The advantage of using a corner measure with $a > 0$ is that the product will focus on high isophote curvatures close to high contrast edges. There are two special values of a that deserve a note: $a = 0$ is invariant under a monotonic transformation of the image intensities (morphological invariance), and $a = 3$ is invariant under affine transformations (the angle of the corner).

Scale is an essential part of detecting corners. At pixel scale, differential corner measures will be dominated by the pixellation, but a simple smoothing reduces the effect of the pixel grid, while still retaining important features. We will investigate the above subclass of corner measures in Linear Scale-Space (see [136] and the references therein):

$$L(x, y, t) = G(x, y, t) * L(x, y),$$

where the original image $L(x, y)$ is convolved with a Gaussian $G(x, y, t) = \frac{1}{4\pi t} \exp(-\frac{x^2+y^2}{4t})$ of variance $2t$. The parameter t is called scale. Henceforth, when we refer to scale-space we will mean Linear Scale-Space. There exist essentially three implementations of scale-space: Spatial convolution, multiplication in the frequency domain, and direct solution of the Heat Equation $L_t = L_{xx} + L_{yy}$. We use here the frequency implementation, since it is fastest for large scales, and it easily extends to the calculation of derivatives. However, we note that for small scales the direct solution of the Heat Equation is more stable [134].

The advantage of such an embedding is that it reduces the grid and noise effects and allows for an uniform analysis of corners of all sizes or resolutions. The disadvantage is that the corners are dislocated at high scale and should be traced back to low scale in order to improve their location. Such an algorithm typically involves the following steps:

1. Produce a stack of images $L(x, y, t)$ for a range of scales.
2. For each scale and some constant $a \in \{0, 1, 2, 3\}$, calculate the corner image $C(x, y, a, t)$.
3. Detect corner points in each corner image.
4. Link corner points across scales.
5. At fine scale, choose the corner points that can be linked to coarser scale.

An essential part of the algorithm is thus the linking of corners across scales. We will show that this process – although common in the literature – is problematic, since new corners may arise as the scale is increased.

Studying the behaviour of corners across scales is conveniently done with Catastrophe Theory. We will sketch the catastrophe structure in two different settings. Firstly, we will examine the unconstrained, spatial singularity structure of $C(x, y, a, t)$. We find that new corners do generically appear for increasing scales. Secondly, we extend Rieger's analysis [106] to single isophotes, since the definition of $C(x, y, a, t)$ is intrinsically linked to isophotes, and since the corners do lie on the curves under study in contrast to Canny edges as noted by Rieger [106]. Also for single isophotes, creations of corners do occur for increasing scales.

Some authors have studied corners by first obtaining edge curves, and then analyse the curves for corners, see e.g. [85] and references therein. These studies do typically not analyse the effect of image resolution.

Related to our work in terms of Catastrophe Theory is Damon [23], Rieger [106, 107], Griffin & Colchester [46], Olsen [97], Johansen [56], and Gauch and Pizer [38].

6.2 Image structure

Corners are spatial point features, and for the differential corner measures (6.2), these points are found as the intersection between curves: $A(x, y) = 0$ and $B(x, y) = 0$. We will investigate two pairs of intersecting curves:

1. Corners defined as critical points of C, i.e. $A = C_x$ and $B = C_y$.
2. Corners defined on the isophote, L_0, as extrema of C along the isophote, i.e. $A = L - L_0$ and $B = C_v$.

We analyse the solution set satisfying $A = B = 0$ and its evolution, when L is evolved in scale-space.

In Figure 6.1(LEFT) are shown typical intersections between curves: Either two curves intersect with non-parallel tangents or their intersection is empty. This typical type of intersection is called transversal. A consequence of the transversal intersection is that closed curves given by zero-crossings of any pair of image derivatives have an even number of intersections or none at all.

The intersection points form smooth curves in scale-space. Such feature curves will have points, where the tangent is perpendicular to the t-axis. This typically corresponds to the interaction of spatial point features, i.e. feature points annihilate or are created in pairs as illustrated in Figure 6.1(RIGHT). We refer to such points as catastrophe points.

It is always possible to come up with an example where the intersection set exhibits extreme and atypical behaviour. However, in this chapter we will

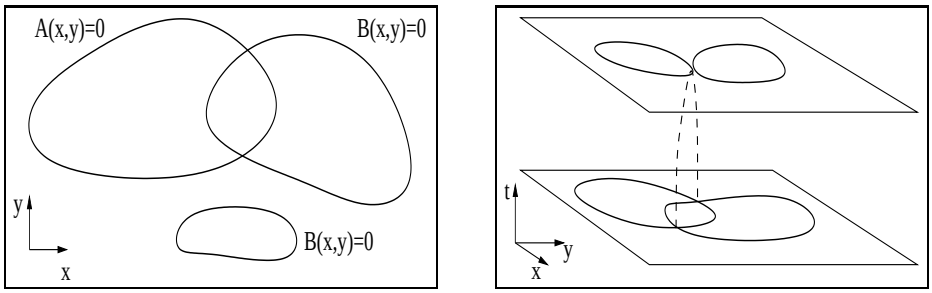


Figure 6.1: The superficial and deep structure of the intersection of curves. LEFT: Two curves A and B intersect generically in an even number of points or not at all. RIGHT: A stack of images illustrating the evolution of an intersection (dashed line). Top image corresponds to an annihilation event. A creation event is identical with reversed scale.

only consider typical behaviour (in a well defined mathematical manner). We will consider properties which hold for an open[2] and dense[3] subset of all images. Such properties are called generic, and corresponds well with the general notion of being typical: An image picked at random will with probability 1 be generic, and any image is infinitesimally close to a generic image [46]. We will now illustrate how transversal intersections in jet space describe geometric properties of our feature points, where after we will discuss the genericity of transversal intersections.

6.2.1 Geometric properties by transverse intersections in jet space

We shall establish generic properties in jet space (the space of k-th order Taylor expansions of solutions to the Heat Equation, $L_t = L_{xx} + L_{yy}$). A property is represented by its defining equations ($A = B = 0$) expressed in terms of image derivatives. This implicitly defines a manifold of some codimension[4] called the property manifold. To establish generic properties, we have to show that the property manifold is defined and determine its codimension.

[2] A set X is open if $\forall x \in X \exists \epsilon > 0 : N_\epsilon(x) \subset X$, where $N_\epsilon(x)$ is an ϵ-neighbourhood of x.

[3] A set X is dense in Y if and only if $\forall y \in Y \forall \epsilon > 0 : N_\epsilon(y) \cap X \neq \emptyset$

[4] The codimension of a manifold is the dimension of the embedding space minus the dimension of the manifold. E.g. the codimension of a sphere is always 1, while its dimension depends on the embedding space: A sphere in 3D is 2 dimensional and in 4D is 3 dimensional.

A scale-space image can be mapped to a smooth manifold in jet space in the following way: The k-th order local Taylor expansion is in each point of the scale-space mapped into jet space, thus sweeping a manifold, which we call the image manifold.

If the image manifold is transverse to the property manifold, then the set of points in scale-space, where the property occurs, is a smooth manifold with the same codimension as the property manifold. Hence we have described our feature manifold in scale-space.

6.2.2 Transversal intersections are generic

The scale-space images which have a transversal intersection between the image manifold and the property manifold form an open and dense set in the space of all scale-space images. This result is due to Damon [23]. Note that a scale-space image satisfies the Heat Equation, hence the evolution of spatial feature points is restricted in comparison to a general evolution. This changes the general program of catastrophe theory slightly [23].

By the above we conclude that typical feature properties can be determined by transversal intersection between images and a property in jet space.

6.2.3 Generic catastrophes for differential corner measures

The codimension of a manifold describes the number of constraints imposed on the degree of freedoms in the embedding space. We will now illustrate how the codimension for corner feature curves is found.

With some local parametrisation of the feature curve is an implicit parametrisation given by:

$$\begin{aligned} A(x(s), y(s), t(s)) &= 0, \\ B(x(s), y(s), t(s)) &= 0. \end{aligned}$$

Differentiating with respect to the parameter s yields:

$$\begin{aligned} A_x x_s + A_y y_s + A_t t_s &= 0, \\ B_x x_s + B_y y_s + B_t t_s &= 0, \end{aligned}$$

In a catastrophe point ($t_s = 0$) and with a non-zero curve tangent we have:

$$A_x B_y - B_x A_y = 0. \tag{6.3}$$

We have analysed (6.3) combined with $A = B = 0$, and we have determined that they define a smooth manifold of codimension 3 in jet space. See

Appendix A for further details. The equations $A = B = 0$ alone define a property manifold of codimension 2. The solutions to $A = B = 0$ therefore are smooth curves in scale-space with catastrophes occurring in isolated points on these curves. Geometrically, catastrophes for critical points in C correspond to the annihilation or creation of a saddle with either a minimum or a maximum. The catastrophes for the isophote constrained measure correspond to the annihilation or creation of a minimum and maximum pair.

The sign of t_{ss} determines the catastrophe type: negative is an annihilation and positive is a creation. Recall that $t_s = 0$ at a catastrophe point therefore the sign of t_{ss} determines if the curve moves towards positive or negative t-values from the catastrophe point and consequently the catastrophe is either a creation or an annihilation.

We have been able to create generic examples of both types, hence both annihilation and creation can occur. If $t_{ss} = 0$, we get a catastrophe of even higher order. Where $t_s = t_{ss} = 0$ defines a manifold of codimension 4 and are not generic, since scale-space images only have 3 dimensions.

We conclude that *tracking a corner over scale can only be performed for an open interval of scales.* At scales outside this interval, the corner does not exist. Hence, both annihilations and creations are generic events. In particular, it is in general not possible to trace back corners to arbitrary fine scales.

6.3 Experiments on two typical images

The above analysis shows that in the continuous case, annihilations and especially creations are likely when scale is increased. To verify these results in the discrete setting we have calculated corners of two typical images shown in Figure 6.2. The number of corners as a function of scale for Figure 6.2(LEFT) is shown in Figure 6.3. The corners are found as points where $|C|$ is larger than its 4 neighbours[5]. We see that it is not a monotonic decreasing sequence and conclude that creations do occur. In Figure 6.4 are shown a sequence of scales for $a = 0$ in which a corner creation occurs approximately in the middle image. The creation events for other a's are similar. We thus conclude that corners defined as $C_x = C_y = 0$ also have creations in the discrete setting.

We have tested the isophote approach ($L = L_0$ and $C_v = 0$) on the image in Figure 6.2(RIGHT). For binary images the mid-isophote (intensity value 0.5) is an edge operator, and it is to be expected that the mid-isophote is close to $L_{ww} = 0$ edges for low scales. However, binary images are not generic.

[5]This neighbourhood size is used to reduce computation time

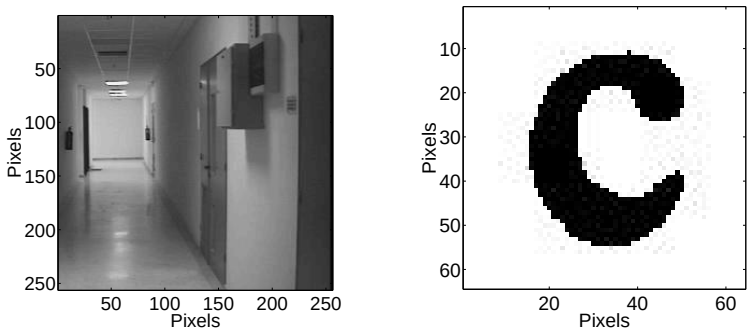


Figure 6.2: Two test images. LEFT: A typical hallway, courtesy C. Colios, FORTH. RIGHT: The letter 'c'.

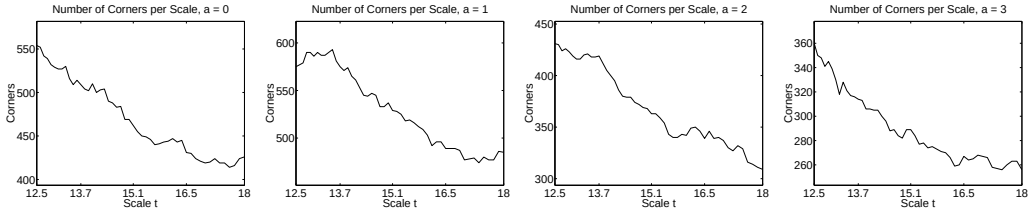


Figure 6.3: Corners of Figure 6.2(LEFT) given by $C_x = C_y = 0$ have frequent creations for all $a = 0, \ldots, 3$.

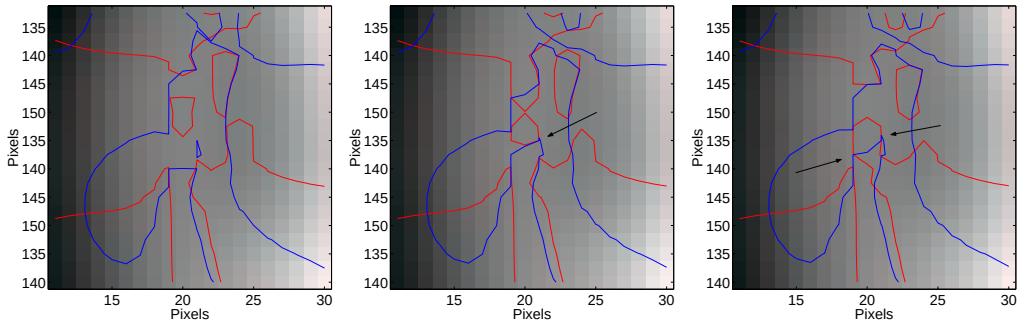


Figure 6.4: A zoom of Figure 6.2(LEFT) showing lines of $C_x = 0$ (red) and $C_y = 0$ (blue) for $a = 0$ and scale $t = 17.60, 17.70, 17.86$ from left to right. Arrows mark creations.

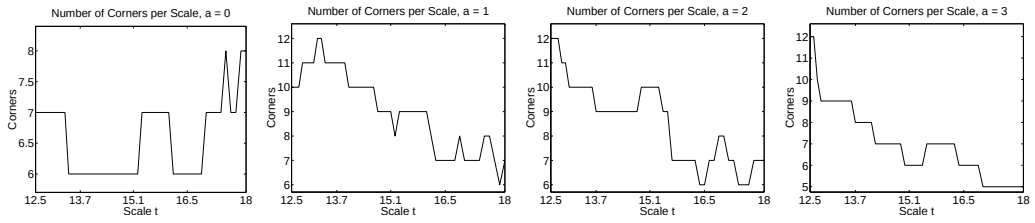


Figure 6.5: Corners given by $L = L_0$ and $C_v = 0$ have frequent creations for all $a = 0, \ldots, 3$.

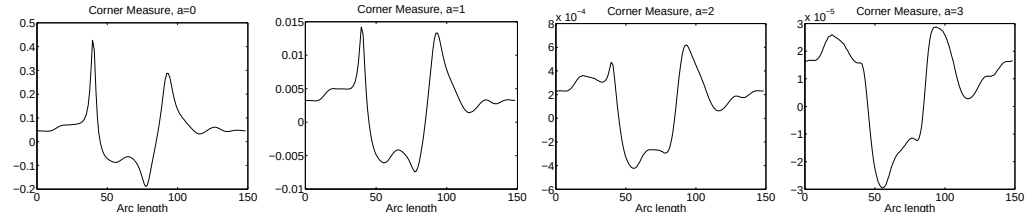


Figure 6.6: The function C on the mid-isophote for Figure 6.2(RIGHT) at $t = 12.5$ and for $a = 0, \ldots, 3$. The arc-length functions starts on the outside of 'c' at its leftmost point and travels downwards.

Fortunately, an arbitrary small scale larger than zero will transform a binary image into a generic gray-valued image.

The number of corners for a range of scales for the mid-isophote is shown in Figure 6.5. Corners are here found as points in which $|C|$ is larger than its 2 neighbours on the isophote. Again we see that this is not a monotonic decreasing sequence and conclude that the isophote approach also has creations in the discrete setting.

The final experiment we have performed concerns the selection of semantically important corners: Some approaches order the extrema of C according to their absolute strength at coarse scales and uses this ordering to choose the semantically important extrema at fine scale [73]. Varying a can alter the ordering dramatically and possibly catastrophically. In Figure 6.6 we show the value of C on the mid-isophote of the letter 'c' at scale 12.5 for various a. Notice the sharp peak for $a = 0$ at arc-length 40. This corresponds to the sharpest turn of the contour of the letter 'c'. For $a = 3$ this peak is practically removed and replaced by the neighbouring peak at arc-length 20. This shift in ordering also occurs for corners given by $C_x = C_y = 0$.

6.4 Summary

We have studied the family of corner measures $\kappa|\nabla L|^a$ for $a \in \{0, \ldots, 3\}$ when the image is embedded in Linear Scale-Space. Two approaches have been analysed: The catastrophe structure of the spatial extrema of the corner measure and the extrema along an isophote.

We have found in both approaches that the evolution of corner points are smooth curves in space and scale, and that both annihilations and creations are generic events at isolated points on these curves. Hence, corners cannot always be tracked to arbitrary coarse or fine scale.

For both approaches we have experienced that the value $a = 1$ seems to generate the fewest number of creations. Coincidentally, this is also the measure with the simplest intrinsic form: $C(x, 1, t) = L_{vv}$. Conversely, we have experienced that especially many catastrophes occur when $a = 0$.

Finally, the isophote approach has been shown to shift the focus away from high isophote curvature points for $a > 0$. The resulting corners do not always correspond well with intuition. A similar result holds for corners defined as spatial extrema of C.

Chapter 7

Singularity theory II: Genericity via equivalence

This chapter build on the theory presented in chapter 3. Here will be introduce concepts related to genericity via equivalence relations and stability under deformations. This chapter is meant to introduce concepts and ideas from the literature to be used in the next chapter on "The Structure of The Optic Flow".

The basic idea is to define an equivalence between functions. Each of the resulting equivalence classes is then considered as one type of structural behaviour. The equivalence is given by a diffeomorphic warping of the domain (see Figure 7.1). A homeomorphism is a bijective mapping which is continuous in both directions, a diffeomorphism is a homeomorphism which is infinitely differentiable in both directions. An ellipse, a circle and a square can be warped into each other by a homeomorphism. A square can not be transformed by a diffeomorphism into an ellipse but an ellipse and circle can transformed into each other by a diffeomorphism.

Mapping a smooth function to a smooth function by a local transformation of the coordinate system requires a smooth transformation. A diffeomorphic transformation cannot introduce or remove singularities in the level set. (see Figure 7.1). This is basically why it is interesting to study diffeomorphic transformations. The singularities in the level set for the transformed function is the same as for the original function. Hence, a horribly formed function can be transformed by a diffeomorphism into a function on a simple form. The topology of the simple form is easy to analyse and the obtained results are directly applicable to the original function. Even though diffeomorphisms preserve the topology they can make tremendous changes to the landscape, the limitations are that the transformation has to be smooth and is not allowed to introduce singularities. An equivalence relation can be

defined in the following way:

Definition 20 (Right-Equivalence)
Let U, V be open subsets of $\mathbb{R}^n$ and let $f, g : U, V \to \mathbb{R}$ be smooth functions. The functions f (at $x \in U$) and g (at $y \in V$) are right equivalent if there exists open neighbourhoods X, Y of x and y, a diffeomorphism $h : X \to Y$ and a constant c such that:

$$h(x) = y \quad \wedge \quad f(x) = g(h(x)) + c \tag{7.1}$$

In the following we will sketch the path from an equivalence relation to a list of normal forms. That is a list of germs which each represents an equivalence class. The list only consists of the generic types of behaviour.

The space of functions defined near zero is referred to as the space of germs. For example, let $f(x, y) = a(x - x_0)^2 + b(y - y_0)^2$. If x_0 or y_0 are changed the germ $ax^2 + by^2$ for this function f doesn't change but changing a or b does change the germ. Hence when considering germs, we consider functions moved to the same origin.

Right equivalence (and other types of equivalence) can be represented as a group of diffeomorphic germs. When a group acts on a germ in jet space a so called orbit is traced in jet space. The orbit consists of all members which can be reached from the original germ. This orbit, a set of jets, is a smooth submanifold [80]. Hence the results on transversal intersection between manifolds can be applied. All the orbits with sufficiently small codimension will be intersected transversely by jets from a residual set of functions. By this, it is established that these equivalence classes describe generic behaviour. The next move is to represent each class with a normal form, a germ with a simple form.

A deep result by Malgrange [80, 40, 24] often referred to as Malgrange preparation theorem allows to state the transversality conditions as infinitesimal stability under deformations. The unfolding theorem connects infinitesimal stability and global stability. The final piece is the determinacy theorem which insures that a stable germ is equivalent to a finite part of the Taylor series for the germ.

Now it is possible to list normal forms for simple germs.

Theorem 8 (Classification by Arnold [4])
A complete list of normal forms of a function in the neighbourhood of a simple critical point.

A_k $f = \pm x_1^{k+1} \pm x_2^2 + Q, k \geq 1$ $codim\ A_k = k-1$
D_k $f = x_1^2 x_2 \pm x_2^{k-1} + Q, k \geq 4$ $codim\ D_k = k-1$
E_6 $f = x_1^3 \pm x_2^4 + Q$ $codim\ E_6 = 5$
E_7 $f = x_1^3 + x_1 x_2^3 + Q$ $codim\ E_7 = 6$
E_8 $f = x_1^2 + x_2^5 + Q$ $codim\ E_8 = 7$
where $Q = -x_3^2 - \ldots - x_s^2 + x_{s+1}^2 + \ldots x_n^2$. [1]

A simple germ does not depend on parameters in contrast to t-modal germs which depend on t parameters. Orbits can either form discrete "stratifications" or continuous families [4]. The discrete case results in a simple germ. The case with moduli we will not consider further.

7.1 Determinacy

Here we will present a version of the determinacy theorem. We follow more or less the presentation by Bruce and Giblin [17]. We will not give proofs here the reader is referred to Bruce and Giblin and references therein[17].

We will need a bit of algebra to present the next result. A nice introduction to the subjects can be found in the book by Cox et al. [20].

The smooth germs $f : \mathbb{R}^n, 0 \to \mathbb{R}$ form a ring (denoted R_n) with simple point wise multiplication and addition. The subset of functions that vanish at 0 is an ideal in R_N, we denote it M_n. It is actually the maximal ideal in R_n.

Proposition 3 (Generating Ideals)

- *The ideal M_n is generated by $x_1, \ldots, x_n$. Which means that any $f \in M_n$ can be written as $f = \sum_{i=1}^n x_i f_i$ for some $f_i \in R_n$. We will write this as $M_n = R_n < x_1, \ldots, x_n >$*

- *The set of germs vanish on the subspace $x_1 = \ldots = x_{n-1} = 0$ is also an ideal, denoted M_{n-1} and is generated by $x_1, \ldots, x_{n-1}$.*

Definition 21 (M_n^k)
If M_n is a maximal ideal in R_n then we can recursively define a finitely generated ideal $M_n^k, k \geq 1$ in the following way:

$$M_n^k = M_n \cdot M_n^{k-1}, M_n^0 = M_n \tag{7.2}$$

A germ f lies in M_n^k if and only if f and its partial derivatives of order $\leq k$ all vanish at 0. Now we are in position to state the determinacy theorem.

