THE UNIVERSITY
of EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

- This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
- A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
- The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
- When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

Scale-Based Surface Understanding Using Diffusion Smoothing

Li-Dong Cai

EX BIBL. UNIV. EDINBURGEN.

PhD
University of Edinburgh
1990

To my mother, sister, wife and daughter

for their love, patience and supports

Acknowledgements

I would like to acknowledge gratefully those people who have contributed toward the work discussed in this thesis:

Primarily my supervisors Jim Howe, Bob Fisher, Mark Orr and John Hallam in different periods, especially, Bob Fisher, for his guidance throughout the study.

When the first draft of this thesis was written up, Mike Cameron-Jones gave valuable comments on the second and third chapters as sincerely as he did on my papers during the past years; John Hallam read carefully the third and fourth chapters and Bob Fisher made good comments on all chapters in spite of their heavy lecturing and research tasks; Hamish Taylor of Heriot-Watt University read and smoothed many chapters with great care and patience. Without their help, this thesis will not be in the present shape and reading it would be like chewing a heap of dry straws, attracting complaints even from the hungriest cow.

Thanks also go to other fellow Ph.D. students James Kwa, Andy Robertson and Chee-Kit Looi and all members of the Edinburgh Robot/Vision research group for their encouragement in academic activities, and David Wyse and others for technical assistance.

Finally, I thank the University of Edinburgh for providing a research studentship.

Abstract

Scale-Based Surface Understanding Using Diffusion Smoothing

by Li-Dong Cai

The research discussed in this thesis is concerned with surface understanding from the viewpoint of recognition-oriented, scale-related processing based on surface curvatures and diffusion smoothing. Four problems below high level visual processing are investigated:

1) 3-dimensional data smoothing using a diffusion process;
2) Behaviour of shape features across multiple scales,
3) Surface segmentation over multiple scales; and
4) Symbolic description of surface features at multiple scales.

In this thesis, the noisy data smoothing problem is treated mathematically as a boundary value problem of the diffusion equation instead of the well-known Gaussian convolution, In such a way, it provides a theoretical basis to uniformly interpret the inter-relationships amongst diffusion smoothing, Gaussian smoothing, repeated averaging and spline smoothing. It also leads to solving the problem with a numerical scheme of unconditional stability, which efficiently reduces the computational complexity and preserves the signs of curvatures along the surface boundaries.

Surface shapes are classified into eight types using the combinations of the signs of the Gaussian curvature K and mean curvature H, both of which change at different scale levels. Behaviour of surface shape features over multiple scale levels is discussed in terms of the stability of large shape features, the creation, remaining and fading of small shape features, the interaction between large and small features and the structure of behaviour of the nested shape features in the KH sign image. It provides a guidance for tracking the movement of shape features from fine to large scales and for setting up a surface shape description accordingly.

A smoothed surface is partitioned into a set of regions based on curvature sign homogeneity. Surface segmentation is posed as a problem of approximating a surface up to the degree of Gaussian and mean curvature signs using the depth data alone How to obtain feasible solutions of this under-determined problem is discussed, which includes the surface curvature sign preservation, the reason that a sculptured surface can be segmented with the KH sign image alone and the selection of basis functions of surface fitting for obtaining the KH sign image or for region growing.

A symbolic description of the segmented surface is set up at each scale level. It is composed of a dual graph and a geometrical property list for the segmented surface. The graph describes the adjacency and connectivity among different patches as the topological-invariant properties that allow some object's flexibility, whilst the geometrical property list is added to the graph as constraints that reduce uncertainty. With this organisation, a tower-like surface representation is obtained by tracking the movement of significant features of the segmented surface through different scale levels, from which a stable description can be extracted for inexact matching during object recognition.

Table of Contents

List of Figures

Chapter 1

Introduction

1.1. Problem posed

Humans effortlessly perceive the 3-D shapes of unfamiliar objects from their 2-D images on the retina of an eye. This ability of the human visual system has intrigued perceptual psychologists for decades, and more recently computer vision scientists, for two reasons. First, the loss of valuable information due to the projection of the 3-D scene onto the 2-D retina has not hindered human's perception, and second, details of objects are captured by perception while noise introduced in this information transmission is ignored. A primary goal of computer vision is to understand this phenomenon at the level of its *computational theory* [Marr 1982]

Traditionally, computer vision research is divided into two levels High-level vision is the portion of the signals-to-symbols paradigm concerned with providing a goal-oriented description of a scene in which purpose, context, and space and time relations are all made explicit. Such a research relies on special high-level information about the scene to be analysed, such as models Low-level vision is the first phase of the signals-to-symbols paradigm. Research at this level relies on general knowledge about the physical world. The processing is local and is independent of purpose or content. It outputs a description of intensity variations and anomalies in the image or the local intrinsic characteristics of the visible scene surfaces, such as depth, orientation and velocity.

Low-level vision processing includes data smoothing, surface segmentation, surface shape classification, shape from shading and photometric stereo, shape from texture and shape from motion, ... *etc.* It uses either intensity data or depth data to recover surface geometry. Research in this thesis will use the depth data which gives the distance to the visible surfaces. The processing will aim at object recognition, a high-level vision processing, but will not include it as a task. So it covers only some stages below high-level vision processing. These stages are illustrated in the box below:

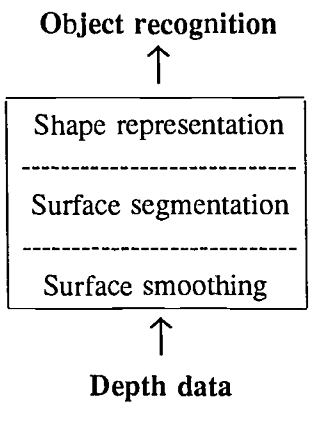


Figure 1.1.1. Visual processing stages covered by the research.

Surface smoothing is set as the first stage, since data smoothing may underpin the rest of low-level vision processing, because noise contamination of data is often a major problem even when the data is measured using a laser range finder. Diffusion smoothing, *i.e.*, smoothing with a diffusion equation, will be used to implement surface smoothing.

Computational studies and perceptional experiments [Marr 1977] suggest that low level vision processes are generic ones that correspond to conceptionally independent modules that can be studied in isolation. Surface segmentation is such an example. Although it is nearly impossible to formulate precisely in terms of the image or even the physical world what the *exact* goals of segmentation are, image segmentation for different tasks has been intensively studied by many researchers (cf. Chapter 5). Surface segmentation will be the second stage of the processing in this research to produce a set of patches rather than edges Each patch will have a uniform shape in a certain sense.

The third and the final stage of the processing in this research is to work out a shape representation as the consequences of surface segmentation. A symbolic description of the partitioned surfaces will be built up based on the geometrical shape of individual patch and the topological relationships among patches. It is desired to supply a basis for further level vision processing, such as object recognition

Considering the above three stages as a whole, it should be noted that they all involve a fundamental problem in visual processing, the scale problem. That is, surface understanding depends on the spatial scales at which the surface is measured and analysed Although with some prior knowledge of a scene and its data, low-level visual processing up to segmentation can be considered at a single scale level, when visual processing is continued further up to shape representation, a single scale treatment no longer seems enough, since the surface shape and its description will change at different scale levels. As problems related to multiple scales will be frequently involved in, scale space processing [Witkin 1983] of the 3-D objects is to be applied to all three stages.