[1] In theorem 8 the index s indicates the transition between negative and positive sign for the quadratic terms

Theorem 9 (Determinacy[17])
Let $f : \mathbb{R}^n, 0 \to \mathbb{R}, 0$ be a smooth germ with the following ideal $I = M_n^2 < \frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n} >$, if M_n^{k+1} is contained in the I then f is right equivalent to any other germ g having the same k-jet (the same Taylor series up to and including the terms of order k). In this case f is called k-right-equivalent.

7.2 How to read a list of normal forms

A list of normal forms gives a canonical representative (the normal form) of the equivalence classes deduced from the equivalence relation. It is the smallest list for which it is true that there is a residual set of functions for which all points in their domain can be locally described by one of the local forms in the list. In this sense the list is complete.

This does not imply that all functions in the residual set exhibit all the normal forms in their domain. For example the list of normal forms for functions include regular points and non-degenerated critical points. The open and dense set A of functions which can exhibit these properties are called Morse functions. The subset B of strictly increasing monotonic functions $(\forall x : f_x(x) > 0)$ are also Morse functions but they do not have critical points at all.

One could ask if the set $A \backslash B$ where all members have at least one critical point cover almost all types of functions. But that is not the case: the set $A \backslash B$ is not dense because B is an open set. Hence the set B can not be ignored without losing genericity.

7.3 Genericity via equivalence for the heat equation

The results presented in this section are due to Damon[23, 24, 25] and references therein. The used equivalence relations are especially designed to capture the special role played by the time parameter in the heat equation. Two equivalence relations, H-equivalence and IS-equivalence, are in play.

Definition 22 (H-equivalence[24])
Germs $f, g : \mathbb{R}^{n+1}, 0 \to \mathbb{R}, 0$ are H-equivalent if there is a local change of coordinates $\phi : \mathbb{R}^{n+1}, 0 \to \mathbb{R}^{n+1}, 0$ of the form $\phi(x,t) = (\phi_1(x,t), \phi_2(t))$ with $\phi_2'(0) > 0$ and a germ $c(t) : \mathbb{R}, 0 \to \mathbb{R}, 0$ such that

$$g(x,t) = f \circ \phi(x,t) + c(t) \tag{7.3}$$

Note how ϕ_2 only depends on t by this, the special role of the variable t is preserved. In order to keep track of both the isointensity surface as well as the intensity level of the critical point Damon introduces a more restrictive equivalence than H-equivalence. The new equivalence class is denoted Intensity-Sensitive equivalence.

Definition 23 (IS-equivalence[24])
Germs $f, g : \mathbb{R}^{n+1}, 0 \to \mathbb{R}, 0$ are IS-equivalent if there is a local change of coordinates $\phi : \mathbb{R}^{n+1}, 0 \to \mathbb{R}^{n+1}, 0$ of the form $\phi(x,t) = (\phi_1(x,t), \phi_2(t))$ with $\phi_2'(0) > 0$ and a local change of coordinates $\psi(y,t) : \mathbb{R}^2, 0 \to \mathbb{R}^2, 0$ of the form $\psi(y,t) = (\psi_1(y,t), t)$ with $\frac{\partial \psi_1}{\partial y}(0,0) > 0$ and $\psi_1(0,t) = 0$ for all t, so that

$$g(x,t) = \psi_1(f \circ \phi(x,t), t) \tag{7.4}$$

Note that the two types of equivalence do not preserve the heat equation. Hence, a solution to the heat equation is equivalent to a function outside the solution space. The equivalences are designed by Damon to capture exactly the structures of interest. The designed equivalences can be viewed as lying in between two limiting equivalences. In one limit, the equivalence results in a too fine classification: All quadratic functions where the coefficients sum to zero are solutions to the heat equation. An equivalence which preserves the heat equation will have to leave the eigenvalues unchanged but this will result in an equivalence class for each possible set of eigenvalues even though the level set only depends on the sign of eigenvalues. In the other limit, Morse theory states that for a function with a critical point with a nonsingular Hessian matrix there exists a diffeomorphism which warps the function to a quadratic form where all the coefficients are plus or minus one. Using this kind of equivalence will result in a too coarse classification. H-equivalence and IS-equivalence are two possible choices in between. IS-equivalence is a refinement of H-equivalence. Not only are the changes in the isosurfaces tracked also the intensity level of the critical point is tracked by IS-equivalence.

Example 4 (H-equivalence and IS-equivalence)
The following functions are H-equivalent

$$\begin{aligned} f(x,y,t) &= x^2 - y^2 \\ g(x,y,t) &= 3x^2 - 2y^2 + 2t \\ h(x,y,t) &= x^2 - 2y^2 - 2t \end{aligned}$$

f and g are H-equivalent by the following transformation:

$$g(x,y,t) = f \circ \phi(x,y,t) + c(t) \text{ where}$$

$$\begin{aligned} \phi(x,y,t) &= ((\sqrt{3}x, \sqrt{2}y), t) \text{ and} \\ c(t) &= 2t \end{aligned}$$

f and h are H-equivalent by the following transformation:

$$\begin{aligned} h(x,y,t) &= f \circ \phi(x,y,t) + c(t) \text{ where} \\ \phi(x,y,t) &= ((x, \sqrt{2}y), t) \text{ and} \\ c(t) &= -2t \end{aligned}$$

None of the three functions are IS-equivalent. IS-equivalence refines the equivalence class induced by H-equivalence and the three functions are members of different equivalence class.

Based on the equivalence relations Damon introduces the notion of stability under deformation. A deformation of a germ $g : \mathbb{R}^{n+1}, 0 \to \mathbb{R}, 0$ is a germ $f : \mathbb{R}^{n+1+q}, 0 \to \mathbb{R}, 0$ such that $f(x,t,0) = g(x,t)$.

Definition 24 (H-stable)
A germ $g : \mathbb{R}^{n+1}, 0 \to \mathbb{R}, 0$ is H-stable if for any deformation f there is a local change of coordinates $\phi : \mathbb{R}^{n+1+q}, 0 \to \mathbb{R}^{n+1+q}, 0$ of the following form $\phi(x,t,u) = (\phi_1(x,t,u), \phi_2(t,u), u)$ with $\frac{\partial \phi_2}{\partial t}(0,0) \geq 0$ and a germ $c(t,u) : \mathbb{R}^{1+q}, 0 \to \mathbb{R}, 0$ so that

$$g(x,t) = f \circ \phi(x,t,u) + c(t,u) \tag{7.5}$$

Definition 25 (IS-stable)
A germ $g : \mathbb{R}^{n+1}, 0 \to \mathbb{R}, 0$ is IS-stable if for any deformation f there is a local change of coordinates ϕ as in definition 24 and $\psi : \mathbb{R}^{2+q}, 0 \to \mathbb{R}^{2+q}, 0$ of the form $\psi(y,t,u) = (\psi_1(y,t,u), t, u)$ with $\frac{\partial \psi_1}{\partial y}(0,0,0) > 0$ and $\psi_1(0,t,u) = 0$ for all (t,u) and a germ $c : \mathbb{R}^q, 0 \to \mathbb{R}, 0$ so that

$$g(x,t) = \psi_1(f \circ \phi(x,t,u), t, u) + c(u) \tag{7.6}$$

The functions in example 4 are all H-stable but only g and h are IS-stable. For f the critical value remains the same for all values of t, intuitively a highly unstable situation. This intuition is formalised by IS-stability.

A germ f is H-stable respectively IS-stable if it is its own versal unfolding. In other words, a sufficiently small perturbation of f is H-equivalent, respectively IS-equivalent to itself. This stability under deformation corresponds to

the stability formulated in topological terms, namely openness. The equivalence relation can "act" in jet space by the corresponding diffeomorphism group. The orbits traced by the group action are submanifolds in jet space and via transversality arguments topological stability can be established for the equivalence.

Damon [23, 24] proves that it is a generic property to be H-stable or IS-stable for solutions to the heat equation. A list of H-stable and IS-stable germs which describe the local structure in the neighbourhood of the point of interest is presented in table 7.1 from [24].

Case	Normal form	Description/conditions
0	x_1	submersion
1	$\sum_{i=1}^{n} a_i x_i^2 + 2ct$	$\forall a_i \neq 0, c = \sum_{i=1}^{n} a_i$
1'	$\sum_{i=1}^{n} a_i x_i^2 \pm E(t^2)$	$\forall a_i \neq 0, \sum_{i=1}^{n} a_i = 0$
2	$x_1^3 + 6tx_1 + \sum_{i=2}^{n} a_i x_i + 2ct$	$\forall a_i \neq 0, c = \sum_{i=2}^{n} a_i$
3	$x_1^3 - 6tx_1 - 6x_1x_2^2 + \sum_{i=2}^{n} a_i x_i^2 + 2ct$	$\forall a_i \neq 0, c = \sum_{i=2}^{n} a_i$

Table 7.1: A complete list of H-stable and IS-stable germs. All cases are in use for IS-equivalence and only 0,1,2 and 3 are used for H-equivalence. For two normal forms to be H or IS-equivalent within the same case the quadratic forms have to have the same index. There are extra conditions for IS-equivalence: If $\sum_{i=1}^{n} a_i = 0$ then case 1' is used. Also the sign of c has to be the same for two forms to be IS-equivalent. The index is determined by counting the signs of a_i's. $E(t^2)$ is the lowest order polynomial solution to the heat equation which has t^2 as one of the terms. In 2D $E(t^2) = t^2 + tr^2/2 + r^4/32, r = x^2 + y^2$

.

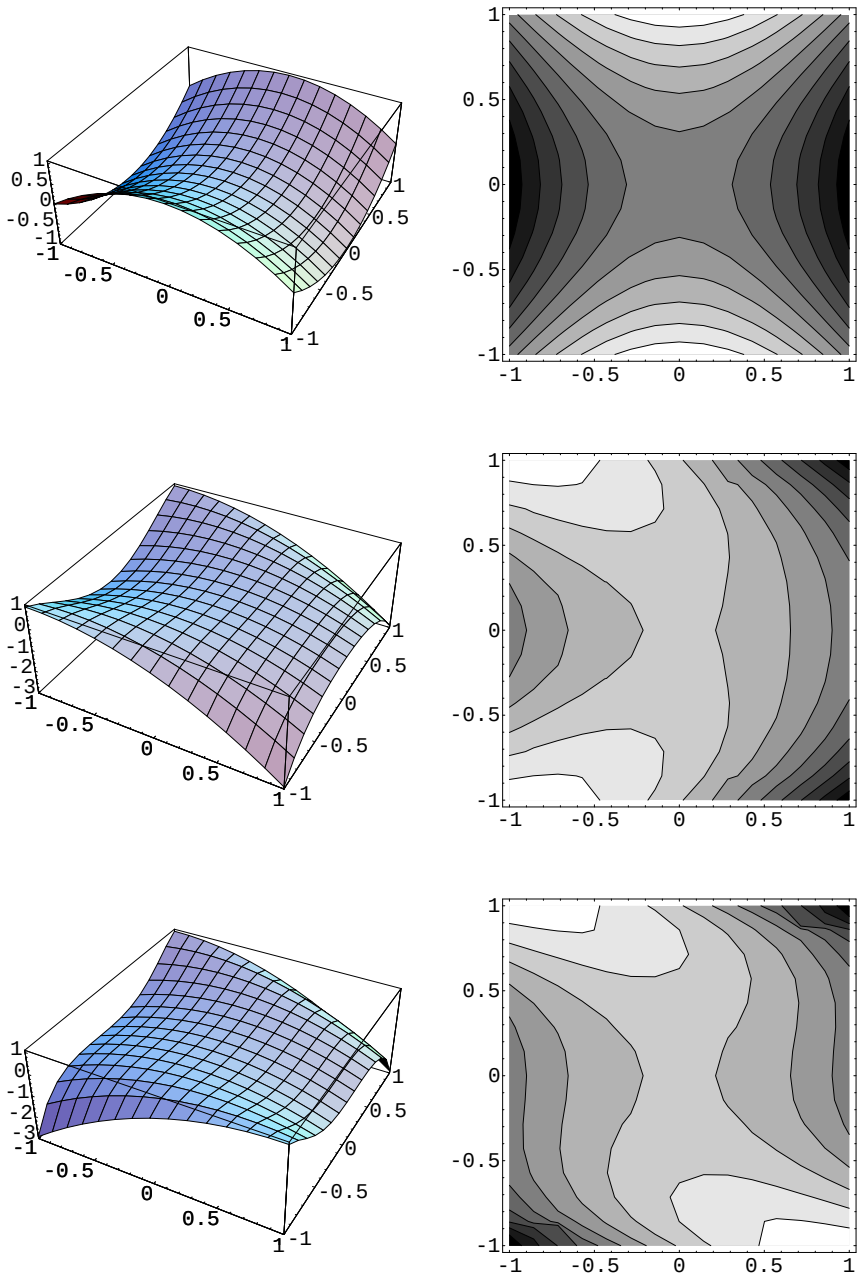


Figure 7.1: A diffeomorphic transformation of a saddle structure. The three graphs are equivalent under right equivalence. From top to bottom the graphs are given by $f(a,b) = -a^2 + b^2$, $g(x,y) = -x^2 + y^2 - 2xy^2 - y^4$ and $h(x,y) = -x^2 + y^2 - 2xy^3 - y^6$. The following transformations warp f to g and h, respectively: $(a,b) = (x + y^2, y)$, $(a,b) = (x + y^3, y)$

Chapter 8

The structure of the optic flow [1]

8.1 Introduction

A lot of work on optic flow has been devoted to its definition [50, 30] and to regularisation schemes for its robust computation [50, 89, 9, 99, 137]. In this thesis, we follow the Horn and Schunck definition of the optic flow field [50]. We do not regularize the solution, but only wish to classify it. The motivation is fourfold: we seek a classification, in mathematical terms, of the flow field and its temporal changes. We want to emphasise that the events we describe or detect in images must be generic events. We will develop mechanisms for detecting these events, and finally we wish to indicate that this purely academic examination of the optic flow field may sub serve the development of algorithms for many different visual tasks. In this chapter we give simple examples using the flow structure for guiding an attention mechanism and for computing the degree of turbulence in a flow field. The inspiration is mainly from the analysis of autonomous dynamical systems [8] to which we will describe the analogy.

An object moving with respect to a camera induces a motion field on the image plane. This motion field is the projection of the motion of physical points fixed to the object, and will only under very restricted lighting and reflectance circumstances directly relate to the optic flow field [54]. We analyse the singularities of the data induced *optic flow field* while the singularities of the *motion field* have earlier been analysed for recovery of object motion parameters [65, 130]. Since we do not in this work relate the optic flow field to the motion field, we can not make similar observations.

[1] This chapter has been published in a conference proceedings [92]

The structure of a general analytic flow field is normally assessed through the singularities of the field, i.e. points where the flow vanishes[8]. The first order structure of the flow field round these points can classify the points as attractors, repellers, or whirls. The optic flow field is a special flow field since computed as a simple intensity preservation. Generically the optic flow field is not everywhere analytical. Furthermore, the tangential component of the flow is not determined by the constraint equation. These differences cause new flow structures to be created generically and ill-define the classical flow structures in an optic flow field. This chapter analyses these differences, show examples, and applications.

A temporarily changing image may be obtained from imaging a dynamic scene. A physical conservation law then defines a spatial vector field temporarily connecting conserved properties. In computer vision the Horn and Schunck (HS) equation [50] expresses the preservation of intensity over time. The derived spatial vector field is the *optic flow field*. The HS equation only solves locally for one component of the vector field, giving rise to the so-called aperture problem. An unique representation of the optic flow field is given by the normal flow: the flow perpendicular to the local isophote.

The normal flow is well defined in all image points with non-vanishing spatial gradient, elsewhere the flow is undefined. However, in a neighbourhood round these singular points the flow field exhibits some typical behaviour. The flow magnitude typically increases towards infinity, the direction will be inwards, outwards, or combinations hereof. In Section 8.7 we analyse the flow around these singularities, categorise it, and see how these poles change over time. The purpose of this exercise is to gain insight in the *structure of the optic flow field*.

In general analytic flow fields, vanishing flow points can describe the flow field structure. In optic flow fields only the normal flow is directly accessible and this will generically vanish at hypersurfaces of codimension 1, i.e. at curves in 2D images. This means, that the standard classification of the flow field structure can not directly be applied to the normal flow fields. We can, however, define whirls in the normal flow field as second order temporal events and apply the detection of these to the quantisation of turbulence.

A proper definition of structure change in the flow field needs a definition of *structure*. We do this through the mathematical concept of equivalence of flow fields under deformations. In Section 8.6 we review this method and its application to structural classification of analytical flow patterns. In Section 8.7 we define and derive the generic structure of the optic flow field.

Normally, the optic flow has been defined and computed directly based on pixel values, so as if they represent the true value of the intensity field. Recently [30] the optic flow definition and computation have been formulated

in a scale-space framework taking the finite extent of pixels and filter-outputs into account. In Subsection 8.10 we comment on some aspects of the change of structure of the optic flow field when scale changes.

In Section 8.11 we describe how changes of flow field structure can be detected from outputs of linear filters. These events may be useful to detect violations of continuation models: structure emerging or disappearing, and thus guide an attention mechanism. We give examples of computations of the flow structure, and apply this to simple examples from computer vision and turbulent flow.

First, however, we look into the necessary notation and definitions.

8.2 Optic flow: notation and definitions

In this section we establish notations of what images, optic flow, normal flow, and the spatio-temporal iso-surface are. Furthermore we link the geometric properties of the spatio-temporal iso-surface to the flow and normal flow-field. We assume that I is sufficiently differentiable.

Definition 26 (Image sequence)
An image sequence $I(x,t) : \mathbb{R}^D \times \mathbb{R} \mapsto \mathbb{R}$ *is a mapping from* D *spatial dimensions* ($x = (x_1, x_2, \ldots, x_D)$) *and a temporal dimension* (t) *into scalar values, normally denoted* intensities.

Definition 27 (Spatio-temporal optic flow field)
The spatio temporal optic flow field $v : \mathbb{R}^{D+1} \mapsto \mathbb{R}^{D+1}$ *is any vector field preserving image sequence intensity along flow lines.*

The preservation of intensity along flow-lines of v is expressed through the full (or Lie) derivative along the flow:

$$\mathcal{L}_v I = v^{x_1} I_{x_1} + v^{x_2} I_{x_2} + \ldots + v^{x_D} I_{x_D} + v^t I_t = 0$$

where upper index denotes component of vector and lower index denotes partial differentiation. In the following we will often use notation from $2D$ (x, y) to simplify expressions: $\mathcal{L}_v I = v^x I_x + v^y I_y + v^t I_t = 0$.

8.3 Temporal gauge

The optic flow equation yields one equation in $D + 1$ unknowns. In general, the length of the vector is unimportant as only the actual connection of spatio-temporal points carries information. Often the length of the vector

field is normalised to unit temporal component, i.e. $v^t = 1$. This normalisation is referred to as the temporal gauge. In this normalised flow field we denote the spatial components u. This reveals the well known Horn and Schunck equation

$$u^x I_x + u^y I_y + I_t = 0$$

We call (u^x, u^y) the spatial optic flow field or simply the optic flow field. This flow field answers the typical question asked by the computer vision programmer: in next frame of my image sequence, where did points move?

Definition 28 (Spatial optic flow field)
The spatial optic flow field $u : \mathbb{R}^D \mapsto \mathbb{R}^D$ *is any vector field preserving image sequence intensity along flow lines of the spatio-temporal vector field* $v = \begin{bmatrix} u \\ 1 \end{bmatrix}$.

The normalisation of the temporal component is, however, only possible (finite) when $v^t \neq 0$. In cases where the temporal component vanishes, the above formulation yields singularities (poles) in the flow field. In Section 8.7, we analyse the flow field round these poles, show that they exist generically, and that they exhibit certain generic behaviours/interactions. In order to do this, even though the spatial flow field is our main concern, we must stay in the spatio-temporal formulation of the flow field. In this way, we can derive properties of the spatial flow field from simple geometric considerations.

8.4 Normal flow

The temporal gauge, i.e. $v^t = 1$, (or another normalisation) results in one equation in D unknowns. This shows the intrinsic degree of freedom in the flow, normally denoted the *aperture problem*. Using the temporal gauge, the spatial optic flow constraint equation reads $u \cdot \nabla I = -I_t$ where u is the spatial flow field and ∇ denotes the spatial gradient. The component of u along the spatial image gradient (the normal flow) is uniquely determined by the constraint equation, while any component in the iso-intensity tangent plane is unresolved. The normal flow will therefore often be considered the solution to the optic flow constraint equation, keeping in mind that any tangential component can be added.

8.5 Spatio-temporal iso-surfaces

When looking at the spatio-temporal flow at a given point in space-time (x_0, t_0) it is constrained to preserve intensity.

Definition 29 (Spatio-temporal iso-surface)
In every point (x_0, t_0) where $I_\alpha \neq 0$, $I(x,t) = I(x_0, t_0)$ defines the corresponding spatio-temporal iso-surface.

Any flow line is constrained to lie within the spatio-temporal surface. This surface is only defined for points where the spatio-temporal gradient I_α does not vanish. Whenever the image is continuous, the spatio-temporal surface is a closed surface differentiable to the same order as the image.

If $I_t(x_0, t_0) \neq 0$ then the spatio-temporal iso-surface can be locally parametrised by the spatial coordinates. In this situation, both the intensity change in the time direction and the tangent plane to the iso-surface will not locally be perpendicular to the spatial directions, i.e. the normal flow is not zero.

Definition 30 (Spatio-temporal iso-function)
The function $s(x) : \mathbb{R}^D \mapsto \mathbb{R}$ is defined in an open set round every point (x_0, t_0) where $I_t(x_0, t_0) \neq 0$ such that $I(x, s(x)) = I(x_0, t_0)$.

The graph of the (spatio-temporal) iso-function is simply the iso-surface. The iso-function is linked to the local flow pattern through proposition 4:

Proposition 4 (iso-function normal flow)
The spatial normal flow through a point (x_0, t_0) is $u_n(x_0, t_0) = \|\nabla s\|^{-2} \nabla s$.

Proof 1
The flow is determined by the equation $u \cdot \nabla I + I_t = 0$. By spatially differentiation of the definition of the iso-function $\partial_x I(x, s(x)) = 0$ we find $\nabla s = -\frac{\nabla I}{I_t}$ and thereby the optic flow equation reads $u \cdot (-I_t \nabla s) + I_t = 0$. This reduces to $u \cdot \nabla s = 1$ which is obviously fulfilled by $u = \frac{\nabla s}{\|\nabla s\|^2}$. Since ∇s is directed along the image gradient ∇I, this is the normal flow.

We are now capable of linking the geometric structure of the iso-function to the local normal flow. In the points where its tangent (hyper)plane coincides with the (hyper)plane spanned by the spatial dimensions, the normal flow is not defined, but in neighbouring points on the iso-surface, we can find the flow and categorise the undefined flow by its limiting structure.

In the following we will analyse the generic shape of the spatio-temporal iso-surfaces in terms of the spatio-temporal iso-function. Especially we will analyse the generic properties and the corresponding generic flow patterns.

8.6 Structure and genericity

We define, as in common catastrophe theory [41], structure as equivalence classes under deformations. That is, given a function $f(x, c) : \mathbb{R}^D \times \mathbb{R}^k \mapsto \mathbb{R}$, where x are the D spatial coordinates and c are the k control parameters of the function, define equivalence classes from a class of allowable deformations of x and c. In common catastrophe theory, we define $x' = \phi(x, c), c' = \psi(c)$, where ϕ and ψ are diffeomorphisms. Now the game is, given a function f, to choose ϕ, ψ such that $f(x', c')$ takes a special algebraic form. As an example, any point where the spatial gradient of f does not vanish, can by a correct choice of ϕ, ψ be put on the form $f(x', c') = x'_1$. We call this representation of f the normal form. Such an analysis of C^∞ functions leads to Thom's classification of catastrophes: regular points, critical points, folds, cusps, swallow tails, etc., each represented by a normal form.

An event is generic if one can not perturb it away with an infinitesimal perturbation. Mathematically that is, the event occurs in an open and dense set of all functions f. We expect to see only generic events in real image sequences, as all other events have measure zero in C^∞. Using the transversality theorem[41], one can argue on genericity simply by counting dimensions. In the product space of spatial coordinates and control parameters we expect an event to occur at a manifold of dimension $D + k$ minus the number of linear constraints on the functions jet to be satisfied for the event to take place.

In the case of flow fields, or dynamical systems, the class of diffeomorphisms is constrained since the flow field represents a connection of physical points. That is, the deformation of the coordinate system is not allowed to change the flow fields connection of physical points, only its coordinate representation. The result of this is that one cannot remove points of vanishing flow, and one cannot alter the eigenvalues of the matrix containing the spatial first order derivatives of the flow [8]. This leads to a very fine classification of flow fields since the eigenvalues index the equivalence class. Generically we find in a flow field at a given time instance points of vanishing flow (fixed points), and in 2D we categorise them according to their eigenvalues: two positive implies an unstable node (repeller), two negative a stable node (attractor or sink), two of opposite sign a bistable node (saddle node), and a pair of complex conjugate eigenvalues yields a spiralling flow called a focus and which may be stable or unstable according to the value of the real part of the eigenvalues.

When time varies the fixed points may interact changing the fixed point topology of the flow. Transitions which takes place generically when a single parameter (the time) is varied are called codimension 1 events. In case of

the flow field, one generically meets three different events at codimension 1: scatter, saddle bifurcation, or Hopf bifurcation.

8.7 Structure of the optic flow field

In this section, we apply the general scheme outline above to the analysis of the structure of an optic flow field. An optic flow is defined through the conservation of image intensity along flow-lines. We define equivalence of the flow-field as identical up to a diffeomorphism of the image sequence:

Definition 31 (Image isophote equivalence)
Two images $I(x,t) : \mathbb{R}^D \times \mathbb{R} \mapsto \mathbb{R}$ *and* $J(x,t) : \mathbb{R}^D \times \mathbb{R} \mapsto \mathbb{R}$ *are isophote equivalent or I-equivalent if* $I(x,t) = \tilde{J}(\tilde{x},\tilde{t})$*, where* η, ψ *and* ϕ *are diffeomorphisms with the following constraints:*

$$\tilde{J}(x,t) = \eta(J(x,t)), \qquad \tilde{x} = \psi(x,t), \qquad \tilde{t} = \phi(t)$$

where $\eta'(0) > 0$*,* $\phi'(0) > 0$ *and* $Jacobian(\psi(.,0)) > 0$ *since we want to distinguish also the direction of flow on flow-lines.*

Notice, that η is not a function of x or t. This is because the optic flow is dependent on the iso-intensity line (isophote) structure, and a varying diffeomorphism would change this structure. η can only change the intensity values, but not change the isophotes. Without further restrictions the classification of flow structure is rather crude, and we make the following class, which leads to a finer classification.

Definition 32 (Image stationary equivalence)
Two images $I(x,t) : \mathbb{R}^D \times \mathbb{R} \mapsto \mathbb{R}$ *and* $J(x,t) : \mathbb{R}^D \times \mathbb{R} \mapsto \mathbb{R}$ *are stationary equivalent or S-equivalent if they are I-equivalent with* $\psi(0,t) = 0$.

This more restrictive equivalence cannot change points of zero flow, unlike I-equivalence. None of them can remove critical points in the iso-functions. We introduce S-equivalence to make the analogy to nodes and foci of analytical flow fields. An even more restrictive equivalence class could be constructed, not allowing the spatial diffeomorphism to vary in time. This would be even more analogous to the classification of the analytical flows since the total first order flow structure would be invariant under the diffeomorphisms. This classification is, however, too fine in our taste and the S-equivalence suffices for our purposes, so we will not pursue this direction further in this thesis.

We define local I-equivalence (local S-equivalence) in (x_0, t_0) as being I-equivalent (S-equivalent) in an open set round (x_0, t_0).

8.8 I-equivalent structure of the optic flow field

The normal flow is uniquely determined by the spatio-temporal iso-surface whenever the spatio-temporal image gradient does not vanish, and in positions where its tangent plane is not parallel to the time axis, the flow is uniquely determined by the spatio-temporal iso-function. Under I-equivalence the tangent plane of the iso-surface can be tilted away from being parallel to the time axis, and we need not treat this case separately. Hence we only analyse for general analytic iso-functions and vanishing spatio-temporal gradient.

Proposition 5 (I-normal forms of 2D optic flow, codim 0)
At a fixed time-slice $t = t_0$ the normal flow is generically in any point I-equivalent to one of the following normal forms:

$$\begin{aligned} n_0(x,y) &= (1,0)^T \\ n_2(x,y) &= \frac{1}{x^2+y^2}\begin{pmatrix} \pm x \\ \pm y \end{pmatrix} \end{aligned}$$

where the sign combinations in n_2: $(+,-)$ and $(-,+)$ are equivalent.

Proof 2
At a fixed time slice, the spatio-temporal gradient of the image will not vanish generically (this happens at codimension 1), and thereby the iso-surface is defined in every point. According to the arguments above we need only to analyse analytic iso-functions. Any regular point on the iso-function is I-equivalent to $s(x,y) = x$, and by using Prop. 4, we find n_0. The only generic critical points are Morse critical points, which are I-equivalent to $s = \pm\frac{1}{2}x^2 \pm \frac{1}{2}y^2$. Again using Prop. 4, we find n_2.

The normal form n_2 has respectively identical spatial flow-lines to the stable node $(-,-)$, the saddle $(-,+)$, and the unstable node $(+,+)$ of an analytical flow field. However, the velocity increases towards plus/minus infinity when the point approaches $(0,0)$.

The classification of the flow follows directly from the classification of critical points in analytical functions making the progress simple. The only twist is that for codimension ≥ 1 there exists generically points where the iso-surface is not defined.

Proposition 6 (I-normal forms of 2D optic flow, codim 1)
At a fixed time-slice $t = t_0$ the normal flow is generically in any point I-equivalent with codimension 1 to the normal forms of Prop. 5 or one of the following normal forms:

$$
\begin{aligned}
n_{2+1}(x,y) &= \frac{1}{x^2+y^2}\begin{pmatrix} xt \\ yt \end{pmatrix} \\
n_3(x,y) &= \frac{1}{(x^2+t)^2+y^2}\begin{pmatrix} x^2+t \\ y \end{pmatrix}
\end{aligned}
$$

where x, y, and t may independently change sign.

Proof 3
At codimension 0 we find the normal forms of Prop. 5. In codimension 1 we divide the analysis into two distinct case. Firstly we analyse the case where the spatio-temporal iso-surface is defined (the fold), secondly the case where it is not defined (the spatio-temporal critical point).

When the spatio-temporal surface is defined, we can using I-equivalence transform it into the normal form of general analytical functions. We use the theorem that iso-surfaces behave as generic functions [62], and find, at codimension 1, the only extra normal form compared to codimension 0 is the fold[41]: $s(x,y) = x^3 + tx + y^2$. By use of Prop. 4, we find n_3. Since I-equivalence only allows positive Jacobians in the diffeomorphisms, we must represent the signs of y and t explicitly.
When the spatio-temporal image gradient vanishes in (x_0, t_0), the iso-surface is not defined in this point. We can bring a spatio-temporal critical point on the following normal form by I-equivalence (up to signs of the individual terms): $I(x,y,t) = x^2 + y^2 + t^2$. We divide into two cases dependent on the sign of t, and find $s = \pm\sqrt{(I_0 - x^2 - y^2)}$, where I_0 is the intensity of the iso-surface. By Prop. 4 we find the normal flow:

$$
n = \frac{\pm\sqrt{I_0 - x^2 - y^2}}{x^2+y^2}\begin{pmatrix} x \\ y \end{pmatrix}
$$

By substitution of the expression for $t = s(x,y)$ into this, the sign cancels out, and we find in both cases n_{2+1}.

Codimension one events take generically place in fixed time slices in a time sequence; The top of Figure 8.1 illustrates these events. A stable pole will always meet the saddle pole in its unstable direction while an unstable pole meets a saddle in its stable direction. This is illustrated in Figure 8.2.

Proposition 7 (I-normal forms of 2D optic flow, codim 2)
At a fixed time-slice $t = t_0$ the normal flow is generically in any point I-equivalent with codimension up to 2 to one of the normal forms in Prop. 5

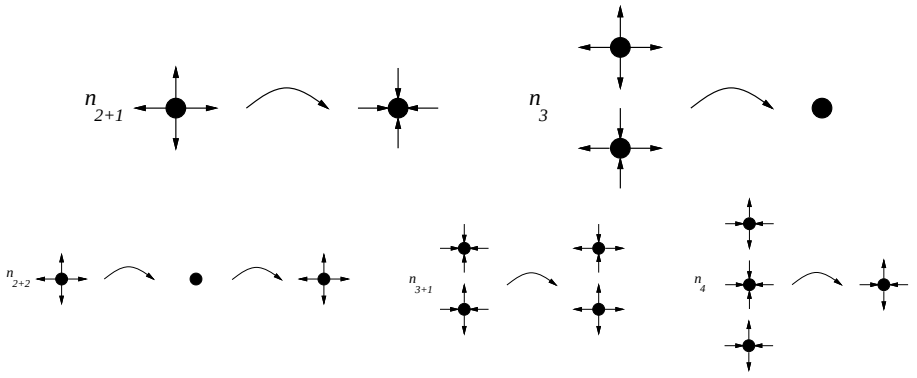


Figure 8.1: **Top**, the generic events of codimension 1. n_{2+1} is a pole, where the directions of the flow are reversed. n_3 describes the interaction of two poles, where a saddle pole and a (un)stable pole interact and annihilate or are created. **Bottom**, the generic events of codimension 2. n_{2+2} is a pole that may change direction twice. n_{3+1} is an annihilation in t_2, but t_1 interchanges the stability of the poles. For instance, a stable pole and a saddle approach like at an annihilation, but they scatter and become a saddle and an unstable pole. n_4 is a pitchfork bifurcation. An example is a stable pole and two saddles approach, interact and become a single saddle.

or Prop. 6 or one of the following normal forms:

$$\begin{aligned} n_{2+2}(x,y) &= \frac{t_1^2+t_2}{x^2+y^2}\begin{pmatrix} x \\ y \end{pmatrix} \\ n_{3+1}(x,y) &= \frac{t_1}{(x^2+t_2)^2+y^2}\begin{pmatrix} x^2+t_2 \\ y \end{pmatrix} \\ n_4(x,y) &= \frac{1}{(x^3+t_2x+t_1)^2+y^2}\begin{pmatrix} x^3+t_2x+t_1 \\ y \end{pmatrix} \end{aligned}$$

where x, y, and t_i may independently change sign. Any of the t_i may correspond to the physical time parameter.

Proof 4

The proof follows the lines of the proof in Prop. 6. First we divide into cases where the spatio-temporal iso-surface is defined or not. n_4 follows easily from a cusp in the iso-function: $s(x,y) = x^4 + t_2x^2 + t_1x + y^2$.

In case of a vanishing spatio-temporal gradient, we subdivide into two cases dependent on whether I_{xx} or I_{tt} vanishes in the spatio-temporal critical point. In the first case we have a spatial fold as in n_3, but augmented by a vanishing temporal derivative, yielding n_{3+1}. In the latter case we have a spatially critical point in which a temporal fold happens, yielding n_{2+2}. The algebraic derivations are similar to the derivation of n_{2+1}.

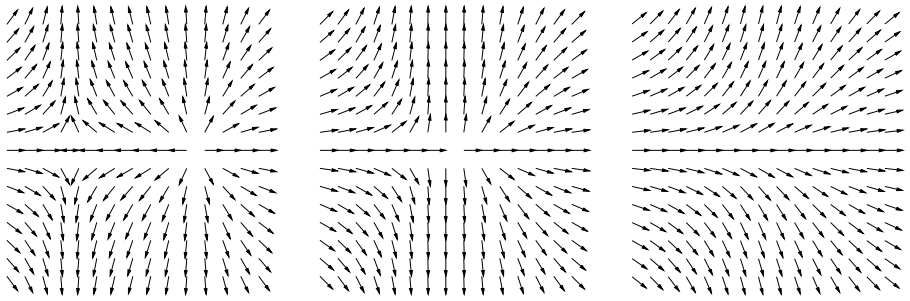


Figure 8.2: The n_3 normal form for $t = -0.25, 0, 0.25$. A saddle and an unstable pole (left) meet (middle) and annihilate (right). Only the orientation of the flow is shown. The magnitude increases towards infinity near the poles.

In these normal forms we use two control parameters of which one is the time parameter and the other are maybe most easily visioned as a scale parameter even though these forms have not been proven to be the normal forms when the evolution along a control parameter is constrained as in the case of Gaussian scale space.

The codimension two events are illustrated schematically at the bottom of Figure 8.1. Assume t_1 is the time parameter, and t_2 is negative. Then when $t_1^2 > |t_2|$ n_{2+2} is an unstable pole, and when $t_1^2 < |t_2|$ it is a stable pole. If t_2 is positive, it is always unstable. Exactly when $t_2 = 0$, the pole disappears for $t_1 = 0$ but reappears with same orientation infinitesimally later.

8.9 S-equivalent structure of the optic flow field

The more restrictive S-equivalence cannot as I-equivalence remove vanishing flow. Hence, points with vanishing flow refines the classification. The intuitive key to the additional normal forms is the spatio-temporal surface of vanishing temporal derivative $T = \{x, y, t | I_t(x, y, t) = 0\}$. Since the image sequence is assumed to be differentiable, generically T will be a differentiable non self-intersecting surface.

Proposition 8 (S-normal forms of 2D optic flow, codim 0)
At a fixed time-slice $t = t_0$ the normal flow is generically in an open spatio-temporal neighbourhood round any point S-equivalent to one of the following normal forms:

$$\begin{aligned} n_0(x,y) &= (1,0)^T \\ m_1(x,y) &= (x-t,0)^T \\ m_2(x,y) &= (x-(y+t)t,0)^T \\ n_2(x,y) &= \frac{1}{x^2+y^2}\begin{pmatrix} \pm x \\ \pm y \end{pmatrix} \end{aligned}$$

where the sign combinations in n_2: $(+,-)$ and $(-,+)$ are equivalent.

Proof 5
The iso-surface is defined everywhere since at a fixed time the spatio-temporal image gradient is not generically zero. For non-vanishing temporal derivative we arrive at n_0 and n_2 for regular respectively critical spatial points. m_1 or m_2 occurs for $I_t = 0$. In a spatial coordinate system (v,w) where w is the image gradient direction, we find the parameters of the diffeomorphism such that the normal flow takes the form of m_1. This form of the diffeomorphism is only valid whenever $I_{tt} \neq 0$. For $I_{tt} = 0$ we find m_2. Examples of the computation can be found in appendix B.

The codimension for a particular form determines the dimension of the set with points equivalent to the form. Hence, n_0, m_1, m_2 and n_2 points group in manifolds of dimension two, one, zero and zero, respectively.

The normal form m_1 shows a line of zero normal flow, denoted a fixed line. In terms of the spatio-temporal iso-surface, the normal form m_1 applies when the tangent plane is perpendicular to the spatial axis. Notice that S-equivalence do not distinguish attracting and repelling lines. m_2 counts for that the fixed line rotates locally, and we denote this event a "whirl". In terms of the spatio-temporal iso-surfaces, the normal form m_2 corresponds to an inflection point.

Proposition 9 (S-normal forms of 2D optic flow, codim 1)
At a fixed time-slice $t = t_0$, the normal flow in a generic one-parameter family is in an open spatio-temporal neighbourhood round any point S-equivalent to one of the normal forms of Prop. 8, Prop. 6, or the following:

$$\begin{aligned} m_{1+2}(x,y) &= (\pm y^2 \pm x^2 + t, 0)^T \\ m_{2+1}(x,y) &= (x \pm (y^2+t)t, 0)^T \\ m_3(x,y) &= (x-(y+t^2)t, 0)^T \end{aligned}$$

where the sign combinations in m_{1+2}: $(+,-)$ and $(-,+)$ are equivalent.

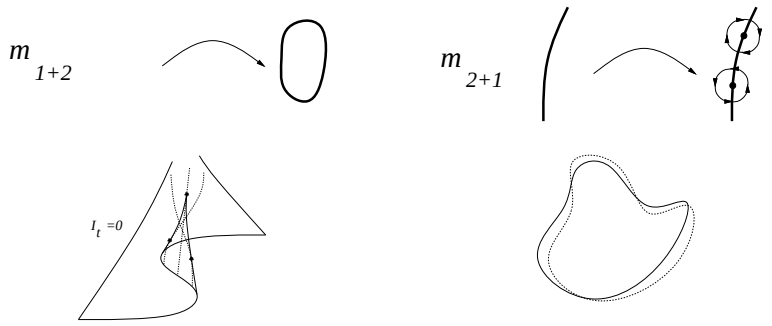


Figure 8.3: **Top**, Two events of codimension 1: m_{1+2} and m_{2+1}. **Bottom**, left is the stationary fixed line whirl. Time is vertical. It corresponds to the cusp point of the surface of $I_t = 0$. The dashed lines are the fixed line at different time instances (horizontal planes cutting the surface). The dots are the corresponding whirls. They move on a parabola open in the direction towards the reader. Right is an illustration of the principle that a perturbed curve crosses the original curve an even number of times, and equally many times from inside as from outside.

Proof 6
m_{1+2} follows from m_1 when also the spatial gradient of I_t vanishes. $m_{2+1}(x, y)$ follows from m_2 with the additional constraint that the spatio-temporal line of a whirl is locally orthogonal to the temporal dimension. m_3 follows from m_2 when $I_{ttt} = 0$. Examples of the algebraic derivations are presented in appendix B

The event m_{1+2} is, for our purposes, the most important event arising from the S-equivalence next to m_1, if one is interested solely in the fixed lines. The latter describes that the normal flow vanishes at lines. m_{1+2} describes topology change of zero flow lines, denoted fixed lines. Depending on the signs, it is either a creation event $(-,-)$, an annihilation event $(+,+)$ or a fixed line saddle event $(+,-)$ or $(-,+)$. During an annihilation or creation event a circular zero flow line vanishes/appears. During the fixed line saddle event the connectivity of two zero flow lines changes. Four incoming lines meet in a cross exactly during the event. Before and after two different pairs of incoming lines are connected.

The event m_{2+1} describes the annihilation/creation of a pair of whirls. The two whirls will have opposite rotation directions. In a point they meet and annihilate. Even though m_2 does not distinguish the rotation direction, since two whirls of opposite rotation are S-equivalent, m_{2+1} constrains the whirls to having opposite rotation since the diffeomorphism can only change direction for both simultaneously. In Figure 8.3 top are m_{1+2} and m_{2+1} illustrated.

m_3 accounts for a locally stationary whirl. This point corresponds to a cusp in the function surface $I_t = 0$. Figure 8.3, bottom-left illustrates this. Going through the cusp, does not in codimension 1 change the rotation direction of the whirl. Whirls may change direction, but this event is not singled out since it is S-equivalent to the whirl itself. If one is interested in orientation of whirls another equivalence class must be constructed to subserve this analysis. However, what one can say directly is that in the real spatial plane, there will always be equally many left and right whirls. This holds for any subset of the plane where all fixed lines form closed curves, since a curve will after an infinitesimal perturbation cross the un-perturbed curve an even number of times, and these crossings will be equally many outwards and inwards crossing. This is illustrated in Figure 8.3 (right).

The S-equivalence implies that on top of the poles, points of zero motion is the basis of the taxonomy of image sequence structure. The S-equivalence first picks up lines of zero flow, fixed lines. Then points where these lines do not move (whirls) and points where the fixed lines changes topology. It does not distinguish attracting and repelling lines.

8.10 A comment on the multi-scale optic flow structure

A scale space is constructed by convolving the image by Gaussians so that the scale-space fulfils the Heat equation $I_s = \triangle I$, where s is the scale parameter and $\triangle$ denotes the spatial, the temporal, or the spatio-temporal Laplacian dependent on which scale-space one constructs. The analysis of structure in scale-space can not be done by simply using transversality arguments and referring to Thom's classification. The proper analysis has been performed by Damon[23]. In conjunction with flow, we believe however that the heat equation does not constrain the spatio-temporal iso-surface in its local deformation, only in its topology changes. The idea is that the second derivative across the surface, may make the surface evolve in any direction in its jet space. The importance of this is that we can argue of genericity simply by counting constraints on the iso-function. If the heat equation does not locally constrain the spatio-temporal iso-surface then in all the above normal forms, time may be exchanged with scale. The only limitation is that the time and the scale parameter may not coincide.

In general, intensities at larger scales will not be preserved since intensities will change weight under the Gaussian aperture functions due to the flow. This is treated in detail by Florack et al. [30]. The above normal forms are

though still valid as they deal with the infinite resolution flow, but argued from a scale-space point of view, they can never be assessed.

8.11 Detection of structural changes

The change of topology in the poles or lines of zero normal flow is characterised by the corresponding normal forms. Thus, to detect change in the structure of the flow field we must detect when and where the normal forms apply. Table 8.1 list the conditions for the events and name the events. The conditions arise from the proofs for the propositions on I- and S-normal forms as necessary conditions on the image structure for bringing the flow at the correspond normal form under the given equivalence. Please see the proofs and appendix B for further details.