It should also be noted that the mode of behaviour of visual processing depends on the task assigned to it. For example, recognition-oriented low-level visual processing can be distinguished from reconstruction-oriented processing although they cover the same stages (smoothing, segmentation and representation). The output of recognition-oriented processing will give precedence to those features playing a salient role in recognition, whereas the output of reconstruction-oriented processing will reproduce the surface as closely as possible without discrimination in favour of any feature.

In short, the problem studied in this thesis can be specified as understanding surfaces from the viewpoint of recognition-oriented, scale-related processing based on surface curvatures and diffusion smoothing at the level of its *computational theory*. Along with the development of this thesis, the goal of the computation, why is it appropriate, and what is the logic of the strategy by which it can be carried out will be further made clear.

1.2. Philosophy and motivation for the study

An important goal of computer vision is to understand human vision from an information processing perspective. Even though there is no proof yet that Marr's paradigm for computational vision is correct, it has been adopted in the belief that something similar should be true. With a philosophy that computer vision involves a general understanding but not an exact mimicry of human vision, seeking a similarity between them is right in principle. As a scientific tool still in progress, it seems neither possible nor necessary that a computer vision system should work identically to the human vision system.

Comparable things have happened in other branches of science. For example, flying like birds has been a dream of human beings for a long time, but making a pair of wings for human flying is still too complex to be successful. Instead, with a study of the aerodynamics of bird flight, the aeroplane has been invented and improved so that the "iron birds" can fly at much higher speeds and for far longer distances than any bird although they do not look as graceful as birds do.

Obviously, to reliably derive properties of the world from image data of it, clues from human perception are highly appreciated. However, the most important requirement here is to keep coherence in the computational theory itself and coherence within the recognition mechanism so that the theory works. In other words, the practical aim is to have an "aeroplane" to pilot rather than a pair of "wings" for flying oneself. Hence, the motivation for the research done in this thesis is to seek an appropriate bridge over the gap between depth data and a surface-based recognition system.

To achieve this aim a rigorous and broad analysis of data smoothing and the behaviour of features is fundamental to understanding vision as an information processing task. For this reason the approach will focus on the problem of surface understanding through scale space and based on curvatures. The feasibility of this approach has been demonstrated in some other studies [Witkin 1983],[Brady *et al* 1985],[Besl 1986],[Besl and Jain 1986].

Another important motivation for this research arises from its potential applications in machine vision systems. This work is directly relevant to moving industrial parts tests, 3-D surface-based recognition and human face recognition [Bruce and Burton 1989].

1.3. A computational approach

Depth data is a 2-D image with 3-D information, which gives distance to the ***visible*** surfaces in a viewer-centred coordinate frame (this is also the reason it is called the 2½-D data rather than the 3-D data). Therefore, the goal of the computational approach is to recover properties of 3-D objects from the data, construct segmented structures and finally output it in the form of a symbolic representation convenient for recognition purposes. This visual control strategy, which indicates the flow of information and activity through representational layers, is mainly a data-driven one; but adopting a part of model-based strategy is still possible [Ballard and Brown 1982]. In total, this computational approach is built around six major propositions:

1. Diffusion smoothing or smoothing with a diffusion equation is an incremental operation with respect to time, that suits a framework for continuous processing from fine scales to large scales corresponding to time.
2. The investigation of diffusion smoothing makes explicit the inter-relationships among diffusion process and other widely used smoothing methods and introduces a fast computational scheme associated with a convenient boundary treatment which can efficiently prevent surface shape distortion at the boundary.
3. Knowing that the significance of surface features and the features themselves will change across scale levels, the investigation of the scale space behaviour of surface

features in terms of curvature signs will supply theoretical guidance to the dynamic process of surface segmentation and representation.

4. Low-level visual processing starts with local depth data and yields the Gaussian and mean curvatures of surfaces. As stringent surface properties they are natural indications of different surface shapes. Such a curvature-based shape classification makes visual processing mainly data-driven and partially model-based. It is a coherent way to bridge the gap between depth data and a model-based recognition system.
5. Surface segmentation based on curvature signs using depth data alone is posed as an under-determined surface approximation problem. This shows the direction along which efforts should be made to get the feasible solutions.
6. Organisation of shape information in representation is particularly emphasised. More flexibility is brought into the shape representation due to the separation of topological properties from geometrical properties, and this favours recognition.

1.4. Organisation of the thesis

This thesis has seven chapters and several appendices. Chapter 1 introduces the argument of the thesis and Chapter 7 summarises the the results obtained, and indicates possible directions for further research Chapter 2 reviews some relevant concepts from differential geometry, where the principal curvatures and directions of surfaces are explained from the viewpoint of the eigenvalues and their eigenvectors of the mapping matrix. The other chapters cover three different stages of the low level visual processes. Essential terminologies and summarises of previous work are given at the beginning of individual chapters concerned.

Chapter 3 and Chapter 4 are the heart of this thesis. Chapter 3 addresses the theory of diffusion smoothing developed in this thesis. Eight theorems are proven. They supply a general theoretical basis to interpret Gaussian smoothing, repeated averaging and spline smoothing in a uniform fashion. A "small leakage" diffusion model is derived to deal with the difficult problem of shape distortion at the surface boundary. It enables a good quality curvature sign image to be obtained and makes the scale space processing of surfaces practical. Other applications of diffusion smoothing are also discussed, *e.g*, the design of optimal Gaussian operators in small neighbourhoods

Chapter 4 elucidates the behaviour of surface shape features in scale space in terms of curvature signs The stability of surface categories or large surface features is discussed, which leads to the formulation of a relationship between the zero thresholds of Gaussian and mean curvatures. The scale space behaviour of small surface features is observed and elucidated in the form of three theorems and three rules as the means to track the movement of single and nested features across scale levels. Based on the stability analysis and behaviour theory developed above, an explanation is given to the phenomenon of the "odd textures" remaining on the curvature sign image of a cylindrical oil bottle, especially when its depth data is poor.

Chapter 5 is an application of both diffusion smoothing and the scale behaviour theory to surface segmentation. Segmentation is approached from the viewpoint of curvature-based processing and posed as an under-determined mathematical problem of approximating surfaces up to the degree of curvature signs using depth data alone. Using the image of Gaussian and mean curvature signs to segment generic surfaces in scale space is advocated. Using the principal directions to detect cylindrical surfaces is proposed. While stress is laid on using local operations to detect surface patches, segmentation using global fitting is also discussed. Six lemmas, a theorem and two corollaries demonstrate the feasibility of approximating surfaces up to curvature signs via a recursive piecewise surface fitting with depth data alone. The essential differences between single scale and multi-scale segmentation methods are also addressed when comments are made on the extension of Besl's single scale segmentation method.

Chapter 6 addresses the organisation of a tower-like surface shape representation for recognition purposes. The advantages of separating topological properties from geometrical properties are emphasised The forms and functions of the adjacency graph and the attribute list and how they are combined in a flexibly constrained graph in order to describe a segmented surface are discussed. The concept of the flexibility of the surface representation and its formal expression are introduced. A stable, symbolic shape description is eventually extracted from the tower of flexibly constrained graphs. All these lead to an extension of Witkin's scale space smoothing from the 1-D case to an appropriate form of the 2½-D case.

Appendices A1 to A6 contain some results moved out from Chapter 3, 4 and 5 to make discussion in each chapter concentrate on its main topic and be compact as well.