We compute derivatives of digital images as scale space derivatives. That is, we observe the image under a Gaussian aperture defining the spatial and temporal scale (inverse resolution). By differentiation of this spatio-temporal Gaussian prior to convolution, the computation of image derivatives is well-posed. The side effect is that it is not the image at grid resolution but at a lower resolution which is the object of analysis. We do not in this thesis take into account the aspects due to the non-commutation of the Gaussian convolution and the deformation due to flow field. These effects have been analysed by Florack et al. [30].

The zero locus of pre-computed differential expression is computed using an algorithm similar to the Marching Cubes algorithm [74]. For each differential expression, the zero locus is computed, and the intersection of loci is computed using an algorithm resembling the Marching Lines algorithm [126]. In this way the normal flow events are detected and their spatio-temporal position simultaneously computed.

In the following we detect some of these in two different image sequences. First, we detect the poles and their temporal interaction in a sequence of a person walking in a hallway. Secondly, we detect the lines of fixed flow and their interaction in a sequence of turbulent flow, and use the scale interaction for quantifying the amount of turbulence in the sequence.

8.12 Temporal pole evolution

Figure 8.4 illustrates the detection of flow poles and their temporal interaction in a sequence of a person appearing in a hallway. We see poles due to

n_0	Regular point	$I_x \neq 0 \vee I_y \neq 0$
n_2	Pole	$I_x = I_y = 0$
n_{2+1}	Pole stability reversion	$I_x = I_y = 0$, $I_t = 0$
n_3	Pole pair creation	$I_x = I_y = 0$, $I_{xx}I_{yy} - I_{xy}^2 = 0$
n_{2+2}	Pole stability fold	$I_x = I_y = I_t = 0$, $I_{tt} = 0$
n_{3+1}	Pole scatter	$I_x = I_y = I_t = 0$, $I_{xx}I_{yy} - I_{xy}^2 = 0$
n_4	Pole pitchfork bifurcation	$I_x = I_y = I_{xx}I_{yy} - I_{xy}^2 = I_{ppp} = 0$
m_1	Fixed line	$I_t = 0$
m_2	Fixed line whirl	$I_t = 0$, $I_{tt} = 0$
m_{1+2}	Fixed line creation	$I_t = 0$, $I_{xt} = 0$, $I_{yt} = 0$
m_{2+1}	Fixed line whirl creation	$I_t = 0$, $I_{tt} = 0$, $I_{ty}I_{ttx} - I_{tx}I_{tty} = 0$
m_3	Stationary fixed line whirl	$I_t = 0$, $I_{tt} = 0$, $I_{ttt} = 0$

Table 8.1: Symbol and name for the generic events and the condition for their occurrence. In the table, the sign p denotes the direction in which the spatial second order image structure vanishes. The connection between the conditions and the normal forms is elaborated in appendix B

critical image points at the scale at hand. These are distributed all over the image, and most have close to constant positions. However, in the center region, where the person appears in the hallway we see poles created and annihilated. These points correspond to points where the topology of the flow pattern changes. That is, these creation/annihilation points are invariant to any additional flow added to the normal flow. In this way they are not influenced by quantitative aspects such as speed, orientation etc. We suggest that they may be used for guiding an attention mechanism.

8.13 Fixed line scale-evolution

In turbulent flow, the degree of turbulence can be assessed through the scaling properties of the "eddies". Kolmogorov introduced the cascade models of turbulent flow, looking at the energy transport from large scale eddies to small scale eddies [68]. Frisch introduced a variant of these called the β-model [35] where the variable of interest is the scaling properties of the space filling of eddies. The so-called structure function characterising the flow is defined in terms of the scaling exponent of the space filling of the eddies. In the following, we sketch how this can be assessed through the multi-scale optic flow structure.

At every scale a number of whirls is present. As an approximation we

assume that a whirl corresponds to an eddy and that its space filling corresponds to its area, that is its spatial scale squared. The scaling exponent of the energy as a function of scale may then be estimated from counting whirls at a number of different scales.

In Figure 8.5, smoke induced into a ventilated pigsty is shown. The smoke is illuminated by a laser scanning through a plane such that the smoke in a vertical 2D plane in the 3D pigsty is imaged. In Figure 8.5 bottom-right the scale evolution of whirls in the Pigsty sequence is shown including annihilation (and the few creation) events. As indicated above the scaling properties of the whirls may be used for assessing the degree of turbulence in the flow. From the number of detected whirls as a function of scale we find approximately that $V \propto s^{0.5}$. That is, only 70 percent of the energy is transported to whirls at the half length scale[2]. This computation is, though, based only on approximately 100 whirls and a scaling interval of a single decade. This is clearly insufficient to state that we have proven selfsimilarity or precisely computed the degree of turbulence. We have merely indicated a direction in which the structure of the optic flow field as suggested in this thesis can be used for more practical exercises.

8.14 Summary

We have introduced two equivalence classes of optic flow and derived normal forms of codimension 0, 1, and 2 (the latter only in case of I-equivalence). I-equivalence leads to a definition of structure as the poles in the flow field whereas the S-equivalence leads also to fixed points.

The major differences to normal analytical flow fields as in autonomous dynamical systems, is the presence of poles and that the tangential component of the optic flow field is undefined. Poles are generic. The only way to avoid them in a solution is to regularize the solution space and thereby confine the solutions to a more restricted space consisting solely of solutions without poles. This is the approach taken by classic computer vision algorithms [50] to compute the optic flow field. An arbitrary "gauge condition" could be imposed to fix the tangential component and in this way fixed points as in dynamical systems can be introduced. Restricting the equivalence class of flows could make the analogy to dynamical systems even larger.

We have introduced a concept of whirls. These however have a very different nature than the nodes in dynamical systems since the whirl include second order temporal structure, and the analogy to nodes is not clear.

[2]Dimensional analysis indicates a scaling of S^0, so this is not a trivial result

The natural continuation of the research presented in this thesis is to look at a gauge fixed tangential flow, and to introduce temporarily constant diffeomorphisms for definition of the equivalence of flow. In this way the only difference to dynamical systems may be the poles.

The theoretical results in this thesis has been applied to two simple examples: computation of spatio-temporal points in which the topology of the flow field changes, as a mechanism for guiding attention, and computation of the scaling properties of whirls as to characterise turbulent flow.

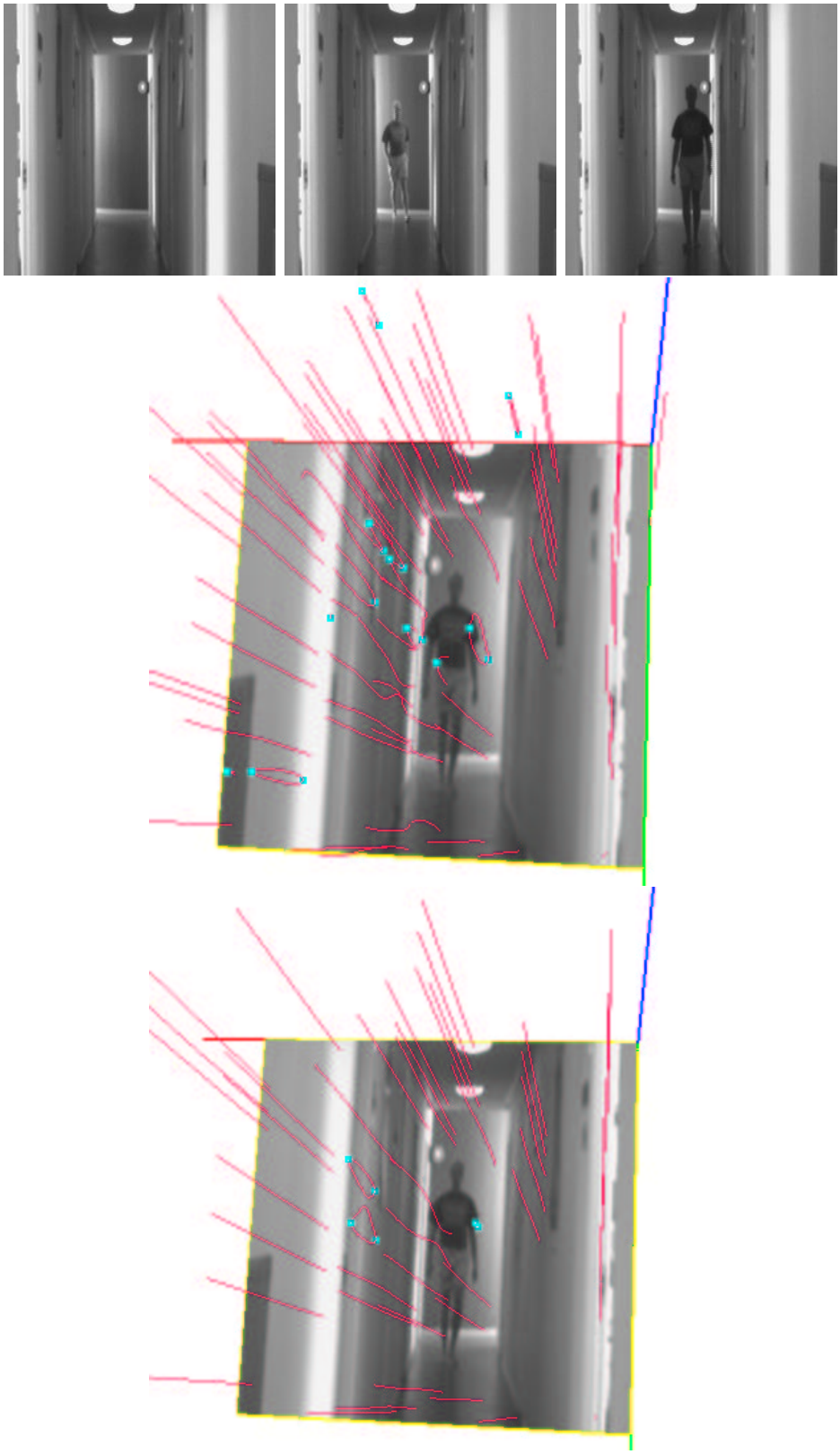


Figure 8.4: **Top**, the first, middle, and last frame of the hallway sequence. **Bottom**,The spatio-temporal curves of the poles and their creation points for spatial scale $s = 5$ (left) and $s = 8$ (right). Temporal scale is 2. Sequence is $256 \times 256 \times 32$ pixels cubed

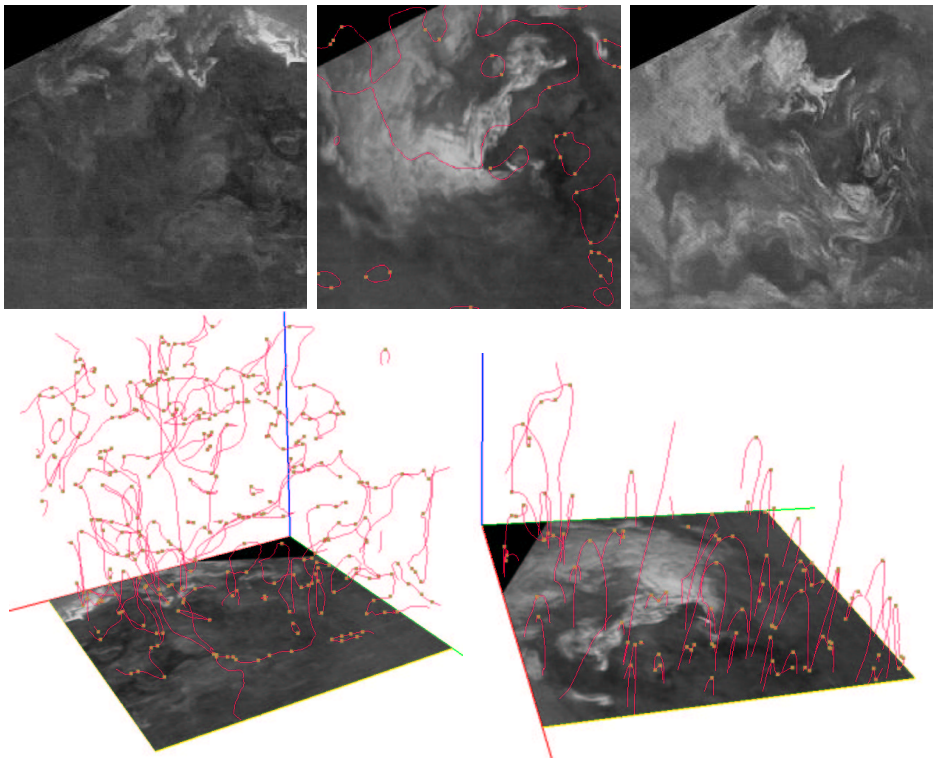


Figure 8.5: **Top**, the first, middle, and last frame of the pigsty sequence. The fixed lines and the whirls are superimposed on the middle frame. **Middle and Bottom**, In the middle is the temporal evolution of whirls and their creation/annihilation points. In the bottom are the whirls in the middle frame as a function of scale. Points mark annihilations or creations.

Part II

Segmentation using deep structure

Chapter 9

Segmentation using deep structure: Introduction

A fundamental problem in image analysis is the partition of the image scene into semantically meaningful regions. This task is usually called segmentation. Examples of tasks that often rely on segmentation are shape analysis, many medical visualisation tasks, quantitative measurements of objects and surgery planning.

The focus in this part is to build a segmentation tool based on a visual front-end like the linear Gaussian scale space. In this sense our segmentation process is uncommitted to any particular scale. The three chapters have an increasing commitment in other aspects. The common denominator is the definition of regions as catchment basins. The borders (called watersheds) between the catchment basins can not be detected locally. One could call it a semi global feature since its detection rely on a neighbourhood with a size similar to the size of the regions on the particular scale. Defining regions in this way is of course a restriction compared to no commitment at all.

In the first chapter the linking across scale between the catchment basins is described. This connection is shared by all three approaches. The linking structure is a tree-like structure where an user or modelling process can select one region by pointing to a leaf or group a collection of regions by point to a branching point of the tree (an internal node).

The first chapter gives examples for minima in the gradient magnitude squared and an example on a texture measure. The scale selection is done interactively. The second chapter is committed to the gradient magnitude as a measure of dissimilarity and the scale selection is interactive. In the third chapter a model is imposed in order to perform scale selection. The minimum description length (MDL) principle is used to select the specific internal nodes in the tree.

9.1 Automatic segmentation

Automatic algorithms for segmentation work without user interaction. Because of the complexity of the general segmentation problem the automatic methods are often dedicated to one specific task. By this restriction of the problem better solutions can be achieved by controlling the partitioning with prior knowledge of the task. The general problem is a hard nut to crack, actually segmentation is a NP-complete problem [18], however reasonable solutions may be found automatically in polynomial time. One coarse categorisation of automatic methods is : thresholding[49, 100], regional split and merge, variational & partial differential equation (PDE) based approaches[87] and combinations thereof. Chapter three falls in the last category being a combination.

Thresholding is very fast but requires that the acquisition process is tuned to have a one to one relation between intensity and objects. This is sometimes possible to achieve for instance for simple objects in an industrial setting or bone tissue in CT scannings.

Variational and PDE-based algorithms have a wide range of usage. A well known approach is the Mumford-Shah functional [87]. The ideal solution to this functional is piecewise smooth segments including a regularisation of the segment boundaries. A fast algorithm for close to global solutions for a class of segmentation functionals has been developed [135]. Well-posedness and stability has been proved for some of these functionals [117].

The split and merge approaches [49, 100, 11] start by splitting the domain into a large number of small segments based on one criterion, and thereafter merge the segments to achieve the overall goal. In our approach, the catchment basin of the gradient squared can be thought of as a splitting criteria. The merging is provided by the scale space. What is missing before one has an automatic method is a stopping criteria. Stopping the merging at an appropriate scale could handle the presence of objects at different scale in the image. Chapter three suggests a solution and thereby an automatic approach and chapter one and two let the user take the decision on when to stop.

9.2 Multi scale watershed segmentation

A lot of segmentation research has been based on watershed either on the image or the gradient magnitude. We will not try to give a complete reference list just pointers into the literature. Gradient watersheds or watershed on the image itself is being used extensively within the field of morphology

[84, 90]. Also scale space approaches to segmentation have been abundantly used [69, 133, 132, 47, 45].

Multi-scale watershed segmentation has been carried out based on the intensity and on ridges: Gauch[36] used the image intensity function directly as local measure of homogeneity. Eberly [28] defined a homogeneity measure based upon local "ridgeness". Griffin [47] used the image intensity or the image intensity gradient and based the segmentation on a multi-grid method. Also Pratikakis et al. [104] has presented an hybrid of scale space, watershed and other methods. Recently Gauch [37] has discussed the combination of scale space and watershed. The combination of watershed on the gradient magnitude image and nonlinear scale spaces has been investigated by Weickert [135] and also by Dam [21] .

The following three sections each introduces one of the next three chapters.

9.3 Multi-scale gradient magnitude watershed segmentation

In chapter 10, 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 chapter describes the multi-scale linking based on the generic events. A prototype of an interactive segmentation tool is presented together with results on synthetic and real 3D medical images.

9.4 The interactive segmentation tool

Chapter 11 presents the latest development towards a clinical use of the ideas presented previously in this thesis. The interactive segmentation tool consists of a preprocessing of the three-dimensional cells in the four-dimensional scale space of medical scannings. The linking and visualisation is partly preprocessed for fast presentation. The preprocessing and a friendly user interface facilitate use of the tool by a non expert in computer supported image analysis. The tool is being validated by regular use in the clinical research laboratory , 3D-Lab, School of dentistry, University of Copenhagen. A prototype and results were presented at the European Conference of Radiology 1999 [22].

9.5 Segmentation by compression

In chapter 12 segmentation is performed with watersheds of the gradient magnitude. Linear Scale-Space is used to generate the neighbourhood structure of the catchment basins in a tree structure, which has previously been shown to be a convenient tool for semi-automatic segmentation. This chapter investigates a fully automatic segmentation procedure using Minimum Description Length. The three tools jointly form a very useful and general algorithm, which incorporates local scale-selection and parametric descriptions. Further, the algorithm can form a basis for a large range of automatic segmentation algorithms when choosing different similarity measures and model priors.

Chapter 10

Multi-scale gradient magnitude watershed segmentation [1]

10.1 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[64]; *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 chapter, we use a Gaussian probe as a linear measurement apparatus (i.e. Gaussian convolution) and thereby introduce the Gaussian scale-space formalism [140, 66]. 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 an energy functional [86, 70]. This is computationally expensive and difficult to tune to prior information unless extensive statistical material is available [44, 19]. Also split and merge techniques[43] have been introduced. However, this strategy is captured more elegantly in multi-scale linking approaches [66, 132, 71, 36, 28].

Locally we compute a measure of dissimilarity of the image, at a certain scale. The watersheds of this measure define 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 func-

[1] This chapter has been published in a conference proceedings[98]

tion. 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 [23] and also analysed ridge measures [27], Lindeberg analysed blob detectors [73], and Rieger analysed edge and corner detectors [107]. We analysed the gradient magnitude[97].

Multi-scale watershed segmentation has been carried out based on the intensity and ridges: Gauch[36] used the image intensity function directly as local measure of homogeneity. Eberly [28] defined a homogeneity measure based upon local "ridgeness". Griffin [47] 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 [97] 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 [71]) 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 10.3. Section 10.4 describes the interactive segmentation tool. Section 10.5 presents experimental results. Finally, in Section 10.6, we summarise.

10.2 Scale-space and local dissimilarity measures

Definition 33 (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

$$m(\cdot,t) = \sum_{k\in K, x_i\in\{x,y,z\}} |\tilde{\partial}_{x_i}\hat{L}(\cdot,k,t)|^2 \quad . \tag{10.1}$$

10.3 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 34 (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 35 (Watersheds)
The watersheds are the boundaries of the catchment basins.

A property especially interesting for segmentation is the fact that *watersheds form closed hypersurfaces 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[41]. We suggest:

Definition 36 (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 [90] is to "flood" the image. Maes et al.[75] post-processed the segmentation by merging neighbouring regions using a MDL Principle. Griffin et al. [47] 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 [84, 48]. 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[97]. The duality between segments and the minima of the gradient magnitude suggests the linking scheme. A cusp is the interaction between three singularities, in

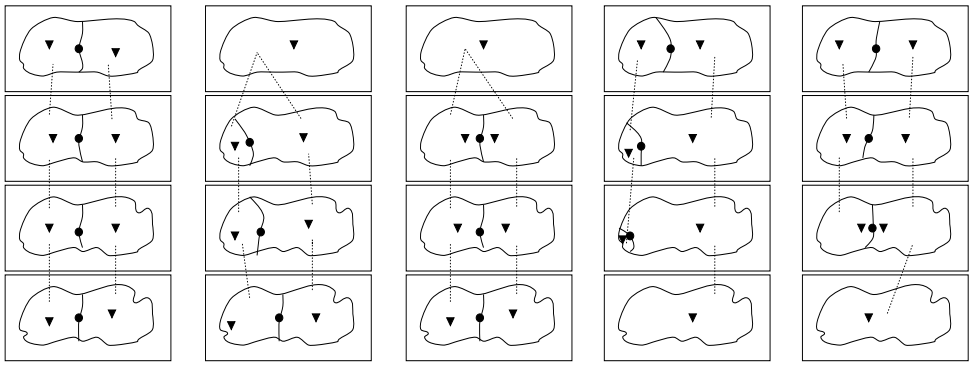


Figure 10.1: Multi-scale linking of generic events in watersheds of the gradient magnitude. The events (annihilation, merging, creation, splitting) are named after the interaction between the saddle and the minimum (or minima). Minima and saddles are symbolised with triangles and circles, respectively. A line from a segment to a segment indicates the linking.

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 10.1 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 [73] used a similar idea for blob linking.

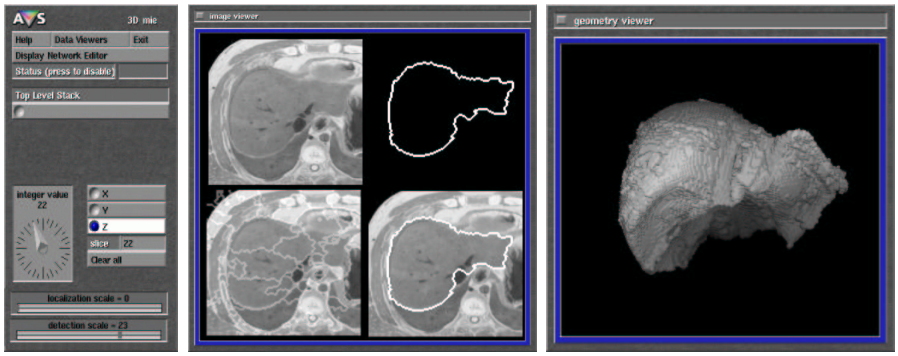
10.4 The interactive segmentation interface

A user-interface has been constructed for assessing the multi-scale segment structure (Figure 10.2). 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.

10.5 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 10.3), a CT head scan of a patient with abnormal growth (Subfigures 10.5.c,d) and digital photos (red channel) from the visible human project (Subfigures 10.5.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 10.6.

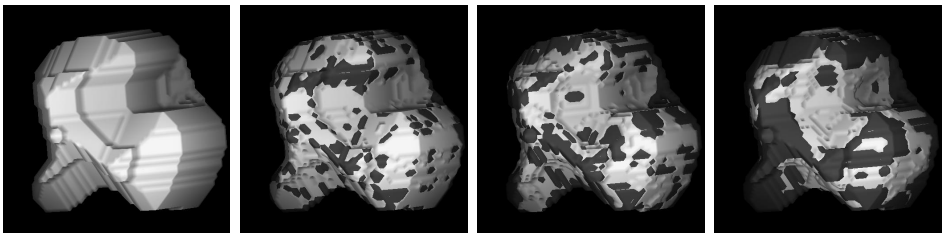
Figure 10.3 (a) displays a rendering of the true phantom. Different levels of noise has been added: respectively 0, 25, 50 and 75 percent of the voxels have been modified with Gaussian additive noise with zero mean and standard deviation of 80% of the phantom to background contrast. We shall refer to 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 de-



(a) (b) (c)

Figure 10.2: User interface windows. In window (*a*), the localisation 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.

fine the object as one single segment, which has automatically been tracked to a lower localisation scale (see Figure 10.3).



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

Figure 10.3: 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.

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 10.3).

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 localisation scale, a smooth boundary can be constructed. However, at this scale, the Gaussian blurring has deformed the object deterministically so that sharp corners (convex or concave)

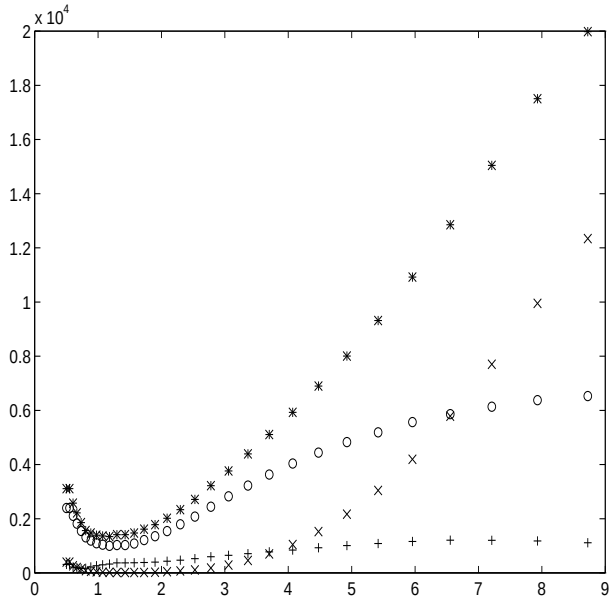


Figure 10.4: Number of erroneous voxels as a function of localization scale for noise level 25. Figure 10.3 (b) shows the segmentation corresponding to the minimum of erroneous voxels. Bold crosses indicate total number of erroneous voxels, circles indicate missing voxels on surface, crosses indicate missing interior voxels, and pluses indicate additional voxels. The sum of the "curves" given by circles, crosses and pluses equals the bold "curve". The qualitative shape of the curves are similar for the other noise levels. The optimal localisation scale and the corresponding number of erroneous voxels for different noise levels are summarised in table 10.1

are rounded. An optimal scale may be established from prior information on noise level, object size, etc. This in done empirically in Figure 10.4. The statistics for the different noise levels are summarised in table 10.1. In Figure 10.3 (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.

10.6 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 localisation scale provides a fast way of

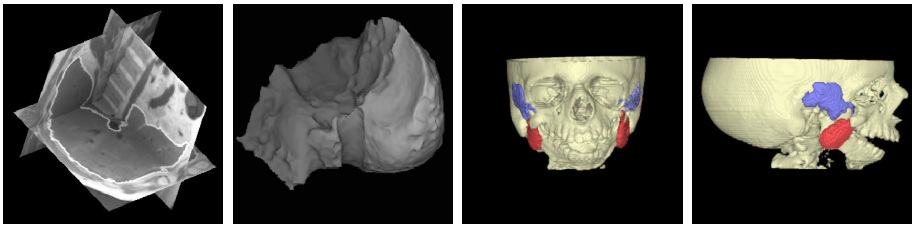
Noise level:	0 %	25 %	50 %	75 %
Best localization scale (σ)	0.605	1.18	1.73	2.78
Number of wrong voxels	79	1342	2830	7121
Wrong voxels / size of phantom	0.0014	0.0233	0.0490	0.1234

Table 10.1: Figure 10.3 shows renderings of the best segmentation result (with one segment selected at detection scale) for different levels of noise. Figure 10.4 plots the number of erroneous voxels as a function of localisation scale for noise level 25. The position and the value of the minimum from Figure 10.4 and similar plots for the other noise levels are summarised in this table.

doing semi–automatic segmentation.

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.



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

Figure 10.5: **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.

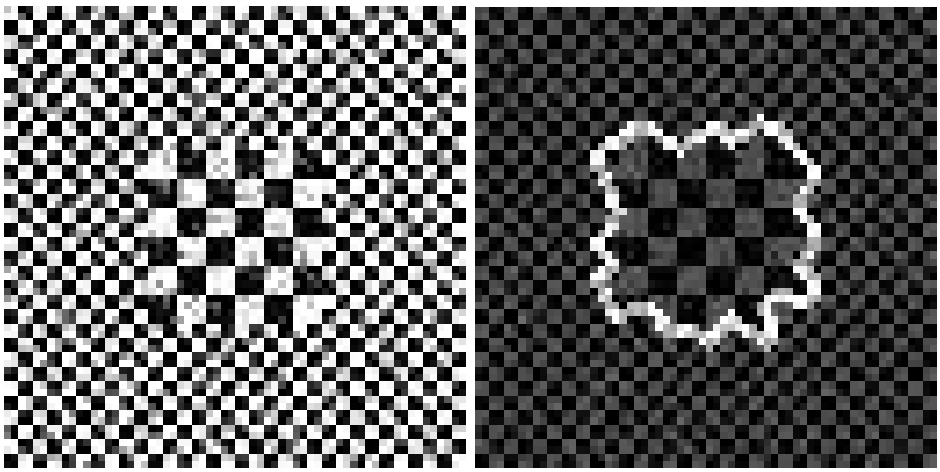


Figure 10.6: Multi-scale texture segmentation based on local frequency differences defined by Gabor functions. Two distinctive textures with Gaussian noise added is segmented using the dissimilarity measure defined in Eq. 10.1. These are preliminary results serving only as indication of the generality of the multi-scale watershed segmentation approach.

Chapter 11

The interactive segmentation tool [1]

This chapter is meant as a short illustration of the practical use of image analysis tools based on scale space theory. After preprocessing of the scale space and the linking structure the application provides real time interaction. The user interacts either directly in a visualisation of the data or in two-dimensional slices through the data cube. By providing a front end with access to the image data at all scales and a flexible geometric structure for segmentation the user has a tool with flexibility as drawing by hand but up to an order of magnitude faster.

The main menu is designed for ease of use (Figure 11.1). Visualisation is done by dragging a loaded object to a viewing window where interaction including selecting segments is possible. Two kinds of viewing windows are provided: a two-dimensional slice through the three-dimensional data and a visualisation of the three dimensional data. Figure 11.2 shows a CT scanning of the abdominal region. The user selects a working scale which can be altered any time. The application responds by presenting the preprocessed geometric structures at the given scale of selection. For a three-dimensional scanning the data domain is fully divided into cells; in the 2D view the intersection between a plane and these cells is presented.

Figure 11.3 displays a visualisation of the same data set and segments as in Figure 11.2. Here the user gets the impression of three-dimensional feedback when rotating, moving and rescaling the objects. Selection of segments is also immediate and direct in this window. Both in 2D and 3D windows the user effectively selects a three-dimensional cell in the data space. Hence, glu-

[1] The software presented in this chapter has been developed in collaboration with Andreas Thomsen and Erik Dam. The prototype was presented at the European Conference of Radiology [22]

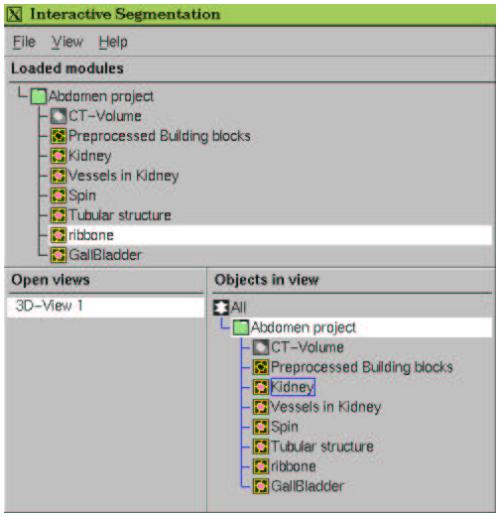


Figure 11.1: The main menu of the interactive segmentation tool. The menu provides three work areas: The top area show available data objects. Bottom left show available visualisation windows. Data objects are visualised by drag and drop to a visualisation window. Bottom right is content of a selected viewing window.

ing two dimensional segmentations together to reach the three-dimensional segments is not an issue. The figure shows three orthogonal slices through the data cube and the rendered surfaces of the selected objects. In Figure 11.4 is shown a full body PET scanning. The brain and tumour in the area of the mouth has been segmented by a few mouse clicks.

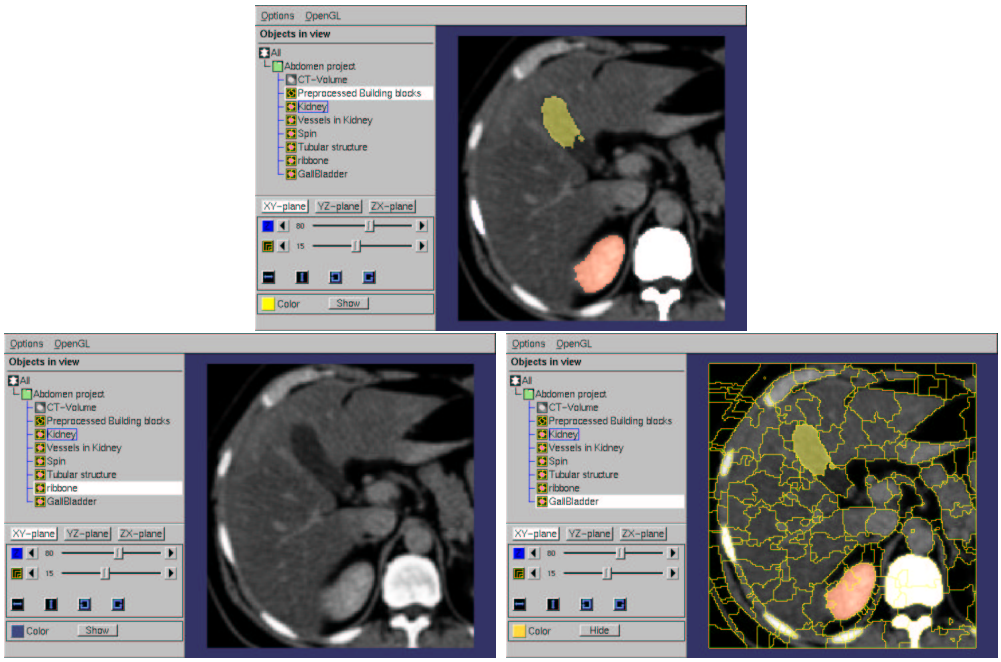


Figure 11.2: Three different settings of a two-dimensional viewing window. The data is a CT scan of the abdominal region. The segmented objects are the kidney, the spin and the gall bladder. At the top the selected anatomical objects are coloured on the raw data. Bottom left a slice of the raw data set is presented. Bottom right, the preprocessed regions are displayed as lines on the data. We thank 3D-lab, school of Dentistry, University of Copenhagen for providing the data set.

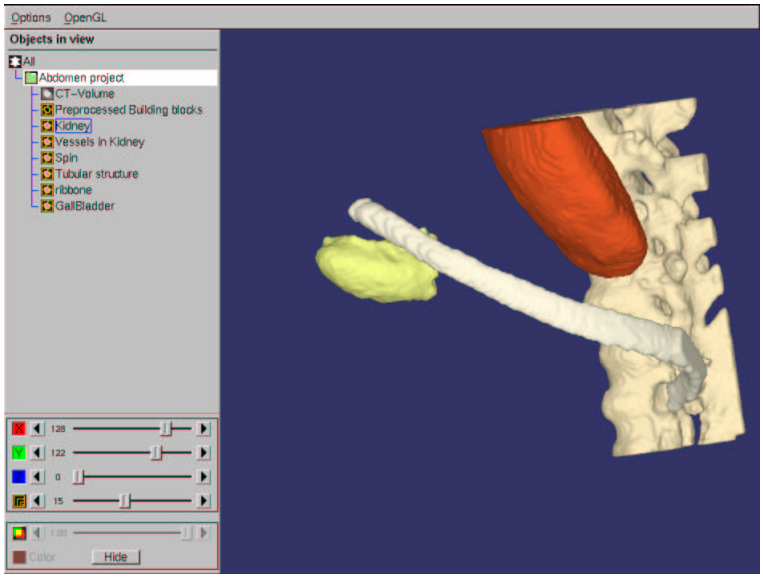


Figure 11.3: A visualisation of the abdominal region. The segmented anatomical objects are a rib, the spin, a kidney and the gall bladder. We thank 3D-lab, school of Dentistry, University of Copenhagen for providing the data set.

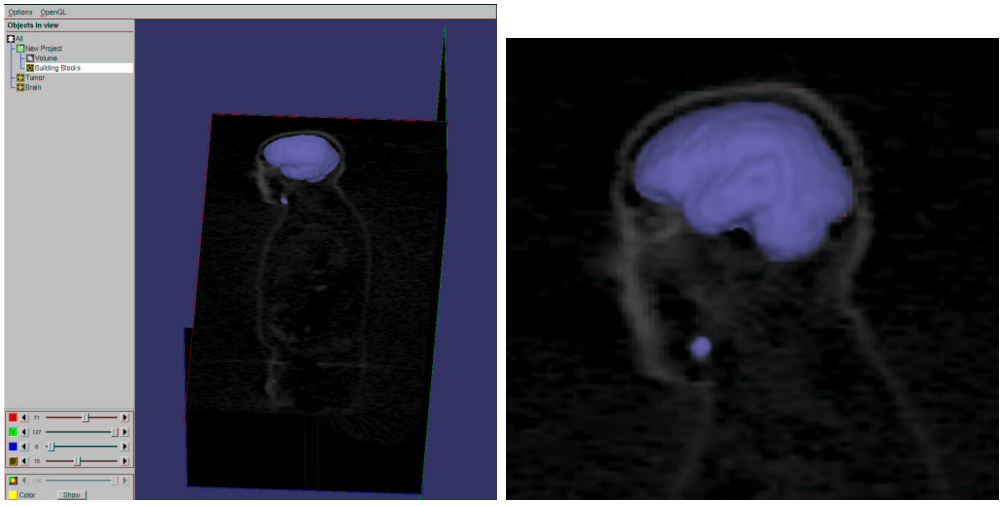


Figure 11.4: A full body PET scanning. The brain and a tumour have been segmented. The tumour is located close to the mouth. The left image show the fully body scanning; to the right is a zoom on the area of interest. We thank 3D-lab, school of Dentistry, University of Copenhagen for providing the data set.

Chapter 12

Segmentation by compression [1]

12.1 Introduction

The goal of a segmentation is to fully partition a data set into a number of non-overlapping segments. For example, a semantically meaningful segmentation of an indoor scene as in Figure 12.1(LEFT) would be piecewise smooth regions corresponding to walls, floor, ceiling, and human. Each of these areas are fairly smooth, and their borders seem to constitute discontinuities in the image. Other types of images also posse regular structure but are very much the subject of experts. E.g. in the brain in Figure 12.1(RIGHT) the location of border of the skull is readily agreed upon, but medical experts will look for more subtle clues to identify important biological structures such as a hypothetical tumour etc.. Some images are not at all smooth. These are called textured images and are not the main subject of this article.

The above examples illustrate that segmentation is strongly task dependent. This is usually solved by segmenting an image indirectly through some similarity measure, and our main interest in this article will be the gradient magnitude. For piecewise smooth images, edges are likely where the gradient magnitude is high and unlikely elsewhere. Hence, our image model is piecewise smooth images.

Generally, segmentation is a NP-complete problem [18] for two dimensional images, however reasonable solutions may be found in polynomial time. Segmentation algorithms may be divided into three broad categories: Intensity thresholding, regional split and merge, and variational and partial differential equation (PDE) based approaches.

Intensity thresholding [49, 100] is a very fast segmentation method. However, they are only applicable in situations where there is a direct relation

[1] This chapter has been published in a conference proceedings [124]

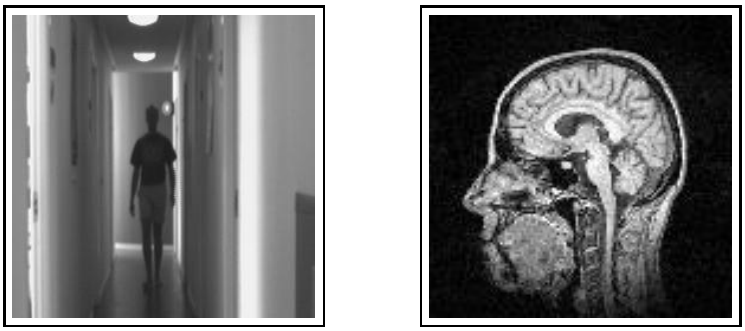


Figure 12.1: Two different images to be segmented. LEFT: An image of an indoor scene with smooth and regular structure. RIGHT: A MR slice of a human head with a different kind of structure.

between the object and its intensity. This can often be achieved in factory settings where the objects can be placed as silhouettes on a bright background.

Variational and PDE-based algorithms have a wide range of usage. For example, the Mumford-Shah functional [87] elegantly directs the solution towards piecewise smooth segments including a regularisation of the segment boundaries, and the results are relatively insensitive to noise and lighting conditions. Further, a fast algorithm has recently been discovered based on the solution of PDEs [135] that achieves a near global minimum for a large class of segmentation functionals. For some of these functionals the algorithms are guaranteed to be well-posed and stable [117]. Segmentations based on PDEs use a non-parametric representation of the segments.

Split and merge algorithms [49, 100, 11] are perhaps the most versatile class of segmentation algorithms. They can either use a parametric or non-parametric representation of segments. The basic idea is to split an image into a large number of small segments based on one criterion, and thereafter merge the segments to achieve the overall goal, which is often different from the splitting criterion. Global optimum has been shown to be intractable [18]: An optimal merging of n segments will have to consider $2^n - 1$ subsets, and is at least exponential in computation complexity. However there exist many algorithms which quite effectively achieve reasonable local optimum [11, 122].

Finally, algorithms also exist which mix various parts of the above [94, 53, 121, 70]. This is also the topic of this article. The algorithm presented here uses a PDE based technique [94] for hierarchical splitting regions, which is based on a *well-founded*, *thoroughly studied*, and *least committed* scale

analysis [136, 47, 97]. These regions will be merged using the modelling *consistent* merging ability of Minimum Description Length (MDL) [109] to find *parametric descriptions* of segments.

12.2 Least Committed Splitting

The gradient magnitude is an indicator for the presence of an edge and will tend to partition the image into regions bordered by edge-like structure. It is a strictly local operator thus very fast to calculate. Defining regions by the gradient magnitude can conveniently be performed by the watershed algorithm, which is only semi-local but still very fast to evaluate [29]. Particularly, it is worth noticing that in contrast to the Mumford-Shah functional, the watersheds are not restricted to intersect in T-junctions at 120 degree angles.

However, without sufficient regularisation, any edge operator will be useless, and so will any segmentation. Several authors have therefore investigated various scale-spaces [47, 97, 53, 135] and their properties with respect to the gradient magnitude and the watershed algorithm. Linear Scale-Space [52, 136] is convenient for studying the deep structure, and the combination of the gradient magnitude and the watershed algorithm has been thoroughly analysed [47, 97]. By this analysis a semi-automatic segmentation tool using the watersheds of the gradient magnitude in the Linear Scale-Space has been developed previously [94], henceforth called Olsen's segmentation tool. This is a convenient tool for many applications, where the segmentation task is user-defined. E.g. many experimental applications need not a complete and automatic system, since such systems are always based on some prior expectancy, which possibly is under development.

Olsen's segmentation tool organises segments at different scales in a hierarchical data structure, which makes it convenient to use for a splitting operation. Since Linear Scale-Space is the least committed tool for scale analysis [136], and watersheds are perhaps the simplest morphological tool for segmentation, Olsen's segmentation tool makes the least committed assumption on the gradient magnitude of the image.

12.2.1 Olsen's Segmentation Tool

In the previous chapter 10 was Olsen's segmentation tool presented [94]. This work forms the basis of the automatic segmentation algorithm to be described in the next section.

Olsen's segmentation tool calculates the watersheds [29] of the gradient

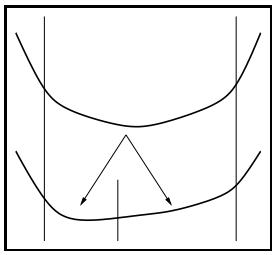


Figure 12.2: A split or merge event for the gradient magnitude. The thick lines show the signal at two different scales, the horizontal lines indicate the approximate positions of the watersheds, and the arrows show the subgraph.

magnitude of an image. This segments the image into a set of catchment basins, where there is exactly one catchment basin for each minimum of the gradient magnitude. The borders constitute a definition of edges and the basins regions of homogeneity up to second order.

When an image is evolved in Linear Scale-Space [52, 136], the evolution of the catchment basins of the gradient magnitude has a simple structure [97]: Only pairwise merging and splitting of basins are generic (probable) events. Since there is an one-to-one relation between the catchment basins and the minima the tracking algorithm only needs to keep track of the minima. However since the saddle-points play an active role in the merging and splitting, it is convenient to take them into account, when inferring the deep structure. We illustrate the basic idea in Figure 12.2, and hint on the corresponding singularity graph data structure, that stores the merging and splitting events. In practice, the number of splittings are small for the gradient magnitude, and the singularity graph can be approximated with a singularity tree. We call this the Scale-Space Tree. Although catchment basins due to splittings will be present at coarse scale, the subset not due to splittings represents the total image domain at scale zero.

Further it has been found practical to perform the tracking by mere area overlap [73], and it is certainly sufficient for sufficiently small logarithmic scale increases. We see that the principle of tracking by area overlap pragmatically extends the algorithm [94] to a large range of similarity measures. Olsen's segmentation tool is linear in the number of scales and number of image pixels: For T scale levels and N pixels the computational complexity is $\mathcal{O}(TN)$.