1.5. Major contributions of this study

This study has developed a computational theory for the use of scale space processing of surfaces from depth data to obtain a symbolic shape representation for the purpose of recognition. The major contributions to computer vision are as follows:

1. Diffusion smoothing theory has been pushed forward a substantial step in this study, making explicit the inter-relationships among diffusion smoothing, Gaussian smoothing, repeated averaging and spline smoothing and making scale space processing practicable by a fast diffusion smoothing algorithm.
2. A "small leakage" diffusion smoothing model is proposed, which treats the external surface boundary properly It prevents surface shapes along the boundary from being distorted and produces good quality curvature sign images, by which sculptured surfaces can be segmented directly using the curvature sign image from diffusion smoothing alone.
3. The stability analysis of surface shape categories and the scale space behaviour theory of shape features in the 2½-D case provide theoretical guidance for tracking the movement of features in scale space.
4. A new organisation of shape information in representation is advocated. Topological properties of a segmented surface are separated from its geometrical properties. It leads to a scale space shape representation for the purpose of recognition, from which a stable shape description can be extracted for model matching (cf. [Winston 1977]). It is thus an extension of Witkin's 1-D scale space smoothing to the 2½-D case.

Chapter 2

A brief review of differential geometry

In this chapter, some concepts from differential geometry are reviewed. As the aim is to provide a basis for understanding later parts of this thesis, such as the sections on surface classification, segmentation and representation, the review is not going to outline the whole of differential geometry here but to describe briefly those ideas relevant to surface curvatures. More details can be found in [Chern 1967],[Do Carmo 1976] and [Faux and Pratt 1979] *etc.*

2.1. A surface, its tangent plane and normal

In differential geometry, a smooth or regular surface S, its 1st and 2nd derivative, and normal can be expressed in the form of parameterised vectors:

$$S = \left\{ \mathbf{r}(u,v) = \begin{bmatrix} x(u,v) \\ y(u,v) \\ z(u,v) \end{bmatrix} \right\}_{(u,v) \in D \subseteq R^2} \tag{2.1.1}$$

$$d\mathbf{r} = \mathbf{r}_u du + \mathbf{r}_v dv = [\,\mathbf{r}_u\ \mathbf{r}_v] \begin{bmatrix} du \\ dv \end{bmatrix} \tag{2.1.2}$$

$$d^2\mathbf{r} = \begin{bmatrix} du \\ dv \end{bmatrix}^T \begin{bmatrix} \mathbf{r}_{uu} & \mathbf{r}_{uv} \\ \mathbf{r}_{uv} & \mathbf{r}_{vv} \end{bmatrix} \begin{bmatrix} du \\ dv \end{bmatrix} \tag{2.1.3}$$

$$\mathbf{n} = \frac{\mathbf{r}_u \times \mathbf{r}_v}{|\mathbf{r}_u \times \mathbf{r}_v|} \tag{2.1.4}$$

$$\mathbf{r}_u = \begin{bmatrix} x_u \\ y_u \\ z_u \end{bmatrix}, \quad \mathbf{r}_v = \begin{bmatrix} x_v \\ y_v \\ z_v \end{bmatrix}, \quad \mathbf{r}_{uu} = \begin{bmatrix} x_{uu} \\ y_{uu} \\ z_{uu} \end{bmatrix}, \quad \mathbf{r}_{uv} = \begin{bmatrix} x_{uv} \\ y_{uv} \\ z_{uv} \end{bmatrix}, \quad \mathbf{r}_{vv} = \begin{bmatrix} x_{vv} \\ y_{vv} \\ z_{vv} \end{bmatrix} \tag{2.1.5}$$

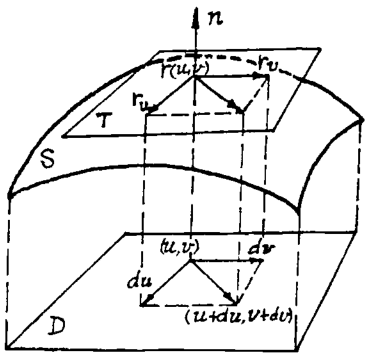


Figure 2.1 1. Parameterised surface, its tangent plane and normal.

here the vector $\mathbf{r}$ denotes a point on the surface S, $d\mathbf{r}$ the first differential of $\mathbf{r}$, $d^2\mathbf{r}$ the second differential; D is the domain of S, u and v are parameters; $\mathbf{r}_u$ and $\mathbf{r}_v$ are linear independent vectors lying in the tangent plane $T(u,v)$ of surface at the point $\mathbf{r}(u,v)$, where the three vectors $\mathbf{n}$, $\mathbf{r}_u$ and $\mathbf{r}_v$ form a local coordinate frame since the dot products of vectors $\mathbf{n}\,\mathbf{r}_u = 0$, $\mathbf{n}\cdot\mathbf{r}_v = 0$ and the mixed product of vectors $\mathbf{n}\cdot(\mathbf{r}_u \times \mathbf{r}_v) \neq 0$.

2.2. The first and second fundamental forms of a surface

The first and second fundamental forms of a regular or smooth surface are defined as the following two quadratic forms:

$$I = d\mathbf{r}\, d\mathbf{r} = \begin{bmatrix} du \\ dv \end{bmatrix}^T \begin{bmatrix} E & F \\ F & G \end{bmatrix} \begin{bmatrix} du \\ dv \end{bmatrix} \equiv \begin{bmatrix} du \\ dv \end{bmatrix}^T \mathbf{A} \begin{bmatrix} du \\ dv \end{bmatrix} > 0 \tag{2.2.1}$$

$$II = \mathbf{n}\cdot d^2\mathbf{r} = \begin{bmatrix} du \\ dv \end{bmatrix}^T \begin{bmatrix} L & M \\ M & N \end{bmatrix} \begin{bmatrix} du \\ dv \end{bmatrix} \equiv \begin{bmatrix} du \\ dv \end{bmatrix}^T \mathbf{B} \begin{bmatrix} du \\ dv \end{bmatrix} \tag{2.2.2}$$

Where "·" denotes the dot product of two vectors, and

$$\mathbf{A} = \begin{bmatrix} E & F \\ F & G \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} L & M \\ M & N \end{bmatrix} \tag{2.2.3}$$

$$E = \mathbf{r}_u\,\mathbf{r}_u\,, \qquad F = \mathbf{r}_u\,\mathbf{r}_v\,, \qquad G = \mathbf{r}_v\,\mathbf{r}_v \tag{2.2.4}$$

$$L = \mathbf{n}\,\mathbf{r}_{uu}\,, \qquad M = \mathbf{n}\cdot\mathbf{r}_{uv}\,, \qquad N = \mathbf{n}\cdot\mathbf{r}_{vv} \tag{2.2.5}$$

The first fundamental form gives the metric of an "element" of arc length of a surface curve:

$$|\Delta\mathbf{r}|^2 \approx d\mathbf{r}\cdot d\mathbf{r} = I \tag{2.2.6}$$

The second fundamental form denotes the deviation between two close surface points, that is, it gives twice the distance h from one point Q to the tangent plane of the other point P

$$h = \mathbf{n}\cdot\Delta\mathbf{r} = \mathbf{n}\, d\mathbf{r} + \tfrac{1}{2}\mathbf{n}\cdot d^2\mathbf{r} + \mathbf{n}\cdot o(d^2\mathbf{r}) \approx \tfrac{1}{2}\mathbf{n}\cdot d^2\mathbf{r} \tag{2.2.7}$$

Their geometrical meaning is illustrated in Figure 2.2.1.