In Figure 12.3 are given examples of partitions at different levels in the Scale-Space Tree. Each image consists of 3 intensity values $(64, 128, 192)$ plus independent and identically distributed (i.i.d.) normal noise with zero mean

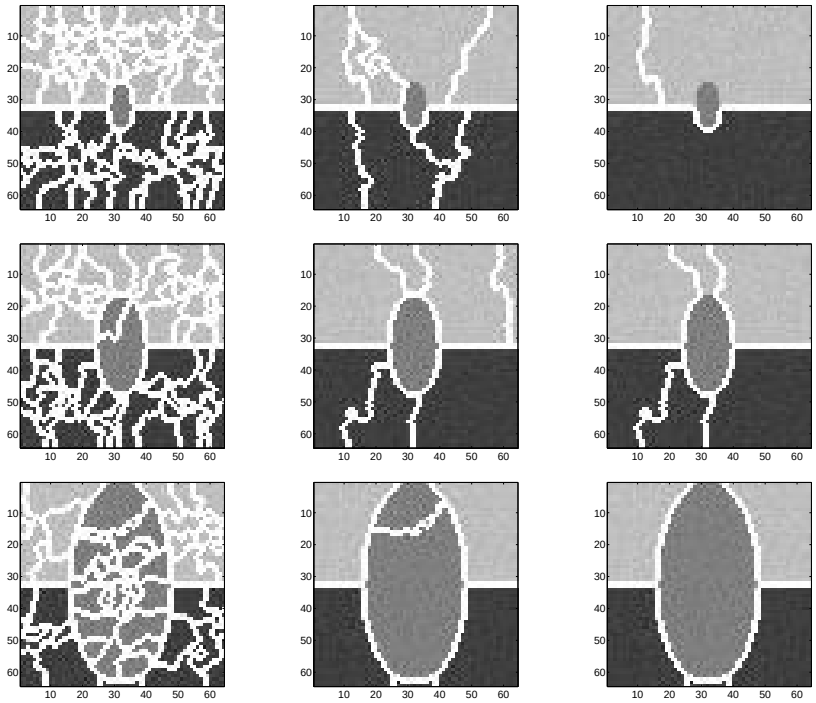


Figure 12.3: The Scale-Space Tree includes segments for ellipses of varying size. Each row shows the watersheds as white lines for the same image using integration scales low, middle, and high and measurement scale zero. The ellipse is well represented in the diagonal. The watersheds on the gradient magnitude do not regularize the segments in anyway that is left for the scale space. From the figure, it is clear that a small peak of noise in a flat area is not removed before a large scale is reached in the linear scale space

and standard deviation 5. The watersheds are the white two pixel wide lines. Each row illustrates integration scales varying from low, medium, and high scales all tracked to measurement scale zero. Note that the ellipses are one pixel further into the light than the dark area.

The integration scales have been chosen such that the watersheds are on good agreement with the ellipses of the diagonal. This illustrates the task of merging: Structure of varying size seems to be sufficiently represented in the Scale-Space Tree, and the task of the merging algorithm is to perform the local scale-selection.

12.3 Consistent Merging

For segmentation algorithms that assume piecewise constant regions, there is no need to consult advanced model selection methods. However, when models have the inclusion property, such that increasing the number of degrees of freedom monotonically decreases the distance to the data set [61], we clearly need some criterion to counter-weigh this monotonicity. An example of a function class with the inclusion property is the class of polynomials: A constant may be modelled by a first degree polynomial, which may be modelled by a second degree polynomial, and so on. Put in other words, it is a commonly known fact that too many degrees of freedom generally overfits data and does not capture the essence of the problem.

There are at present three competing model selection methods: Akaike's Information Criterion (AIC) [1, 61], Schwarz's Bayes Information Criterion (BIC) [118], and Rissanen's Minimum Description Length (MDL) [109, 110]. In the original formulation by Akaike AIC is known to be inconsistent in the sense that it will not always converge to the correct model with increasing samples. This may have been improved by Kanatani's recent work on geometric reasoning. In contrast, both BIC and MDL have been shown to be consistent and converge to each other, but MDL is special since it is the only method that is derived from a principle outside the problem of model selection in itself. Hence in contrast to AIC and BIC does MDL give a clear interpretation of the resulting model selection, as that which achieves optimal compression [10].

Due to the above considerations we will use MDL to merge segments created by Olsen's segmentation tool to yield parametric models of segments.

12.3.1 Specifying Semantics by Compression

The essence of MDL is lossless coding. That is, for the model selection criterion to be consistent, every investigated model, must include everything that is needed to completely reproduce the data set modelled. Mixing models from deterministic and stochastic domains is quite natural, since every physical signal, e.g. as produced by a CCD-chip, contains a portion of randomness. A generic MDL functional is the sum of $L(\vec{x}|\vec{\theta})$ and $L(\vec{\theta})$, where $\vec{x}$ and $\vec{\theta}$ denotes the data and the model (parameters), $L(\cdot)$ is the number of bits used to describe the model, and $L(\cdot|\cdot)$ is the number of bits used to describe the deviation from the model – the noise. Model selection is performed as,

$$\tilde{\theta} = \arg\min_{\vec{\theta}} L(\vec{x}|\vec{\theta}) + L(\vec{\theta}) \tag{12.1}$$

This is the Two-Part MDL version [109], which is apt for model selection. By the pioneering work of Shannon [120] we known that for all practical coding schemes the optimal code lengths of data are proportional to their probability as, $L = -\log P$.

A key point to note is that in almost all cases, the model parameters must be rational numbers, since almost all irrational numbers do not have a finite code length, and it is thus quite natural to study a quantisation of the parameter space with respect to the total code length: It is standard practice in physics to evaluate the precision of a parameter, however, the MDL formalism takes this practice one step further. Centred around the Maximum Likelihood estimate this implies that the needed precision is in practice inversely proportional to the second order structure of the sum in (12.1). The second order structure of an estimator around its minimum is inversely proportional to the variance of the estimator, and for almost all estimators this variance is found to be $1/|\vec{x}|$. Except for the square root we are intuitively led to the classical result of [109]:

$$\lim_{|x|\to\infty} L(\vec{x}|\bar{\vec{\theta}}) + L(\bar{\vec{\theta}}) = L(\vec{x}|\hat{\vec{\theta}}) + L(\hat{\vec{\theta}}) + \frac{|\vec{\theta}|}{2}\log n + \mathcal{O}(|\vec{\theta}|) \tag{12.2}$$

where $\bar{\vec{\theta}}$ denotes the truncated parameters, the maximum likelihood estimates are denoted by $\hat{\vec{\theta}}$ and n is the size of the data. This limit has recently been sharpened to be an $o(1)$ estimate [110, 10]. The per data point difference between the improved estimate and (12.2) can be ignored, when $|\vec{x}| \gg |\vec{\theta}|$. Hence, for computational efficiency (12.2) suffice.

Choosing a coding scheme for segmentation naturally divides into a code for the border and interior [70]. For many large segments there will be a natural tendency for the code length of the border to be diminished by the code length of the interior, because of the simple fact that the fraction of the length of the border over the area of the interior grows as $\frac{1}{|\vec{x}|}$. There is a large group of shapes, where this is not the case. E.g. for a triangle where the length of one side goes to zero, and the length over area fraction goes to infinity. However we do not expect these shapes to be typical, but expect by experience blob like segments. Therefore a simple chain code for the border will suffice. A better and model driven code for borders may be found in [122]. For the interior, the choice of model class is much more interesting. In the piecewise smooth case, low order polynomials are obviously suitable and have a nice interpretation in terms of the extended local structure, but it is unlikely that higher order polynomials are suitable. A harmonic basis is definitely also a valid possibility, and cosine waves may be versatile enough to handle both smooth regions and texture like regions. Certainly have discrete

wavelets proven their usage for compression tasks. For simplicity however, we will use the class of lower order polynomials plus i.i.d. normal noise with zero mean. We will use the centroid of a segment as the origin of the basis functions. The parameters will be coded as the universal prior of integers [109]. We thus derive the total functional for a segment as,

$$L_k = \frac{|\vec{x}|}{2}\left(\log 2\pi e + \log \sum_i \left(x_i - f(x_i, \vec{\theta})\right)^2\right) + \sum_j \log^*(\theta_j) + \frac{|\vec{\theta}|+1}{2}\log|\vec{x}| + |\partial\vec{x}| \tag{12.3}$$

where f is a function in the function class, the maximum likelihood estimate for the variance, $\sigma^2 = \sum_i \left(x_i - f(x_i, \vec{\theta})\right)^2 / |\vec{x}|$, has been used, $\log^*$ is minus the logarithm to the universal distribution, and using a 4-connected chain code who's length is given by the border, $\partial\vec{x}$, of the two dimensional segment. We have divided the code length estimate for the chain code by two, since almost all border points are used for exactly two segments. The total code length for the image is given by independence as, $L = \sum_k L_k$. The task of the merging algorithm is thus to find a minimum for L over the number and placement of segments. The computation complexity of calculating L_k for typical segments is limited by the functional fit: For an orthonormal basis, calculating each parameter is linear with respect to the number of pixels. Hence, for $|\vec{\theta}|$-dimensional orthonormal basis and a segment of $|\vec{x}|$ pixels where $|\vec{\theta}|, |\partial\vec{x}| \ll |\vec{x}|$ we need $\mathcal{O}(|\vec{\theta}||\vec{x}|)$ operations per segment.

12.3.2 A General Merging Algorithm

We will now describe a general algorithm for merging segments given by a tree. The algorithm will be similar to a previously studied merging algorithm [11], but the model class is more powerful, and hence the merging algorithm is more involved.

A general problem of the watersheds in scale-spaces is the sensitivity to noise: For constant image added with i.i.d. normal noise, the watersheds of the gradient magnitude will be present at random location. To illustrate this point, we have conducted the following experiment: Since watersheds of the gradient magnitude are stable at near high gradients, we can use an image as a source by adding i.i.d. normal noise, and averaging the results of the watersheds. An example is shown in Figure 12.4. An original image was produced as Figure 12.3(BOTTOM-LEFT). For each iteration, this image was added i.i.d. normal noise with the same standard deviation, such that the

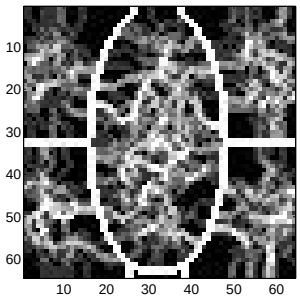


Figure 12.4: Large gradient magnitudes have stable watersheds. An image as Figure 12.3(BOTTOM-LEFT) was added noise 10 times, and the average watersheds were found.

resulting standard deviation became $\sqrt{50}$. This is a different scale analysis than the traditional scale-spaces, and we see that the watersheds due to noise are grayish. However, this does not seem to lead to an useful algorithm, since we need to know the noise source in advance. This is not the case for traditional scale-spaces. However, it is quite likely that there exists *no* linear or non-linear scale-space algorithm that does not produce random watersheds for such an image, since these scale-spaces are all smooth. In the non-linear scale-spaces studied in [135], this problem is elegantly handled by merging proportional to the intrinsic contrast parameter of the scale-spaces. However, such a solution is outside the domain of the Linear Scale-Space and must be handled by the MDL formalism.

For the following algorithm we assume that at any given scale t, the image $I(t)$ is partitioned into a number of non-overlapping segments $S_i(t)$, where image I is given by the union of the segments : $I(t) = \bigcup_i S_i(t)$ with $S_i(t) \cap S_j(t) = \emptyset$ when $i \neq j$.

Further, we will assume that there exists a linking tree T governed by inclusion, such that for a given pair of scales s, t, where $s < t$, and any segment $S_i(s)$ there will be one and only one segment $S_j(t)$ which includes $S_i(s)$. We will write inclusion as $S_i(s) \prec S_j(t)$. This implies that at any scale t we may find the representation of $S_j(t)$ as a set of segments at scale s using the tree: $T[S_j(t)] = \{S_i(s) | S_i(s) \prec S_j(t)\}$. We call $S_j(t)$ the parent or ancestor and $T[S_j(t)]$ its children or descendants. Further the elements of $T[S_j(t)]$ we will say are related as siblings, and children that have no descendants are called leaves of the tree.

To merge segments we need a cost function C. E.g. for two sibling segments $S_i(s)$ and $S_j(s)$ we may calculate $C[S_i(s)]$, $C[S_j(s)]$, and $C[S_i(s) \cup S_j(s)]$, and we will merge the two segments if the cost is decreased by the process, i.e. if $C[S_i(s) \cup S_j(s)] < C[S_i(s)] + C[S_j(s)]$. Given a cost function

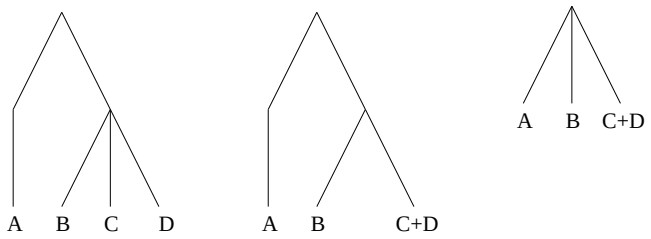


Figure 12.5: A single step of the merging algorithm. LEFT: the original tree, MIDDLE: Node A is not merged since it does not have any siblings, and B,C,D is merged into segments B and C+D (as an example). RIGHT: Immediate ancestors are replaced with the optimally merged children.

such as (12.3) we may now sketch the algorithm:

1. For all sibling tuples that also are leaves, find the merging that has the least description length, taking into account also the no merging situation.

2. For each newly merged sibling tuple replace the ancestor with the merged siblings.

3. Return to step 1 until the root has been reached.

A single iteration of the algorithm is illustrated in Figure 12.5. Since at step 1 there is no direct cross-talk between neighbouring sibling tuples, the tree *defines* a hierarchical neighbourhood structure. The merging process is restricted by this neighbourhood structure therefore the final segmentation result relies heavily on the tree structure.

This algorithm does in no way guarantee that we obtain the global minimum with respect to C. A hypothetical example, where we end in a local minimum, is given by a three-valued signal, $I = ABCA$. Assume that the global minimum is given by $A, B + C, A$ and that the tree is given by the sibling tuples (A, B) and (C, A). Then if $C(A + B) < C(A) + C(B)$ the algorithm will immediately merge A and B in first iteration and never later consider merging $B + C$. However, this is a general problem with all merge algorithms: Considering all possible merges of a given segment set is computationally intractable even for small sets of segments. Since one step of an optimal merging algorithm is to consider the cost of each element in the set of all sets of the segments, we get a lower bound on the computational complexity as 2^n where n is the number of segments. A tractable algorithm thus has to suffice with a local minimum. As in [11] we choose a greedy algorithm that merges the best two siblings iteratively. For n leaves at the lowest scale the

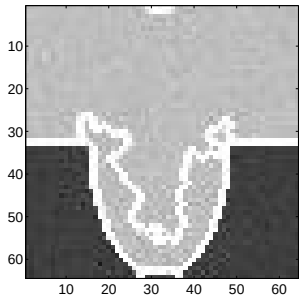


Figure 12.6: Corners generate outliers. The 'extra' segment is caused by two misplaced corner pixels.

cost function will have to be evaluated maximally $n(n-1)+(n-1)(n-2)/2$ times.

12.4 Corners Generate Outliers

The least squares fitting procedure is well suited for the normal distributed noise assumed above. However, it is very ill suited in the case of outliers, and has the worst possible break down point: It takes just a single deviating point to make the fit arbitrarily bad. Such outliers do occur for simple image structure such as corners and T-junctions, near which a few pixels may be misplaced no matter how careful the gradient magnitude is calculated. There will thus be a tendency to oversegment images due to outliers as shown in Figure 12.6.

Outliers are large scale phenomena because the statistics for the data has to be stable in order for the notion of outliers to have a meaning. However, the problem of finding outliers is similar to that of segmentation: We need a model to identify outliers. MDL is suitable for outlier detection, but it is not without problems in the present context: A relatively large area may be identified as an outlier for a small segment early in the merging procedure and will then never again be considered as an independent segment. In that sense, outlier detection counteracts the merging algorithm described in the previous sections. We will therefore apply outlier detection only in the last merge iteration in the Scale-Space Tree.

In the spirit of Least Median of Squares [114], we have implemented a method that uses 1% of a segment's pixels as test-inliers (an elemental subset), refits the model on this subset, and calculates the median of squared deviation of the inliers as a quality measure. This process is iterated till sufficient confidence, and the parameters for the subset that minimises the

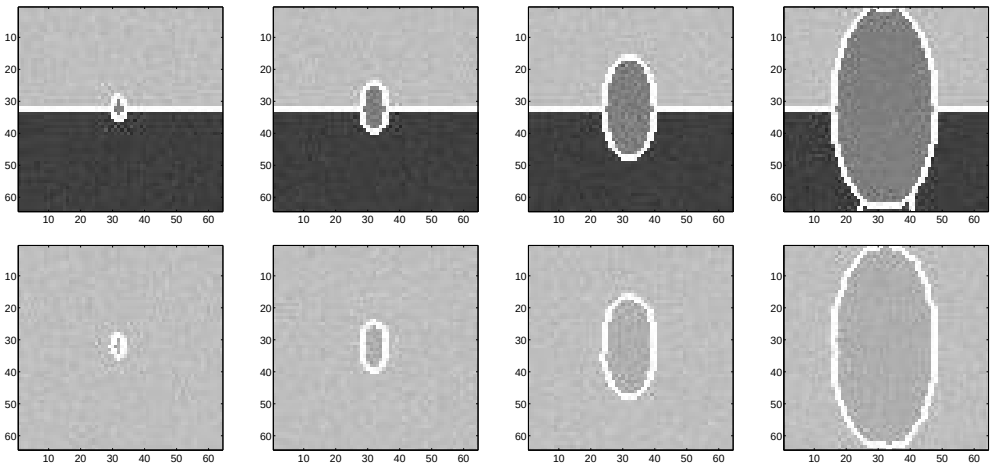


Figure 12.7: The segmentation algorithm segments simple structures. The top row shows images as Figure 12.3. The bottom row shows light dark ellipses on a lighter background (intensity values 112 and 128 respectively) added normal noise of standard deviation 5.

quality measure are used to dissect the segment into inliers and outliers.

The above yields an ordering of the pixels according to Least Median of Squares. Starting from the worst fitted pixel we may grow the set of outliers iteratively, and stop when the combined codelength of the inliers and the outliers is minimised. The inliers are coded as described previously, and the outliers are coded by their position and their value using the universal prior of integers. This has proven to be a very effective outlier detector.

12.5 Shapes in Data

Interpreting data has two basic steps: Firstly, a proper syntax must be found, which can contain all data sets to be considered. Secondly a sentence must be composed that describes a particular data set. This is the basic philosophy behind data interpretation by compression. In this article we have described a system that uses the Scale-Space Tree to define the neighbourhood structure of regions, and we have given an algorithm that seeks the particular combination of neighbourhoods that reduces the description length according to a prespecified preference. In Figure 12.7 are shown several examples of what this preference yields on very simple image structure. With these results we see that the structures captured at various levels of the Scale-Space Tree can be collected such that the basic properties of an image are respected. These

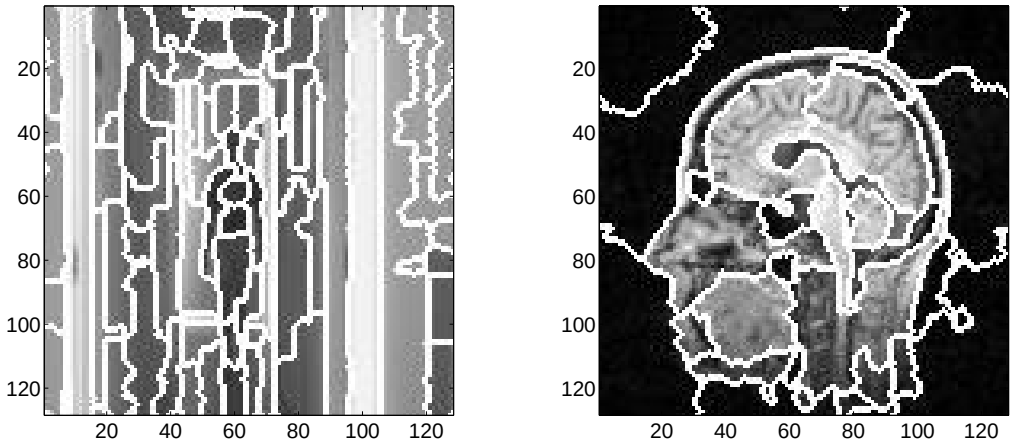


Figure 12.8: The segmentation of two different types of images. The images are 128 times 128 pixels. The linking tree was generated by sampling in the scale direction 30 times starting at 0.5 pixels and ending 32 pixels using exponential steps in between.

results are further remarkable in the smallness of detail detectable both with respect to size and intensity difference. This illustrates that the watersheds of the gradient magnitude in the Linear Scale-Space do capture a wide variety of image structures, and it seems that the algorithm presented in this chapter correctly deciphers the tree for these simple structures.

On more complex images such as shown in Figure 12.8 the algorithm displays a range of behaviours. It is difficult if not impossible to obtain the ground truth or the 'correct' segmentation of such images. However, we note that the algorithm does distinguish a number of significant regions.

This ends a preliminary study of a very general algorithm. Our future work will concentrate on the obvious and simple extensions with a range of scale-spaces such as non-linear diffusion scale-spaces and morphological scale-spaces, and we will consider other model classes such as harmonic basis' and Markov random fields. A regularisation of the segment boundaries will also be the topic of future research.

Appendix A

Smoothing images creates corners

This appendix provides mathematical details of the derivations presented in chapter 6. It is elaborated how to establish that the definition of the corner detectors correspond to a manifold of codimension 2 in jet space and how the catastrophe points for the corner detectors correspond to a manifold of codimension 3.

The following functional of the image is used as basis for detecting corners where $a = 0, 1, 2, 3$.

$$\begin{aligned} C(x,y,a) &= |\nabla L(x,y)|^a \kappa(x,y) \\ &= L_w(w,v)^{a-1} L_{vv}(w,v) \\ &= (L_i L_i)^{\frac{a-3}{2}} (L_{ii} L_j L_j - L_{ij} L_i L_j) \end{aligned} \tag{A.1}$$

Here we study the type of corners defined as critical points for C. In the following we will only go through the case (a = 3). The calculated derivatives can be reused in the other cases due to the structure of C.

Let $H(x,y) = C(x,y,3)$ we have the following first and second order derivatives in tensor notation. Summation over all double index is implied.

$$H_k = L_{iik} L_j L_j + 2 L_{ii} L_j L_{jk} - L_{ijk} L_i L_j - 2 L_{ij} L_i L_{jk} \tag{A.2}$$

$$\begin{aligned} H_{kl} &= L_{iikl} L_j L_j + 2 L_{iik} L_j L_{jl} \\ &+ 2 L_{iil} L_j L_{jk} + 2 L_{ii} L_{jl} L_{jk} + 2 L_{ii} L_j L_{jkl} \\ &- L_{ijkl} L_i L_j - 2 L_{ijk} L_{il} L_j \\ &- 2 L_{ijl} L_i L_{jk} - 2 L_{ij} L_{il} L_{jk} - 2 L_{ij} L_i L_{jkl} \end{aligned} \tag{A.3}$$

In the following equations, the expressions for the first and second order

derivatives have been partly expanded in the (v, w) coordinate system where w is in the direction of the image gradient.

$$\begin{aligned} H_k &= L_w(L_{vvk}L_w + 2L_{vv}L_{wk} - 2L_{wv}L_{vk}) \\ H_{kl} &= L_{vvkl}L_w^2 + 2L_{vvk}L_wL_{wl} \\ &+ 2L_{vvl}L_wL_{wk} + 2L_{ii}L_{jl}L_{jk} + 2L_{vv}L_wL_{wkl} \\ &- 2L_{vwk}L_{vl}L_w \\ &- 2L_{wwl}L_wL_{wk} - 2L_{ij}L_{il}L_{jk} - 2L_{wv}L_wL_{vkl} \end{aligned}$$

A full expansion of the expressions for the first and second order derivatives in the (v, w) coordinate system follows:

$$\begin{aligned} H_w &= L_w(L_{vvw}L_w + 2L_{vv}L_{ww} - 2L_{wv}L_{vw}) \\ H_v &= L_w^2L_{vvv} \\ H_{ww} &= L_{vvww}L_w^2 \\ &+ L_wL_{ww}(4L_{vvw} - 2L_{www}) \\ &+ 2L_{ww}(L_{vv}L_{ww} - L_{vw}^2) \\ &+ L_w(2L_{vv}L_{www} - 4L_{vww}L_{vw}) \\ H_{vv} &= L_{vvvv}L_w^2 + 2L_wL_{wv}(L_{vvv} - L_{wwv}) \\ &+ 2L_{vv}(L_{ww}L_{vv} - L_{vw}^2) \\ H_{vw} &= L_{vvvw}L_w^2 \\ &+ 2L_{vw}(L_{vv}L_{ww} - L_{vw}^2) \\ &+ 2L_w(L_{vv}L_{wvw} + L_{vvv}L_{ww}) \\ &- 2L_wL_{wv}(L_{www} + L_{vvw}) \end{aligned}$$

The variety of interest is defined by $0 = H_w = H_v = H_{ww}H_{vv} - H_{wv}^2$. This variety in jet space corresponds to the catastrophe points for H. First assume $L_w = 0$ then the defining equations can be reduced to the following:

$$\begin{aligned} 0 &= L_w \\ 0 &= (L_{vv}L_{ww} - L_{vw}^2) \end{aligned}$$

These equations state that the first order derivatives of the image have to be zero and the determinant of the Hessian of the image has to be zero. One can solve the equations and use this locally as a smooth parametrisation. This determines a manifold of codimension three outside the point $L_w = L_{vv} = L_{ww} = L_{vw} = 0$. The first equation defines a corner point and corresponds

to a manifold of codimension 2. This is so because the equation gives linear constraints therefore the resulting variety is a vector space which is a smooth manifold.

Next assume $L_w \neq 0$ then the defining equations can be reduced to the following:

$$\begin{aligned}
0 &\neq L_w \\
0 &= L_{vvv} \qquad \text{(A.4)} \\
detL &\equiv L_{vv}L_{ww} - L_{vw}^2 \\
L_{vvw}L_w &= 2detL \qquad \text{(A.5)} \\
0 &= H_{ww}H_{vv} - H_{vw}^2 \text{ where} \qquad \text{(A.6)} \\
H_{ww} &= L_{vvww}L_w^2 \\
&+ 10L_{ww}detL - 2L_wL_{ww}L_{www} \\
&+ 2L_wL_{vv}L_{www} - 4L_wL_{vww}L_{vw} \\
H_{vv} &= L_{vvvv}L_w^2 - 2L_wL_{vw}L_{vww} + 2L_{vv}detL \\
H_{vw} &= L_{vvvw}L_w^2 - 2L_{vw}detL \\
&+ 2L_wL_{vv}L_{vww} - 2L_wL_{vw}L_{www}
\end{aligned}$$

The equation A.6 can be rewritten as follows. So the defining equations become:

$$\begin{aligned}
0 &\neq L_w \\
0 &= L_{vvv} \\
detL &\equiv L_{vv}L_{ww} - L_{vw}^2 \\
L_{vvw}L_w &= 2detL \\
0 &= aL_{www}^2 + bL_{www} + c \text{ where} \\
a &= -4L_w^2L_{vw}^2 \\
b &= 4L_wL_{vw}(L_{vvvw}L_w^2 - 2L_{vw}detL + 2L_wL_{vv}L_{vww}) \\
&+ 2L_w(L_{vv} - L_{ww})(L_{vvvv}L_w^2 - 2L_wL_{vw}L_{vww} + 2L_{vv}detL) \\
c &= (L_{vvww}L_w^2 + 10L_{ww}detL - 4L_wL_{vww}L_{vw}) \\
&\cdot (L_{vvvv}L_w^2 - 2L_wL_{vw}L_{vww} + 2L_{vv}detL) \\
&- (L_{vvvw}L_w^2 - 2L_{vw}detL + 2L_wL_{vv}L_{vww})^2
\end{aligned}$$

Locally one can make a parametrisation of the variety by solving the equations and by this parametrise some of the derivatives by the rest of derivatives.

It can be seen from the equations above that there exists solutions where L_{vvv}, L_{vvw} and L_{www} are parametrised smoothly by the rest of the derivatives. In the case of L_{vvv}, L_{vvw} and L_{www} one has to assume that $L_{vw} \neq 0$ to solve for L_{www}.

If one is only interested in the corner points and not the catastrophe the defining equations are :

$$\begin{aligned} 0 &\neq L_w \\ 0 &= L_{vvv} \\ L_{vvw}L_w &= 2(L_{vv}L_{ww} - L_{vw}^2) \end{aligned}$$

Solving for L_{vvv} and L_{vvw} yields a smooth parametrisation of the variety. Hence we have a smooth manifold of codimension 2.

Appendix B

The structure of optic flow

This appendix provides mathematical details of the derivations presented in chapter 8. It is shown how to come from the definition of the equivalence class to the normal forms. We will shall focus on the derivations for S-equivalence.

We start with a Taylor series for a general image f. Coefficients are denoted with a letter followed by three numbers which count the order for the different variables. The zeroth order derivative for f is ignored since S-equivalence allows to change the value using the diffeomorphism η.

$$\begin{aligned} f &= a100\,p + a200\,p^2 + a300\,p^3 + a010\,q + a110\,p\,q \\ &+ a210\,p^2\,q + a020\,q^2 + a120\,p\,q^2 + a030\,q^3 \\ &+ a001\,s + a101\,p\,s + a201\,p^2\,s + a011\,q\,s + a111\,p\,q\,s \\ &+ a021\,q^2\,s + a002\,s^2 + a102\,p\,s^2 + a012\,q\,s^2 + a003\,s^3 + O(4) \end{aligned} \tag{B.1}$$

Next we make Taylor expansions for the local changes of coordinates. The coordinate changes for p and q do only have mixed terms involving t but no terms only with t. This is so because of the following restriction on S-equivalence : $\forall t \quad \psi(0,t) = 0$.

$$\begin{aligned} newP &= x\,x100 + t\,x\,x101 + t^2\,x\,x102 + x^2\,x200 + t\,x^2\,x201 + x^3\,x300 \\ &+ x010\,y + t\,x011\,y + t^2\,x012\,y + x\,x110\,y + t\,x\,x111\,y + x^2\,x210\,y \\ &+ x020\,y^2 + t\,x021\,y^2 + x\,x120\,y^2 + x030\,y^3 + O(4) \end{aligned} \tag{B.2}$$

$$\begin{aligned} newQ &= y\,y010 + t\,y\,y011 + t^2\,y\,y012 + y^2\,y020 + t\,y^2\,y021 + y^3\,y030 \\ &+ x\,y100 + t\,x\,y101 + t^2\,x\,y102 + x\,y\,y110 + t\,x\,y\,y111 + x\,y^2\,y120 \\ &+ x^2\,y200 + t\,x^2\,y201 + x^2\,y\,y210 + x^3\,y300 + O(4) \end{aligned} \tag{B.3}$$

$$newS \;=\; t\,t1 + t^2\,t2 + t^3\,t3 + O(4) \tag{B.4}$$

The Jacobian $x100y010 - x010y100$ has to be greater than zero and $t1 > 0$ to fulfil the restrictions on the equivalence.

In equation B.1 we substitute p, q and s with $newP, newQ$ and $newS$ from equations B.2, B.3 and B.4. The flow equals :

$$\begin{pmatrix} u \\ v \end{pmatrix} \;=\; |\nabla s|^{-2}\nabla s = |-\frac{\nabla f}{f_t}|^{-2} - \frac{\nabla f}{f_t} = -f_t \nabla f/(\nabla f \cdot \nabla f) \tag{B.5}$$

where ∇ denotes the spatial gradient.

The job is to bring u on a normal form. The coefficients a_{ijk} are delivered by the enemy and we can change the x_{ijk}, y_{ijk} and t_i under the restrictions given by S-equivalence.

In the following we assume to have a regular point for f. Under this assumption we can assume $a010 = 0$ and $a010 \neq 0$ in order to simplify the algebraic expressions.

Without further assumptions the zeroth order term for u is

$$\begin{aligned} u &= -\frac{a001\,t1\,x100}{a100(x010^2 + x100^2)} \\ v &= -\frac{a001\,t1\,x010}{a100\,(x010^2 + x100^2)} \end{aligned} \tag{B.6}$$

If $a001 \neq 0$ then by a suitable choice of the parameters $t1, x100$ and $x010$ one can reached the normal form n_0. The higher order terms of u can be eliminated by successively solving for the free parameters x_{ijk}, y_{ijk} and t_i, and by this remove the terms in u order by order.

Normal from m_1

If $a001 = 0$ one has the situation where $I_t = 0$ and we can derive the normal form $m_1 = (x - t, 0)^T$. The zeroth order terms are zero. The following equations (see B.7 and B.8) must have a solution in order to reach the normal form m_1. From the equation $u_t = -1$ one can derive that $a002 \neq 0$ must be true for a solution to exist. That is $I_{tt} \neq 0$. The solution is printed in the equations B.9. Again the higher order terms can be removed by successively solving for the free parameters x_{ijk}, y_{ijk} and t_i.

$$u_x = 1 \wedge u_y = 0 \wedge u_t = -1 \wedge v_x = 0 \wedge v_y = 0 \wedge v_t = 0 \tag{B.7}$$

$$
\begin{aligned}
u_x &= \frac{x100\ (a101\,t1\,x100 + a100\,x101 + a011\,t1\,y100)}{a100\,\left(x010^2 + x100^2\right)} \\
u_y &= \frac{x100\ (a101\,t1\,x010 + a100\,x011 + a011\,t1\,y010)}{a100\,\left(x010^2 + x100^2\right)} \\
u_t &= \frac{2\,a002\,t1^2\,x100}{a100\,(x010^2 + x100^2)} \\
v_x &= \frac{x010\ (a101\,t1\,x100 + a100\,x101 + a011\,t1\,y100)}{a100\,\left(x010^2 + x100^2\right)} \\
v_y &= \frac{x010\ (a101\,t1\,x010 + a100\,x011 + a011\,t1\,y010)}{a100\,\left(x010^2 + x100^2\right)} \\
v_t &= \frac{2\,a002\,t1^2\,x010}{a100\,(x010^2 + x100^2)}
\end{aligned}
\tag{B.8}
$$

$$
\begin{aligned}
x010 &= 0 \\
x101 &= -\left(\frac{t1\,(2\,a002\,a100\,t1 + 2\,a002\,a101\,t1^2 + a011\,a100\,y100)}{a100^2}\right) \\
x011 &= -\left(\frac{a011\,t1\,y010}{a100}\right) \\
x100 &= \frac{2\,a002\,t1^2}{a100}
\end{aligned}
\tag{B.9}
$$

Normal form m_2

Next we analyse if the above assumption $I_{tt} \neq 0$ is not true. Hence, we assume now that $I_t = 0$ and $I_{tt} = 0$. We aim for the normal form $m_2 = (x - t(y+t), 0)^T$. First we solve the following equations:

$$
u_x = 1 \wedge u_y = 0 \wedge u_t = 0 \wedge v_x = 0 \wedge v_y = 0 \wedge v_t = 0 \tag{B.10}
$$

The equations $u_t = 0$ and $v_t = 0$ are trivially true since u_t and v_t equals zero. Solving the rest gives the following bounds on the diffeomorphisms.

$$
\begin{aligned}
x010 &= 0 \\
x100 &= -\frac{a100x101 + a011t1y100}{a100 + a101t1} \\
y010 &= -\frac{a100x011}{a011t1}
\end{aligned}
\tag{B.11}
$$

We have to assume $a011 \neq 0$. Next we solve for the second order derivatives for the flow:

$$\begin{aligned} u_{yt} &= -1 \\ u_{tt} &= -2 \\ u_{ij} &= 0 \text{ otherwise} \\ v_{ij} &= 0 \end{aligned}$$

The solution restricts the parameters $x021, x102, x111, x012, x020, x201, x110$ and $x101$ as a function of the remaining parameters and the coefficients for f. The higher order formal forms m_{1+2}, m_{2+1} and m_3 can be constructed in the same manner.

Bibliography

[1] H. Akaike. Fitting autoregressive models for prediction. *Ann. Inst. Stat. Mat.*, 21:243–247, 1969.

[2] L. Alvarez, P. L. Lions, and J. M. Morel. Image selective smoothing and edge detection by nonlinear diffusion. *SIAM J.Numer.Anal.*, pages 845–866, 1992.

[3] L. Alvarez and F. Morales. Affine morphological multiscale analysis of corners and multiple junctions. *International Journal of Computer Vision*, 25:95–107, 1997.

[4] V.I. Arnol'd. Normal forms for functions, nondegenerate critical points, the Weyl groups A_n, D_n, E_n and lagrangian singularities. *Func.Anal.Appl.*, pages 254–272, 1972.

[5] V.I. Arnol'd. Wave front evolution and equivariant morse lemma. *Comm. Pure Appl. Math.*, pages 557–582, 1976.

[6] V.I. Arnol'd. *Singularity Theory - Selected papers.* Cambridge University Press, 1981. London Math. Soc. Lecture Notes 53.

[7] Vladimir I. Arnol'd. *Catastrophe Theory.* Springer Verlag, Berlin, 1984. Third Edition 1992. ISBN 0-387-54811-4.

[8] Vladimir I. Arnol'd. *Ordinary Differential Equations.* Springer-Verlag, 1984. ISBN 3-540-54813-0, 3rd ed. 1992.

[9] J. Arnspang. Notes on local determination of smooth optic flow and the translational property of first order optic flow. Technical Report 88-1, Institute of Datalogy, University of Copenhagen, Denmark, 1988.

[10] Andrew Barron, Jorma Rissanen, and Bin Yu. The minimum description length principle in coding and modeling. *IEEE Transactions on Information Theory*, 44(6):2743–2760, 1998.

[11] Jean-Marie Beaulieu and Morris Goldberg. Hierarchy in picture segmentation: A stepwise optimization approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(2):150–163, 1989.

[12] V. Berzins. Accuracy of Laplacian edge detectors. *Computer Vision, Graphics, and Image Processing*, 27:195–210, 1984.

[13] Johan Blom. *Topological and Geometrical Aspects of Image Structure.* PhD thesis, University of Utrecht, Department of Medical and Physiological Physics, Utrecht, The Netherlands, 1992.

[14] Th. Bröcker and L.Lander. *Differentiable Germs and Catastrophes.* Cambridge University Press, 1975. London Math. Soc. Lecture Notes 17.

[15] J.W. Bruce and P. Giblin. Growth, motion and 1-parameter families of symmetry sets. *Proc.Royal Soc. of Edinburgh*, pages 179–204, 1986.

[16] J.W. Bruce, P. Giblin, and C. Gibson. Symmetry sets. *Proc.Royal Soc. of Edinburgh*, pages 163–186, 1985.

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

[18] Martin C. Cooper. The tractability of segmentation and scene analysis. *International Journal of Computer Vision*, 30(1):27–42, 1998.

[19] T.F. Cootes, J. Taylor, D.H. Cooper, and J Graham. Active shape models–their training and application. *Computer Vision and Image Understanding*, 61(1):38–59, January 1995.

[20] David Cox, John Little, and Donal O'Shea. *Ideals, Varieties, and Algorithms.* Springer Verlag: New York, 1997. Second Edition: ISBN 0-387-94680-2.

[21] Erik Dam. Evaluation of diffusion schemes for watershed segmentation. Master's thesis, University of Copenhagen, 2000. Technical Report 2000/1 on http://www.diku.dk/research/techreports/2000.htm.

[22] Erik Dam, Peter Johansen, Ole Fogh Olsen, Andreas Thomsen, Tron Darvann, Andy B. Dobrzeniecki, Nuno V. Hermann, Noriyuki Kitai, Sven Kreiborg, Per Larsen, Martin Lillholm, and Mads Nielsen. Interactive multi-scale segmentation in clinical use. In *European Congress of Radiology 2000*, March 2000. Abstract and video accepted for CompuRAD.

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

[24] James Damon. Local morse theory for gaussian blurred functions. In Sporring et al. [123], chapter 12.

[25] James Damon. Generic properties of solutions to partial differential equations. *Archive for Rational Mechanics and Analysis*, 140(4):353–403, 1997.

[26] James Damon. Generic structure of two dimensional images under gaussian blurring. *SIAM Journal of Applied Math.*, 59(1):97–138, 1998.

[27] James Damon. Properties of ridges and cores for two-dimensional images. *Journal of Mathematical Imaging and Vision*, 10:163–174, 1999.

[28] David Eberly and Stephen M. Pizer. Ridge flow models for image segmentation. Technical Report TR93-056, University of North Carolina, Dept. of Computer Science, 1993.

[29] John Fairfield. Toboggan contrast enhancement for contrast segmentation. In *Proceedings of the 10th International Conference on Pattern Recognition*, pages 712–716, Atlantic City, NJ, USA, June 1990. IEEE Computer Society Press.

[30] L. M. J. Florack, W. J. Niessen, and M. Nielsen. The intrinsic structure of optic flow incorporating measurement duality. *International Journal of Computer Vision*, 18(3):263–286, 1998.

[31] Luc Florack. *The Syntactical Structure of Scalar Images*. PhD thesis, Universiteit Utrecht, 1993.

[32] Luc Florack, Bart ter Haar Romeny, Jan Koenderink, and Max Viergever. General intensity transformations and differential invariants. *Journal of Mathematical Imaging and Vision*, 4(2):171–187, 1994.

[33] Luc Florack, Bart ter Haar Romeny, Max Viergever, and Jan Koenderink. The gaussian scale-space paradigm and the multiscale local jet. *International Journal of Computer Vision*, 18:61–75, 1996.

[34] Luc M J Florack and Arjan Kuijper. The topological structure of scale-space images. *Journal of Mathematical Imaging and Vision*, 12:65–79, 2000.

[35] U. Frisch, P. Sulem, and M. Nelkin. A simple model of intermittent fully developed turbulence. *Journal of Fluid Mechanics*, 87(4):719–736, 1987.

[36] J M Gauch, W R Oliver, and S M Pizer. Multiresolution shape descriptions and their applications in medical imaging. In *IPMI 10*, 1988.

[37] John M. Gauch. Image segmentation and analysis via multiscale gradient watershed hierarchies. *IEEE Transactions on Image Processing*, 8(1):69–79, January 1999.

[38] John M. Gauch and Stephen M. Pizer. Multiresolution analysis of ridges and valleys in grey-scale images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(6):635–646, June 1993.

[39] Peter J. Giblin and Benjamin K. Kimia. On the local form and transitions of symmetry sets, medial axes and shocks. In *Proceedings of seventh international conference on computer vision*, pages 385–391, 1999.

[40] C.G. Gibson, K. Wirthmüller, A.A. du Plessis, and E.J.N. Looijenga. *Topological Stability of smooth mappings.* Springer Verlag, 1976. Lecture Notes in Mathematics 552.

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

[42] M. Golubitsky and V. Guillemin. *Stable mappings and their singularities.* Springer Verlag, 1974.

[43] Rafael C Gonzales and Richard E Woods. *Digital Image Processing.* Addison Wesley, 1993.

[44] U. Grenander, Y. Chow, and D.M. Keenan. *Hands. A Pattern Theoretic Study of Biological Shapes.* Springer Verlag, 1991.

[45] L.D. Griffin, G.P.Robinson, and A.C.F. Colchester. Multi–scale hierarchical segmentation. In *Proceedings of British Machine Vision Conference*, pages 289–298, 1993.

[46] Lewis Griffin and Alan C F Colchester. Superficial and deep structure in linear diffusion scale space: isophotes, critical points and separatices. *Image and Vision Computing*, 13(7):543–557, September 1995.

[47] Lewis Griffin, Alan C F Colchester, and G P Robinson. Scale and segmentation of grey-level images using maximum gradient paths. *Image and Vision Computing*, 10(6):389–402, July/August 1992.

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

[49] Robert Haralick and Linda G. Shapiro. Image segmentation techniques. *Computer Vision, Graphics, and Image Processing*, 29:100–132, 1985.

[50] B. Horn and B. Schunck. Determining optical flow. *Artificial Intelligence*, 23:185–203, 1981.

[51] R.A. Hummel and R. Moniot. Reconstructions from zero-crossings in scale-space. *IEEE Transaction on Acoustics, Speech, and Signal Processing*, 37(12):2111–2130, December 1989.

[52] T. Iijima. Basic theory on normalization of a pattern (in case of typical one-dimensional pattern). *Bulletin of Electrotechnical Laboratory*, 26:368–388, 1962. (in Japanese).

[53] Paul T. Jackway. Gradient watersheds in morphological scale-space. *IEEE Transactions on Image Processing*, 5(6):913–921, 1996.

[54] B. Jähne. *Spatio-Temporal Image Processing-Theory and Scientific Applications*, volume 751 of *Lecture Notes in Computer Science*. Springer, 1993.

[55] Peter Johansen. On the classification of toppoints in scale space. *Mathematical Imaging and Vision*, 4, 1994.

[56] Peter Johansen. Local analysis of image scale space. In Sporring et al. [123], chapter 10, pages 139–148.

[57] Peter Johansen, Mads Nielsen, and Ole Fogh Olsen. Singular points in one-dimensional gaussian scale space. Accept for publication in Journal of Mathematical Imaging and Vision, 2000.

[58] Peter Johansen, Stig Skelboe, Klaus Grue, and Jens Damgaard Andersen. Representing signals by their toppoints in scale space. In *Proceedings of the International Conference on Image Analysis and Pattern Recognition*, pages 215–217, 1986.

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

[60] Stiliyan N. Kalitzin, Bart M.Ter Haar Romeny, Alfons H. Salden, Peter FM Nacken, and Max A. Viergever. Topological numbers and singularities in scalar images. scale space evolutio n properties. *Journal of Mathematical Imaging and Vision*, 9:253–269, 1998.