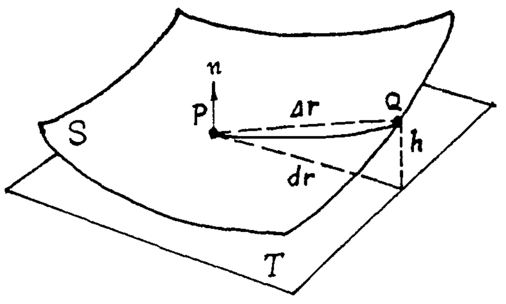


Figure 2.2.1. Geometrical meaning of the 1st and 2nd fundamental forms of surface.

Due to the invariance of the 1st order differential form, dr is invariant under surface rotation, translation or change of (u,v)-parameterisation. Hence, the first fundamental form I, as the dot product of dr, maintains the same property, which is referred to as an *intrinsic* property of the surface since measurements on the surface are made without "leaving" the surface itself; whereas the second fundamental form, which may change sign due to these transformations since the surface normal depends on the embedding of the surface into R^3, is referred to as an *extrinsic* property.

In a Cartesian coordinate system, the surface S is usually in the form $z = z(x,y)$ known as Monge's form. Replacing u with x, v with y gives rewritten expressions as below:

$$S = \left\{ \begin{bmatrix} x \\ y \\ z(x,y) \end{bmatrix} \Bigg|_{(x,y) \in D \subseteq R^2} \right\} \tag{2.2.8}$$

$$dz = z_x dx + z_y dy = [\; z_x \; z_y] \begin{bmatrix} dx \\ dy \end{bmatrix} \tag{2.2.9}$$

$$I = \begin{bmatrix} dx \\ dy \end{bmatrix}^T \begin{bmatrix} E & F \\ F & G \end{bmatrix} \begin{bmatrix} dx \\ dy \end{bmatrix} \tag{2.2.10}$$

$$II = \begin{bmatrix} dx \\ dy \end{bmatrix}^T \begin{bmatrix} L & M \\ M & N \end{bmatrix} \begin{bmatrix} dx \\ dy \end{bmatrix} \tag{2.2.11}$$

where

$$E = 1 + z_x^2 , \qquad F = z_x z_y , \qquad G = 1 + z_y^2 \tag{2.2.12}$$

$$L = \frac{z_{xx}}{\sqrt{1 + z_x^2 + z_y^2}} , \quad M = \frac{z_{xy}}{\sqrt{1 + z_x^2 + z_y^2}} , \quad N = \frac{z_{yy}}{\sqrt{1 + z_x^2 + z_y^2}} \tag{2.2.13}$$

Thus, setting dx, $dy = \pm 1$ pixel, the second fundamental form II gives the expression of deviations from each 8-connected neighbours in terms of L, M and N:

$(L+N-2M)_{i-1,j-1}$	$(N)_{i,j-1}$	$(L+N+2M)_{i+1,j-1}$
$(L)_{i-1,j}$	0	$(L)_{i+1,j}$
$(L+N+2M)_{i-1,j+1}$	$(N)_{i,j+1}$	$(L+N-2M)_{i+1,j+1}$

Figure 2.2.2 Estimation of deviations from the central point (i,j) to its 8-connected neighbours using the 2nd fundamental form

2.3. The Gauss formula and the Weingarten formula of surfaces

As mentioned before, $\mathbf{n}$, $\mathbf{r}_u$ and $\mathbf{r}_v$ can form a local coordinate frame at each surface point. Their derivatives can be expressed as a linear combination of $\mathbf{r}_u$, $\mathbf{r}_v$ and $\mathbf{n}$ themselves with coefficients dependent only on the first and second fundamental forms. These expressions are called the Gauss formula

$$[\; \mathbf{r}_{uu} \; \mathbf{r}_{uv} \; \mathbf{r}_{vv} \;] = [\; \mathbf{r}_u \; \mathbf{r}_v \; \mathbf{n} \;] \begin{bmatrix} \Gamma_{11}^1 & \Gamma_{12}^1 & \Gamma_{22}^1 \\ \Gamma_{11}^2 & \Gamma_{12}^2 & \Gamma_{22}^2 \\ L & M & N \end{bmatrix} \tag{2.3.1}$$

and the Weingarten formula

$$[\; \mathbf{n}_u \; \mathbf{n}_v] = - [\; \mathbf{r}_u \; \mathbf{r}_v] \; \mathbf{W} \tag{2.3.2}$$

where the six coefficients Γ^k_{ij}, $(i,j,k=1,2)$, named Christoffel symbols, depend only on the first fundamental form as below, where symbols E_u and E_v are the first derivatives of E with respect to u and v. Thus, all geometrical concepts and properties which can be expressed in terms of the Christoffel symbols are isometrical, i.e., they are invariant under any transformation preserving the first fundamental form.

$$\Gamma^1_{11}=\frac{GE_u-2FF_u+FE_v}{2(EG-F^2)}\qquad \Gamma^1_{12}=\frac{GE_v-FG_u}{2(EG-F^2)}\qquad \Gamma^1_{22}=\frac{2GF_v-GG_u-FG_v}{2(EG-F^2)}$$

$$\Gamma^2_{11}=\frac{2EF_u-EE_v-FE_u}{2(EG-F^2)}\qquad \Gamma^2_{12}=\frac{EG_u-FE_v}{2(EG-F^2)}\qquad \Gamma^2_{22}=\frac{EG_v-2FF_v+FG_u}{2(EG-F^2)} \tag{2.3.3}$$

Matrix W is termed the Weingarten mapping matrix, which is of great importance in determining the surface shape.

$$\mathrm{W}=\frac{1}{EG-F^2}\begin{bmatrix} GL-FM & GM-FN \\ EM-FL & EN-FM \end{bmatrix}=\mathrm{A}^{-1}\mathrm{B} \tag{2.3.4}$$

2.4. The surface curvatures

At a point (u,v) on the surface S, the intersection of all normal planes $\mathrm{N}(u,v)$ and the surface S yields infinitely many surface curves along different directions. The curvatures of these curves are called the normal curvatures C_n shown in Figure 2.4.1.

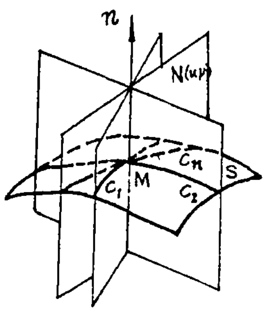


Figure 2.4.1 The principal curvatures and directions of surface.