[61] Kenichi Kanatani. Geometric information criterion for model selection. *International Journal of Computer Vision*, 26(3):171–189, 1998.

[62] Y. L. Kergosien and R. Thom. Sur les points parabolique des surfaces. Technical Report 290:705–710, C.R. Acad. Sci. Paris t., 1980.

[63] Les Kitchen and Azriel Rosenfeld. Gray-level corner detection. *Pattern Recognition Letters*, 1:95–102, 1982.

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

[65] J. J. Koenderink and A. J. van Doorn. Second order optic flow. *Journal of the Optical Society of America*, 8(2):530–538, 1992.

[66] Jan J. Koenderink. The structure of images. *Biological Cybernetics*, 50:363–370, 1984.

[67] Jan J. Koenderink and A.J. van Doorn. Dynamic shape. *Biological Cybernetics*, 53:383–396, 1986.

[68] A.N. Kolmogorov. The local structure of turbulence in incressible s viscious fluids for very large reynolds numbers. Technical report, C R Acad. Sci. USSR 30, 301, 1941.

[69] Andre Koster. *Linking Models for Multiscale Image Segmentation*. PhD thesis, Universiteit Utrecht, 1995.

[70] Yvan G. LeClerc. Constructing simple stable descriptions for image partitioning. *International Journal of Computer Vision*, 3:73–102, 1989.

[71] Lawrence M. Lifshitz and Stephen M. Pizer. A multiresolution hierachical approach to image segmentation based on intensity extrema. *pami*, 12(6):529–540, June 1990. Also in Xth IPMI 1988.

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

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

[74] William E. Lorenson and Harvey E. Cline. Marching cubes: A high resolution 3d surface reconstruction algorithm. *Computer Graphics*, 21(4), 1987.

[75] Frederik Maes, Dirk Vandermeulen, Paul Suetens, and Guy Marchal. Computer–aided interactive object delineation using an intelligent paintbrush technique. In N. Ayache, editor, *Medical computer vision, virtual reality and robotic*, pages 77–83. Springer–Verlag, 1995. LNCS 905.

[76] J. Mather. Stability of C^∞-mappings 3.: Finitely determined map germs. *Inst.Hautes Etudes Sci. Publ.Math.*, 36:127–156, 1968.

[77] J. Mather. Stability of C^∞-mappings 1.: The division theorem. *Ann.Math*, 89:89–104, 1969.

[78] J. Mather. Stability of C^∞-mappings 2.: Infinitesimal stability implies stability. *Ann.Math*, 89:254–291, 1969.

[79] J. Mather. Stability of C^∞-mappings 4. : Classification of stable germs by r-algebras. *Inst.Hautes Etudes Sci. Publ.Math.*, 37:223–248, 1969.

[80] J. Mather. Stability of C^∞-mappings 5.: Transversality. *Advances in Math.*, 4:301–336, 1970.

[81] J. Mather. Stability of C^∞-mappings 6.: The nice dimensions. *Liverpool Singularities Symposium I*, 192:207–253, 1970. Springer Lecture Notes in Math.

[82] J. Mather. How to stratify mappings and jet spaces. *Singularites d'Applications Differentiables*, 535:128–176, 1975. Springer Lecture Notes in Math.

[83] 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.

[84] Fernand Meyer. Topographic distance and watershed lines. *Signal Processing*, 38(6):113–125, July 1994.

[85] Farzin Mokhtarian and Riku Suomela. Robust image corner detection through curvature scale space. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(12):1376–1381, 1998.

[86] D Mumford and J Shah. Boundary detection by minimizing functionals. In *IEEE Conf. on CVPR*, San Francisco, 1985.

[87] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and asociated variational problems. *Comm. on Pure and Applied Mathematics*, 42, July 1989.

[88] 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.

[89] Hans-Hellmut Nagel. Displacement vectors derived from second-order intensity variations in image sequences. *Computer Vision, Graphics, and Image Processing*, 21:85–117, 1983.

[90] Laurent Najman and Michel Schmitt. Watershed of a continuos function. *Signal Processing*, 38(6):99–112, July 1994.

[91] Mads Nielsen, Peter Johansen, Ole Fogh Olsen, and Joachim Weickert. *Scale-Space Theories in Computer Vision.* Springer, 1999. Lecture Notes in Computer Science 1682.

[92] Mads Nielsen and Ole Fogh Olsen. The structure of the optic flow field. In *ECCV*, volume 1407 of *Lecture Notes in Computer Science*, pages II.271–287, Jun 1998.

[93] 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.

[94] Ole Fogh Olsen. Multi-scale watershed segmentation. In Sporring et al. [123], pages 191–200.

[95] Ole Fogh Olsen. Multi-scale segmentation of volumes. In *Øresyn Workshop*, 1999. (http://www.maths.lth.se/matematiklth/vision/oresyn99.html).

[96] 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.

[97] 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.

[98] 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.

[99] M. Otte and H. H. Nagel. Optical flow estimation: Advances and comparisons. In J.-O. Eklundh, editor, *Proc. Europ. Conf. on Computer Vision*, pages 51–60, Stockholm, Sweden, 1994.

[100] N. R. Pal and S. K Pal. A review on image segmentation techniques. *Pattern Recognition*, 26:1277–1294, 1993.

[101] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12:629–639, 1990.

[102] Ian R. Porteous. *Geometric Differentiation for the intelligence of curves and surfaces.* Cambridge University Press, 1994. ISBN 0521 39063X.

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

[104] I. E. Pratikakis, H. Saleli, and J. Cornelis. Hierarchical segmentation using dynamics of multiscale gradient watersheds. *Signal Processing*, 75:173–195, 1999.

[105] Krishnan Rangarajan, Mubarak Shah, and David Van Brackle. Optimal corner detector. *Computer Vision, Graphics, and Image Processing*, 48:230–245, 1989.

[106] J.H. Rieger. Generic properties of edges and corners on smooth grey-value surfaces. *Biological Cybernetics*, 66:497–502, 1992.

[107] J.H. Rieger. Generic evolutions of edges on families of diffused grey-value surfaces. *Journal of Mathematical Imaging and Vision*, 5(3):207–217, 1995.

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

[109] J. Rissanen. *Stochastic Complexity in Statistical Inquiry.* World Scientific, Singapore, 1989.

[110] Jorma Rissanen. Fisher information and stochastic complexity. *IEEE Transactions on Information Theory*, 42(1):40–47, January 1996.

[111] Carl Rohr. Modelling and identification of characteristic intensity variations. *Image and Vision Computing*, 10(2):66–76, 1992.

[112] Karl Rohr. Localization properties of direct corner detectors. *Journal of Mathematical Imaging and Vision*, 4:139–150, 1994.

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

[114] P. J. Rousseeuw and A. M. Leroy. *Robust Regression and Outlier Detection.* John Wiley & Sons, 1987.

[115] P. T. Saunders. *An introduction to Catastrophe Theory.* Cambridge University Press, 1980.

[116] D. Schaeffer. A regularity theorem for conservation laws. *Adv. in Math.*, 11:368–386, 1973.

[117] C. Schnörr. Unique reconstruction of piecewise smooth images by minimizing strictly convex non-quadratic functionals. *Journal of Mathematical Imaging and Vision*, 4(2):189–198, 1994.

[118] G. Schwarz. Estimating the dimension of a model. *Ann. Statist.*, 6(2):461–464, 1978.

[119] M. A. Shah and R. Jain. Detecting time-varying corners. *Computer Vision, Graphics, and Image Processing*, 28(3):345–355, 1984.

[120] Claude E. Shannon and Warren Weaver. *The Mathematical Theory of Communication.* The University of Illinois Press, Urbana, 1949.

[121] Judith F. Silverman and David B. Cooper. Bayesian clustering for unsupervised estimation of surface and texture models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(4):482–495, 1988.

[122] Jon Sporring. *Measuring and Modelling Image Structure*. PhD thesis, DIKU, Datalogisk Institut ved Københavns Universitet, Copenhagen, Denmark, 1999.

[123] Jon Sporring, Mads Nielsen, Luc Florack, and Peter Johansen, editors. *Gaussian Scale–Space*. Kluwer Academic Publishers, 1996.

[124] Jon Sporring and Ole Fogh Olsen. Segmenting by compression using linear scale-space and watersheds. In *Scale-Space Theory in Computer Vision, Proc. 2nd International Conference*, Lecture Notes in Computer Science 1682, pages 513–518, Corfu, Greece, September 1999.

[125] Jon Sporring, Ole Fogh Olsen, Mads Nielsen, and Joachim Weickert. Smoothing images creates corners. *Image and Vision Computing*, 18:261–266, 2000.

[126] Jean Philip Thirion and Alexis Gourdon. The marching lines algorithm: new results and proofs. Technical Report 1881, INRIA, 1993.

[127] R. Thom. Un lemme sur les applications differentiable. *Bol Soc Math Mexicana*, pages 59–71, 1956.

[128] R. Thom. *Stabilite Structurelle et Morphogenese*. Benjamin: Paris, 1972.

[129] R. Thom. *Structural Stability and Morphogenesis (Translated by D.H.Fowler)*. Benjamin-Addison Wesley: New York, 1975.

[130] A. Verri, F. Girosi, and V. Torre. Mathematical properties of the two-dimensional motion field: from singular points to motion parameters. *Journal of the Optical Society of America-A*, 6(5):698–712, 1989.

[131] 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.

[132] 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 *1. Conf. on Visualization in Biomedical Computing*, pages 115–122. IEEE Computer Society Press, Los Alamitos, CA, 1990.

[133] Koen Vincken. *Probabilistic Multiscale Image Segmentation by the Hyperstack*. PhD thesis, Universiteit Utrecht, 1995.

[134] J. Weickert. *Anisotropic diffusion in image processing.* Teubner Verlag, Stuttgart, 1998.

[135] J. Weickert. Efficient image segmentation using partial differential equations and morphology. Technical Report DIKU-98/10, Dept. of Computer Science, University of Copenhagen, Universitetsparken 1, DK-2100 Copenhagen, Denmark, 1998.

[136] J. Weickert, S. Ishikawa, and A. Imiya. On the history of Gaussian scale-space axiomatics. In Sporring et al. [123], chapter 4, pages 45–59.

[137] P. Werkhoven. *Visual Perception of Successive Order.* PhD thesis, Utrecht University, University of Utrecht, Dept. of Physics of Man, Princetonplein 5, 3508 TA Utrecht, the Netherlands, May 1990.

[138] H. Whitney. Local properties of analytic varieties. In *Differential and Combinatorial Topology*, pages 205–244. Princeton, 1965.

[139] H. Whitney. Tangents to an analytic variety. *Annals. of Maths.81*, pages 496–549, 1965.

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

[141] O. A. Zuniga and R. Haralick. Corner detection using the facet model. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 30–37, Arlington, VA, USA, June 1983.

Sammenfatning (in danish)

Afhandlingens titel er "Generic Image Structure". *Generisk* er et centralt begreb for afhandlingen. Generisk er i løse vendinger det samme som at være typisk. Projektet handler om hvad typisk billedstruktur er og dermed også om hvad struktur i billeder er. Fokus har været på en teoretisk forståelse og praktisk udnyttelse af den dybe struktur i skalarum.

Afhandlingen er delt i to hovedområder det første med fokus på teorien og det andet med fokus på det praktiske.

B.1 Den teoretiske del

Et skalarumsbillede er en kontinuert varierende familie af billeder. Familien genereres ud fra det optagede billede for eksempel ved at udjævne det mere og mere. Genereringen er kendt i form af en eller anden partiel differentialligning med billedet som begyndelsesbetingelse. Ofte beskrives skalarumsbilledet som en kontinuert stabel af billeder. For hvert billede i familien beregnes den samme type billedstruktur (for eksempel hjørner). Den gennemgående opgave er at bestemme de typiske ændringer i billedstrukturen, når man bevæger sig igennem familien. For eksempel kan der både opstå og forsvinde hjørner i de fleste skalarumsbilleder.

Billedstruktur beskrives i termer af differential geometriske operatorer, og til bestemmelse af *generiskhed* bruges singularitetsteori (katastrofeteori).

Indenfor billedhandling er det en almindelig metode at opfinde en eller anden differential geometrisk operator, der er god til eksempelvis at finde kanter eller hjørner eller liner. Næste skridt er at undersøge egenskaber ved resultatet fra en sådan operator. Hvilken type af resultat er generisk? Kan kanterne for eksempel krydse sig selv og så videre?

Indenfor matematikkens verden er en almindelig metode at gå udfra såkaldte ækvivalens relationer. Man starter med at definere, hvornår to lokale områder af en funktion er ens. Det sker ved, at man tillader en bestemt type af transformation mellem de to områder, nemlig en diffeomorfi (en glat en-til-en afbildning). Alle ækvivalente funktioner samles i en ækvivalens klasse,

og denne klasse defineres til at være en type af struktur. Med et endeligt antal af ækvivalens klasser kan man dække en åben og tæt delmængde af hele funktionsrummet, som består af de uendeligt differentiable funktioner. Disse resultater er postuleret og bevist af mange forskellige personer, nogle af nøglepersonerne er Thom[127, 128, 129], Mather[76, 77, 78, 79, 80, 81] og Arnold[4, 6, 7].

Det viser sig, at man kan benytte nogle af sætningerne fra den matematiske tilgang til at bevise generiske egenskaber for operatorerne fra billedbehandlings tilgangen. Denne sammenhæng behandles i kapitel 3.

De generelle sætninger er ikke umiddelbart gyldige, når man skifter fokus til en speciel mængde af funktioner. I dette tilfælde har løsninger til varmeledningsligningen speciel interesse. For at kunne udlede resultater om generiske egenskaber i dette tilfælde er det nødvendigt at bruge resultater udviklet af James Damon[23, 24, 25]. Disse resultater tillader tilsvarende ræsonnementer og konklusioner som for mængden af alle glatte funktioner.

Kapitlerne 4,5 og 6 bruger disse resultater til at udlede generiske egenskaber for gradientstørrelsen og en klasse af hjørnemål for skalarumsbilleder.

- Generiske overgange for kvadratet på gradient størrelsen gennemgåes i kapitel 4
  Et skalarum genereres med lineær diffusion $L_t(x, y; t) = L_{xx}(x, y; t) + L_{yy}(x, y; t), L(x, y; 0) = f$. Billedstrukturen er singulariteter for gradienten kvadreret: $\nabla(|\nabla L|^2) = 0$. De typiske strukturændringer er:

  1. Extremum og sadel for $|\nabla L|^2$ forsvinder eller opstår
  2. To extrema og en sadel for $|\nabla L|^2$ forsvinder eller opstår.

- Kapitel 5 relaterer maxima for gradientstørrelsen til det klassiske kantmål. Desuden knyttes singulariteterne for gradientstørrelsen til geometriske egenskaber ved billedet. Der findes to typer af singulariteter for gradientstørrelsen: (1) singularitetspunkter for billedet selv og (2) punkter, hvor andenordens strukturen for billedet degenererer i samme retning som gradienten for billedet.

- Kapitel 6 viser, at hjørner bestemt ved maxima for krumningen af konturkurver gange gradient størrelsen opløftet til en potens både kan opstå og forsvinde. En ordning af hjørnerne efter deres værdi gør det muligt at vælge de mest fremtrædende hjørner. Men kapitel 6 viser også, at de forskellige hjørnemål indenfor klassen giver forskellige ordninger.

Kapitel 7 beskriver den sædvanlige matematiske tilgang til struktur, og præsenterer desuden Damons udvidelser[23, 25] til den klassiske teori.

Kapitel 8 bestemmer den generiske struktur for det optiske flow ved at definere struktur udfra ækvivalensklasser. Dernæst opbygges listen over ækvivalensklasser, som dækker de typiske tilfælde.

B.2 Den mere praktiske del

En segmentering af et billede inddeler (i det idelle tilfælde) billedet i områder, der svarer til objekter i den verden, som billedet afspejler.

En type segmentering kan laves ved at knytte et område til hvert minimum i gradientstørrelsen. Da de typiske ændringer for minima kendes, kan ændringerne af områder udledes. Denne viden bruges til at forbinde områderne mellem nabomedlemmer i billedfamilien. Det mest udtværede medlem er nemmest at segmentere, fordi der kun er de største strukturer tilbage i billedet, mens det mindst udtværede medlem har grænserne mellem områderne bedst placeret. Via forbindelsen kan de større strukturer i billedet findes med en god placering.

Kapitel 10 præsenterer en protoype på et sådant system. Kapitel 11 kommer med en opfølgning, hvor en egentlig applikation til brug for klinikere i forskningøjemed er udviklet.

Kapitel 12 kombinerer ovenstående med en modellering af hvornår områderne hører sammen; altså at områderne faktisk burde sammenlægges til et område, selvom geometrien fanger en mindre forskel. Denne ekstra modellering gør det muligt at lave en fuldautomatisk segmenteringmetode.

Index

Algebraic Set, 26
Analytic, 18
Aperture Problem, 72
Atlas, 18
Attractor, 74

Base Set, 16
Basin of Attraction, 31
Bifurcation Set, 12

Catastrophe Theory, 10
Catchment Basin, 39
    Definition, 97
    Duality, 31
Caustics, 6
Chart, 18
Classification
    Arnold, 62
    Normal Forms, 62
Closed
    Under Operation, 17
Code
    Length, 115
Coding, 114
Condition
    Frontier, 28
    Regularity, 28
Contact
    Order, 22
Corner
    Definition, 51
    Generic Property, 12
    Generic Events, 55
    Linking, 52
    Measure, 12
    Ordering, 12
    Outliers, 119
    Tracking, 56
Curvature Evolution, 6

Damon, 2
Dense, 16
Detection
    Structure Change
        Optic Flow, 83
Determinacy, 64
Diffeomorphism, 12
    Topology Change, 61
Differentiation
    Well-Posed, 32
Dissimilarity Measure, 95
Distance Measure, 24

Edge
    Classical, 47
Equivalence
    Flow, 75
    H, 64
    Image Isophote, 75
    Image Stationary, 75
    IS, 65
    K-Right, 64
    Right, 62
Equivalent
    Taylor Series, 62, 64
Euclidean Shortening Flow, 6

Feature

Detector, 26
Flow
Analytic, 69
Aperture Problem, 72
Equivalence, 75
Motion Field, 69
Normal, 70, 72
Geometric Structure, 73
Normal Form
I-equivalent Codim 0, 76
I-equivalent Codim 1, 76
I-equivalent Codim 2, 77
S-equivalent Codim 0, 79
S-equivalent Codim 1, 80
Optic, 69
Spatial, 72
Spatio-temporal, 71
Structure of Optic, 70
Flow Lines, 45
Form
Normal, 64
Function
Distance Measure, 24

Gauge Coordinates, 34
Generic, 17
Generic Events
Gradient Squared, 39
Generic Properties
PDE's, 5
Genericity, 17
Germ, 62
Gradient Squared
Generic Events, 39
Linking, 41
Grass Fire Evolution, 6
Group
Diffeomorphic Germs, 62

Heat Equation
List of Normal Forms, 67
Homeomorphism, 61

Ideal, 63
Maximal, 63
Image, 2
Features, 2
Structure
Definition, 4
Task Dependent, 5
Image analysis, 1
Implicit Function Theorem, 19
Infinitely Differentiable, 15
Intersection
Countable Infinite, 17
Manifolds
Transversal, 21
Mappings
Transversal, 21
Non Transversal, 20
Transversal, 20, 21
Inverse Function Theorem, 19
Inverse Image
Transversal Intersection, 21
Isophotes, 45

Jacobian Matrix, 19
Jet
Extension Map, 10, 23
K, 23
Space, 10, 23
Junction
Detector, 46
Watershed, 46
Conditions, 47
Edges, 47
Geometric Description, 47

Linking, 32, 41, 91
Corner, 52
Local
Model, 33

Malgrange Preparation The., 62
Manifold
"Function", 10, 23
"Property", 10
Analytic, 18
Differentiable, 17
Smooth, 18
Topological, 18
Transversal Intersection, 21
Mappings
Transversal Intersection, 21
Mather, 2
Max Norm, 24
Maxwell Set, 12
MDL, 114
Measure
Dissimilarity, 95
Texture, 96
Zero, 19
Medial Axis, 6
Minimum Description Length, 114

Norm, 24
Normal Form, 64
Heat Equation, 67
List of, 62
Optic Flow
I-equivalent Codim 0, 76
I-equivalent Codim 1, 76
I-equivalent Codim 2, 77
S-equivalent Codim 0, 79
S-equivalent Codim 1, 80

Open, 16
Orbit, 66
Order
of Contact, 22
Outlier Detection, 119
Overfitting, 114

Partial Differential Equations
Generic Properties, 5
Perturbation, 16
Plateaus, 45
Point
Critical, 19
Non-regular, 19
Regular, 19
Pole, 72
Property
Manifold, 10

Regular
Whitney, 30
Repeller, 74
Residual, 17
Resolution
Continuum, 2
Inner, 2
Outer, 2
Right-Equivalence, 62
Ring, 63

Saddle Node, 74
Sard, 19
Scale Space, 1, 32, 92
Optic Flow, 82
Segmentation, 91
Automatic, 92
Compression, 94, 111
Multi Scale, 92
NP-Complete, 109
Over, 42
Split and Merge, 92
Thresholding, 92
Variational Approaches, 92
Watershed, 92
Separatrix, 39
Configuration, 45
Set
Algebraic, 26
Base, 16
Bifurcation, 12

Cantor, 18
Maxwell, 12
Shape, 31
Shocks, 6
Singularity Theory, 10
Classic, 2
Sink, 74
Skeletons, 6
Slope Lines, 39
Space
of Jets, 23
Taylor Coefficients, 10
Splitting Lemma, 33
Stability, 16
Deformation, 66
Global, 62
H, 66
Infinitesimal, 62
IS, 66
Strata, 26
Stratification, 15, 26
Definition, 27
Whitney, 28
Stratum, 27
Structure, 1
Change, 4
Optic Flow, 83
Task Dependent, 1, 5
Submersion, 18, 19
Surjective, 19

Thom, 2
Thom's Transversality Lemma, 24
Topology
Change, 28, 61
Regular C-Infinity, 24
Residual Set, 24
Toppoints
Reconstruction, 9
Scale Space, 9
Transversal
Intersection, 20, 21
Inverse Image, 15, 21
Manifolds, 21
Mappings, 21
Transversality, 20
Damon's Theorem, 25
Heat Equation, 24
Linear Diffusion, 24
Nonlinear PDE, 25
Regular Values, 20
Transversality Lemma
Thom, 24
Typicality, 17

Uncommitted, 91

Value
Regular, 19
Variety, 10, 26, 28
Visual Front-End, 91

Warping, 61
Watercourses, 45
Watershed, 39
Definition, 97
Segmentation, 92
Whitney
Regular, 30
Stratification, 28

Zero Crossings
Laplacian, 9