Among the normal curvatures C_n, the least and the greatest curvatures are referred to as the principal curvatures C_1 and C_2, and their directions are referred to as the principal directions, given in the tangential forms d_1 and d_2 respectively. The surface curvature K and mean curvature H are then defined in terms of principal curvatures C_1 and C_2, thus:

$$K=C_1\cdot C_2$$

$$H=\tfrac{1}{2}(C_1+C_2) \tag{2.4.1}$$

However, preferred are a definition of the Gaussian curvature K and mean curvature H in terms of the matrix W:

$$K=det(\mathrm{W})=\frac{LN-M^2}{EG-F^2}$$

$$H = \tfrac{1}{2} tr(\mathbf{W}) = \tfrac{1}{2} \frac{EN - 2FM + GL}{EG - F^2} \qquad (2.4.2)$$

and a definition of the principal curvature and the direction pairs $(C_p\,, \phi_p)$ $(p{=}1{,}2)$ from the same matrix **W**:

$$C_p \phi_p = \mathbf{W} \phi_p \qquad p = 1\,,2 \qquad (2.4.3)$$

Usually, the principal direction ϕ_p is denoted in the tangent form as below:

$$d_p = \left[\frac{dv}{du} \right]_p \qquad p = 1, 2 \qquad (2.4.4)$$

The preference for definitions from the matrix **W** is based on the reason that from the theory of matrix algebra, it is easy to reconfirm that K and H are characteristics of the surface shape in terms of **W** and that K, H, C_1, C_2 are algebraic invariants to surface rotation, translation or (u,v)-parameterisation because the eigenvalues of the matrix will not be changed by such linear transformations. Hence, the previous definition (2.4.1) of K and H is now a self-evident relationship between K, H and C_1, C_2.

Also from the definition of principal directions based on matrix **W**, it is easy to show that ϕ_1 is orthogonal to ϕ_2 and derive the following quadratic equation, of which d_1 and d_2 are two roots:

$$(GM{-}FN) \left[\frac{dv}{du} \right]^2 + (GL{-}EN) \frac{dv}{du} + (FL{-}EM) = 0 \qquad (2.4.5)$$

All the surface curves satisfying the above ordinary differential equation are called lines of curvature, which form an orthogonal curve network upon the surface. All the tangent directions along the lines of curvature are principal directions.

2.5. The Gaussian curvature as an intrinsic property of a surface

There exist relationships between the first and second fundamental forms, which are referred to as the compatibility, or the Gauss equations

$$\begin{aligned}
KF &= (\Gamma^1_{12})_u - (\Gamma^1_{11})_v + \Gamma^2_{12}\Gamma^1_{12} - \Gamma^2_{11}\Gamma^1_{22} \\
&= (\Gamma^2_{12})_v - (\Gamma^2_{22})_u + \Gamma^1_{12}\Gamma^2_{12} - \Gamma^1_{22}\Gamma^2_{11} \\
KE &= (\Gamma^2_{11})_v - (\Gamma^2_{12})_u + \Gamma^1_{11}\Gamma^2_{12} + \Gamma^2_{11}\Gamma^2_{22} - \Gamma^1_{22}\Gamma^2_{11} - (\Gamma^2_{12})^2 \\
KG &= (\Gamma^1_{22})_u - (\Gamma^1_{12})_v + \Gamma^2_{22}\Gamma^1_{12} + \Gamma^1_{22}\Gamma^1_{11} - \Gamma^2_{12}\Gamma^1_{22} - (\Gamma^1_{12})^2
\end{aligned} \qquad (2.5.1)$$

and the Mainardi-Codazzi equations:

$$\begin{aligned}
\frac{\partial L}{\partial v} - \frac{\partial M}{\partial u} &= L\Gamma^1_{12} + M(\Gamma^2_{12}{-}\Gamma^1_{11}) - N\Gamma^2_{11} \\
\frac{\partial M}{\partial v} - \frac{\partial N}{\partial u} &= L\Gamma^1_{22} + M(\Gamma^2_{22}{-}\Gamma^1_{12}) - N\Gamma^2_{12}
\end{aligned} \qquad (2.5.2)$$

A striking fact is that although the Gaussian curvature K is defined by the six terms E, F, G, L, M and N of both the 1st and 2nd fundamental forms in (2.4.2), it can be seen from the above Gauss equations that K actually depends on only three functions E, F, G and their derivatives, and thus depends on only the first fundamental form. One of the most important facts of differential geometry is therefore:

1) **Gauss theorem egregium:** The Gaussian curvature K of a surface is invariant under local isometries.
2) It only depends on the metric structure of the surface not its position in R^3, hence, under isometries, a surface point and its corresponding point on the deformed surface have the same Gaussian curvature K.

Hence, the Gaussian curvature is a fundamental property of a surface. Any two regular surfaces are locally conformal. However, surfaces with different Gaussian curvatures cannot reach globally conformal through arbitrary isometric deformation.

On other hand, the shape of a surface cannot be determined using only the first fundamental form To get a unique shape, an extrinsic constraint is required to prevent the surface from being deformable. The following theorem makes explicit this uniqueness of the first and second fundamental form for any smooth surface:

Bonnet theorem: Given any two quadric forms:

$$I = \begin{bmatrix} du \\ dv \end{bmatrix}^T \begin{bmatrix} E & F \\ F & G \end{bmatrix} \begin{bmatrix} du \\ dv \end{bmatrix} > 0 , \qquad II = \begin{bmatrix} du \\ dv \end{bmatrix}^T \begin{bmatrix} L & M \\ M & N \end{bmatrix} \begin{bmatrix} du \\ dv \end{bmatrix} \tag{2.5.3}$$

where the coefficients of I and II satisfy the Gauss equations and the Mainardi-Codazzi equations, then, up to a translation and rotation, there must exist a unique surface which takes I and II as its first and second fundamental forms.

So, surface shape information is concentrated in the six scalar functions E, F, G, L, M and N, or in the matrices $\mathbf{A}$ and $\mathbf{B}$, or the Weingarten mapping matrix $\mathbf{W} = \mathbf{A}^{-1}\mathbf{B}$.

2.6. Surface classification based on curvatures

Using the sign of Gaussian curvature K, surface patches can be categorised into three types:

1) Ellipsoidal, if $K > 0$; and
2) Parabolic, if $K = 0$; and
3) Hyperbolic, if $K < 0$.

As the first fundamental form is always positive, $EG - F^2 > 0$, from (2.4.2) $sgn(K) \equiv sgn(LN{-}M^2)$. The K sign category seems an extrinsic classification since it only refers to the second fundamental form. But this is *not* true. Due to the compatibility equations (2.5.1), the second fundamental form is not an arbitrary quantity once the first fundamental form is given, so the K sign category is virtually an intrinsic classification.

Using the sign of principal curvatures C_1 and C_2, surfaces can be categorised into six types:

C_1 \ C_2	–	o	+
–	peak	ridge	saddle
o	ridge	flat	valley
+	saddle	valley	pit

Figure 2 6 1 Surface shapes from only the signs of C_1, C_2.

This (C_1, C_2) sign category provides more types than the K sign category, but the curvature magnitude information of C_1 and C_2 is completely ignored. To involve the magnitudes of principal curvatures in the classification and for clarity too, C_1 is chosen as the principal curvature which has the smaller magnitude, associated with its tangential direction d_1, and C_2 is chosen as the principal curvature of the larger magnitude, associated with the tangential direction d_2. So, $0 \leq |C_1| \leq |C_2|$.

In this way, a part of the curvature magnitude information is merged into the (C_1,C_2) *sign* category to form a new category, the (C_1,C_2) category, which increases the surface types from six to seven.

C_1 \ C_2	–	o	+
–	peak	(none)	saddle valley + minimal
o	ridge	flat	valley
+	saddle ridge + minimal	(none)	pit

Figure 2 6.2. Surface shapes from both the signs and magnitudes of C_1, C_2.

Alternatively, note that the sign of Gaussian curvature K contains only information from the first fundamental form, but the sign of the mean curvature H contains information from the second fundamental form, hence the (K,H) sign pair may also be used to categorise surfaces.

K \ H	–	o	+
–	saddle ridge	minimal	saddle valley
o	ridge	flat	valley
+	peak	(none)	pit

Figure 2 6 3 Surface shapes from the signs of K, H.

This time the surface is divided into eight types, which include the (C_1,C_2) categories with one additional type. This type is the minimal patch which is separated from both saddle ridge and saddle valley types. Although the minimal surface is rarely encountered and difficult to separate from the saddle valley or saddle ridge surface as will be shown in Chapter 4, it is an indication that the (K,H) sign category combines more information than the (C_1,C_2) category, thus is probably a better surface category for theoretical classification.

There are also two uniqueness theorems from differential geometry which support use of the (K,H) sign category whereas there is no counterpart for the (C_1,C_2) category so far.

Gaussian curvature uniqueness theorem:

For compact convex surfaces (where $LN>M^2$ at every point), there is a single scalar function $K(u,v)$ (Gaussian curvature) that uniquely specifies the surface shape [Chern 1957],[Hsiung 1981].

Mean curvature uniqueness theorem:

The solution of the following partial differential equation, $f(u,v)$ must be unique if it exists on the domain D and at the boundary ∂D [Gilbarg and Trudinger 1983]:

$$\nabla\left[\frac{\nabla f}{\sqrt{1+|\nabla f|^2}}\right]=2H \tag{2.6.1.a}$$

$$f\,\Big|_{\partial D}=f_o \tag{2.6.1.b}$$

Applying the (K,H) sign category to the "Monkey saddle" surface [Do Carmo 1976]

$$x=u\,,\qquad y=v\,,\qquad z=u^3-3uv^2 \tag{2.6.2}$$

gives six patches around an umbilic point (0,0) in Figure 2.6.4, each patch being a homogeneous region in the (K,H) sign category. (This result will be referred to in Section 5.4.)

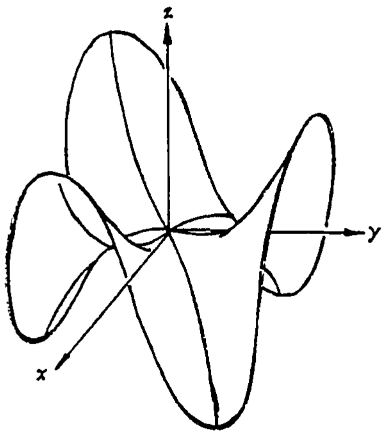


Figure 2.6.4. Monkey saddle surface and its patches corresponding to the (K,H) sign category.

2.7. Other definitions of the Gaussian curvature K

There are some alternative definitions of Gaussian curvature K. In the Gaussian mapping for convex surfaces, K is defined as:

$$K=\lim_{\Delta A\to 0}\frac{\Delta S}{\Delta A} \tag{2.7.1}$$

where the surface normal **n** of an object is mapped, preserving the direction, onto a radius vector **r** of a unit sphere, and an area ΔA on the object surface is mapped to an area ΔS on the unit sphere.

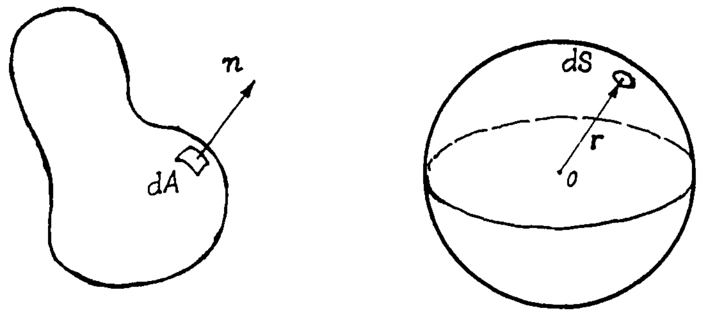


Figure 2.7.1 Surface and its Gaussian mapping.

K can also be defined by the surface normal **n** and its derivatives along u and v in (2.7.2). This definition is closely related to (2.7.1) since the geometrical meaning of the magnitude of a cross product is the area:

$$K = \frac{|\mathbf{n}_u \times \mathbf{n}_v|}{|\mathbf{r}_u \times \mathbf{r}_v|} \tag{2.7.2}$$

2.8. Other intrinsic and extrinsic properties of surface

There are many surface properties, such as the arc length of a surface curve, the area of a surface patch and the angle between two intersecting surface curves, *etc.*, whose computation depends on only the first fundamental form. Hence they are intrinsic too.

The formula for computing the arc length L of a surface curve is:

$$L = \int_{t_0}^{t_1} dl = \int_{t_0}^{t_1} \sqrt{Eu^2 + 2Fu\dot{v} + Gv^2}\,dt \tag{2.8.1}$$

where

$$u = u(t)\ , \quad v = v(t)\ , \quad u = \frac{\partial u}{\partial t}\ , \quad v = \frac{\partial v}{\partial t} \qquad t \in [t_0\ ,\ t_1] \tag{2.8.2}$$

The formula for computing area A of surface patch S is:

$$A = \iint_S dA = \iint_S \sqrt{EG - F^2}\,dudv \tag{2.8.3}$$

And to compute the angle α between two surface curves at the intersection point $M(u,v)$,

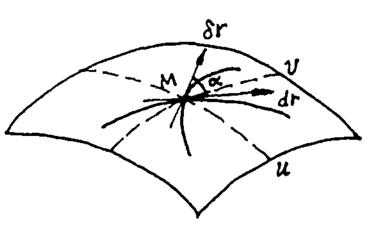


Figure 2.8.1. Joint angle between curves or coordinate lines.

the formula is:

$$\cos\alpha = \frac{d\mathbf{r}\ \delta\mathbf{r}}{|d\mathbf{r}|\ |\delta\mathbf{r}|} = \frac{Edu\delta u + F(du\delta v + dv\delta u) + Gdv\delta v}{\sqrt{Edu^2 + 2Fdudv + Gdv^2}\ \sqrt{E\delta u^2 + 2F\delta u\delta v + G\delta v^2}} \tag{2.8.4}$$

So the intersection angle θ between coordinate curves $u = const$ and $v = const$ is:

$$\cos\theta = \frac{F}{\sqrt{EG}} \tag{2.8.5}$$

Hence, the coordinate curves are orthogonal if and only if

$$F = 0 \tag{2.8.6}$$

Note that the second fundamental form expresses the deviation between two neighbouring points on the surface and that the matrix eigenvalues are invariant to rotation. Also note that the eigenvalue λ_p and its associated unit eigenvector ψ_p of matrix $\mathbf{B}$ satisfy:

$$\psi_p^T\psi_p = 1 \tag{2.8.7}$$

$$\mathbf{B}\psi_p = \lambda_p\psi_p \qquad p = 1,2 \tag{2.8.8}$$

The second fundamental form may be expressed as:

$$II = \psi_p^T\mathbf{B}\psi_p = \lambda_p \qquad p = 1,2 \tag{2.8.9}$$

So the sum of the squared eigenvalues of matrix $\mathbf{B}$ can be used to estimate the local flatness in the following form:

$$Q = \lambda_1^2 + \lambda_2^2 = tr(\mathbf{B}^T\mathbf{B}) \tag{2.8.10}$$

When this quantity Q is small, both λ_1 and λ_2 are small, so are the second fundamental form II and the deviation h in (2.2.7), thus the surface must locally flat (see Figure 2.2.1).

The review of other intrinsic surface properties such as geodesic curvature, geodesic lines and the Gauss-Bonnet theorem (perhaps the most important result in differential geometry), are omitted as they will not be used in this thesis.

Chapter 3

Diffusion Smoothing

3.1. Introduction

Suppression of noise in raw data is a fundamental problem in early visual processing. Many papers have been written on this issue and many methods have been proposed. Even today, with the laser range finder significantly raising the precision of measurements, this problem remains of high importance. Thus any effort to obtain even a small improvement in noise suppression would be justified. The investigation of diffusion smoothing in this chapter will make explicit its relationships to other smoothing methods. It also provides a faster algorithm and a proper boundary treatment so as to set up a framework for scale space processing.

3.1.1. Gaussian filter

Among a broad class of globally applied filters, the Gaussian filter is well-known as an elegant and efficient one. It uses the Gaussian distribution as the kernel of convolution. Not only can it produce images of different resolutions (scales of measurement) over which significant features of the object will appear and disappear, it is also the unique filter possessing the property that a "zero-crossing will not be created as the scale increases" [Yuille and Poggio 1983].

When the scale increases, noise is suppressed to become smaller and smaller. Unfortunately, the signal distortion also becomes larger and larger. But important information can still be retained by tracking the movement of features from fine scales to coarse scales owing to the uniqueness property of the Gaussian filter.

3.1.2. Diffusion Model

As well as using the Gaussian filter in image processing, a diffusion filter can be used for the same task. The physics of diffusion makes the understanding of this filter intuitive.

Rosenfeld and Kak [1976] argued that "the blur in the picture is a result of a diffusion process that satisfies the well-known partial differential equation $\partial g/\partial t = k\nabla^2 g$ where g is a function of x, y, and t (time), and k>0 is a constant", and "a point source blurs into a spot with a Gaussian distribution of brightness". Actually, this implies that Gaussian filtering can be seen as a diffusion process although Rosenfeld applied this diffusion model in the inverse time direction — the image sharpening or the inverse diffusion process.

Observing that "diffusion only destroys structure but cannot generate it", Koenderink [1984] used the diffusion process to show the uniqueness of Gaussian smoothing for multi-scale filtering and argued that "Gaussian blurring is the only sensible way to embed a primal image into a one-parameter family" in both 1-D and 2-D cases. He also mentioned the numerical stability problem in the diffusion process.

Gourlay [1985] proposed to use the implicit numerical schemes of the diffusion equation with a fixed boundary (Dirichlet condition) to do Gaussian smoothing. The numerical stability problem was mentioned too, but not so much attention was paid to this topic, perhaps due to a lack of motivation of practical application in multiscale space. As a result, in his paper there was neither a feasible 2-D implicit scheme nor complexity analysis nor boundary treatment for shape preservation in multiscale filtering.

3.1.3. Problem Posed

To make diffusion smoothing practicable, especially in the scale space, some simple, stable and efficient numerical schemes are needed to solve the diffusion problem with less computational complexity than the Gaussian convolution or the repeated averaging algorithm, and treat boundaries properly so as to preserve some special properties of surfaces.

For these reasons, that will be further addressed in Section 3.3, an implicit scheme is preferred so as to get rid of the restriction from the stability limit of a numerical scheme; and analogy with physical phenomenon is also preferred so as to deal with the boundary problem intuitively and reasonably.

Mathematically speaking, the diffusion equation will eventually be solved using a 2-D implicit scheme together with a "natural boundary condition" treatment. Meaning and details of such a solution will be made clear along with the development of the diffusion smoothing theory in subsequent sections.

3.2. Gaussian Smoothing (GS)

There are two traditional ways to implement Gaussian smoothing. One way is the direct Gaussian convolution and the other is repeated averaging, an iterative approximation form.

3.2.1. Gaussian smoothing using Gaussian convolution (GC)

A given surface f(x,y) can be smoothed by Gaussian convolution using Gaussian distribution as the kernel:

$$F(x,y,\sigma) \equiv f(x,y) * g(x,y,\sigma) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\xi,\eta) exp[-\frac{(\xi-x)^2+(\eta-y)^2}{2\sigma^2}]\, d\xi d\eta \qquad (3.2.1)$$

Using the separability and the associativity of Gaussian convolution, there will be

$$f(x,y)*g(x,y,\sigma) = f(x,y)*g(x,o,\sigma)*g(o,y,\sigma) \qquad (3.2.2)$$

As the Gaussian function is a negative exponential function which decays to zero very rapidly, typically, an 8σ width square window is enough for the discrete Gaussian convolution as used by Canny [1983]. His approximation leads to the number of multiplication operations at each pixel (point) in both x and y directions being:

$$8\sigma + 8\sigma = 16\sigma \qquad (3.2.3)$$

And for an M×M pixel image, the number of multiplications will be $16\sigma M^2$.

As Gaussian convolution is not a scale-incremental operation, the volume of multiplication over all $\sigma = 1, 2, 3, .., K$ will be

$$C_{GS} = \sum_{\sigma=1}^{K} 16\sigma M^2 = 8K(K+1)M^2 = O(K^2M^2) \qquad (3.2.4)$$

3.2.2. Gaussian smoothing using repeated averaging (RA)

Brady *et al* [1985] proposed a method called repeated averaging which "appeals to the central-limit theorem and implements Gaussian filtering using repeated averaging with the 3x3 mask" below:

$\frac{1}{24} \times$

1	2	1
2	12	2
1	2	1

Figure 3.2.1 Brady's repeated averaging (RA) mask.

They showed that "iterating n times approximately corresponds to filtering with a Gaussian whose standard deviation is proportional to $\sqrt{n}$". For convenience, this mask will be referred as the *RA* mask afterwards.

There are some *RA-like* masks, such as masks which appeared in [Hummel and Moniot 1987] and [Cai 1988c] as shown in Figure 3.2.2. They are different from the RA mask in weighting coefficients, but all the three masks can be uniformly derived from the same diffusion smoothing scheme shown in the next section.

$\frac{1}{16} \times$

1	2	1
2	4	2
1	2	1

$\frac{1}{36} \times$

1	4	1
4	16	4
1	4	1

Figure 3.2.2. Hummel's square mask (left) and Cai's spline smoothing mask (right).

Note that some papers used larger symmetric square masks, such as 5x5 and 7x7, even 9x9 *etc*, and some of them can be decomposed into smaller square masks. A square mask propagation theorem will be proven here as a complement to the repeated averaging method.

According to this theorem, a 3x3 square mask can be used to construct, by iteration, larger odd-order square smoothing masks of the same degrees of freedom. And the decomposition of a large square mask can be tested by comparing it with the same large mask which is derived from a smaller mask Since using the smallest width masks in smoothing can ease the surface boundary treatment and the analysis of scale space behaviour of small features, it is preferred to investigate the behaviour of a 3x3 mask instead of those larger masks in repeated averaging.

Theorem 3.1 *(Square mask propagation)· Two iterations of a 3x3 symmetric square mask produce a 5x5 symmetric square mask with the same degree of freedom.*

proof: A normalised 3x3 symmetric square mask must be of this form:

a	*b*	*a*
b	*c*	*b*
a	*b*	*a*

Figure 3.2.3. A 3x3 symmetric square mask is the smallest width mask unit in smoothing.

which depends on three parameters a, b and c and is constrained to the normalisation condition that the sum of all mask elements should be equal to one. Thus it has at most two degrees of freedom. Using this mask twice is equivalent to using the following 5x5 square mask once:

a^2	$2ab$	$2a^2+b^2$		
	$2ac+2b^2$	$4ab+2bc$		
		$4a^2+4b^2+c^2$		

Figure 3.2.4. A 5x5 symmetrical square mask obtained from two iterations of a 3x3 mask.

where the blank boxes are filled according to the symmetry.

Obviously, the degree of freedom of this 5x5 mask is not higher than two. In fact, all the mask elements are combinations of the three parameters a, b and c. They can be *uniquely* determined by three items a^2, $2a^2+b^2$ and $4a^2+4b^2+c^2$ together with three other items $2ab$, $2ac+2b^2$, $4ab+2bc$ and the normalisation constraint.

Q.E.D.

Remark 3.1: *Whether or not a given 5x5 square mask is derived from a 3x3 mask can be identified by comparing the given mask with the 5x5 mask in Figure 3.2.4 to form a system of equations. If the system has a solution, these unknown parameters a, b, c (and the unknown 3x3 mask) can be determined, and a decomposition of the given 5x5 mask is achieved.*

Remark 3.2: *The same idea applies to cases of generating larger masks, such as, producing a 7x7 symmetric square mask from a 3x3 mask by three iterations; and a 9x9 symmetric mask from the 3x3 mask by four iterations, or from a 5x5 mask by two iterations.*

3.3. Diffusion Smoothing (DS)

3.3.1. Mathematical physics background

A. Diffusion phenomenon

In the physical world, diffusion and heat conduction in a continuous medium are similar phenomena. For example, in the 1-D case, given an initial triangular distribution, the top of the wave form will decrease and the shape will tend to be more and more smooth no matter how sharp and uneven the initial distribution is. Although the diffusion effect spreads over an infinite range in theory, it decays to zero fairly fast as a negative exponential function. So it affects only a limited range in practice. This physical phenomenon is subject to a second order parabolic partial differential equation — the diffusion equation with an initial value condition:

$$\frac{\partial u}{\partial t} = b\nabla^2 u \tag{3.3.1e}$$

$$u\big|_{t=0} = f \tag{3.3.1i}$$

B. Relationship between the diffusion process and Gaussian convolution

As is well-known, the analytic solution of the diffusion problem (3.3.1) is

$$u(x,y) = \frac{1}{4\pi bt} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\xi,\eta) exp[-\frac{(\xi-x)^2+(\eta-y)^2}{4bt}] \, d\xi d\eta \qquad (3.3.2)$$

Substituting σ in (3.2.1) with

$$\sigma = \sqrt{2bt} \qquad (3.3.3)$$

it is easy to see that the Gaussian convolution $f(x,y)*g(x,y,\sigma)$ is just the above analytic solution $u(x,y)$. So, the implementation of the Gaussian convolution can be replaced with the solution of a diffusion equation. The time variable t of the diffusion process plays a similar role to the scale parameter σ of the Gaussian convolution.

3.3.2. Motivation of research

Based on the above observations, the motivation of this research on diffusion smoothing is to provide an efficient numerical scheme with proper boundary treatment as a better alternative for performing noise suppression in early visual processing. Possible advantages are as follows:

A. Intuition

The equivalence between the Gaussian convolution and the diffusion process is a convenient bridge in understanding how the same problem can be solved using a different approach. By analogy with a heat distribution, surface smoothing can be implemented by solving the diffusion problem. A smoothed surface can be prevented from "melting" into the background by imposing a fixed boundary. And a new diffusion model for discontinuous media can be set up to preserve some important properties of the surface. Even a "heat convection" term can be easily introduced into the diffusion equation for smoothing a moving object's data. These might be tedious to do with Gaussian smoothing.

B. Data to data, free-form approach

Note that surface fitting methods, such as spline fitting and least-squares fitting can give a satisfactory result in many cases with the cost of imposing some constraint conditions, *e.g.*, continuity of surface derivatives, or empirically choosing a set of specific basis functions, *e.g.*, a plane or a sphere or a cylinder *etc.* before the fitting. However, these might be rather difficult to do when fitting in an irregular domain.

In contrast, discrete surface smoothing is a free-form, data to data approach. It does not need any extra step, such as choosing the fitting basis functions, computing their coefficients and so on. So it is more convenient for a free-form surface. Meanwhile, diffusion smoothing is governed by a partial differential equation rather than the central-limit theorem. This ensures that the smoothing works on a *deterministic* basis with more "confidence" than on a statistical basis

C. More efficient methods

Since Gaussian smoothing is applied to the whole surface, the computational expense is fairly large. Note that the topic of partial differential equations has been deeply investigated and many sophisticated means for solving them have been explored. It is worth trying to implement Gaussian convolution by solving a differential equation with these efficient techniques.

D. Uniform processing

Numerical analysis methods also provides the possibility of dealing with some common problems in smoothing from a uniform viewpoint. These problems may be the choice of the smoothing mask size, the optimal mask in a smallest neighbourhood, the comparison among behaviour of different square masks and so on.

E. Uniform interpretation

The theoretical tool developed in the above investigation would be helpful in handling other smoothing methods even though their appearances might be quite different. The diffusion smoothing method could even provide a uniform interpretation of these different methods. As it is more general, it can do more.

F. Framework of scale space processing

As the diffusion time t is related to the spatial scale σ of the Gaussian, it provides a continuous parameter for incremental smoothing. So, subsequent processing can be naturally implemented within this framework, leading to a scale space representation for surface shape.

3.3.3. Diffusion explicit smoothing scheme (DES)

Roughly speaking, there are two kinds of numerical schemes for solving time-dependent partial differential equations. One is the *explicit* scheme, which advances a solution to the next time from those known values at the current time (and the past time, if necessary). The other is the *implicit* scheme, which advances a solution to the next time using both known values at the current time (and the past time, if necessary) and unknown values at the next time. Hence, it needs to solve a linear algebra system which implicitly specifies the solution at the next time.

A. Explicit numerical schemes of diffusion equation

Since a surface is given in discrete form, it is possible that no derivatives exist at some sampling points where discontinuities occur. But the numerical derivative always exists at any point, and it will approximate the exact derivative, if any, as the sampling step tends to zero.

For any point (p,q), its 8-connected neighbours on a square mesh are:

o o o
o * o
o o o

Figure 3.3.1. The central point (p,q) and its 8-connected neighbours on a square mesh.

Let the mesh step unit be h and the time step unit τ. Applying Taylor series at the centre point (p,q) to expand the value from time k+1 to time k and expand the values of all the neighbours around the point (p,q) and time k gives the following expressions:

$$u^{k+1}_{p,q} = [I + \tau\frac{\partial}{\partial t}]u^k_{p,q} + O(\tau^2) \qquad (3.3.4)$$

$$u^k_{p-1,q} = [I - h\frac{\partial}{\partial x} + \frac{h^2}{2!}\frac{\partial^2}{\partial x^2} - \frac{h^3}{3!}\frac{\partial^3}{\partial x^3}]u^k_{p,q} + O(h^4) \qquad (3.3.5)$$

$$u^k_{p+1,q} = [I + h\frac{\partial}{\partial x} + \frac{h^2}{2!}\frac{\partial^2}{\partial x^2} + \frac{h^3}{3!}\frac{\partial^3}{\partial x^3}]u^k_{p,q} + O(h^4) \qquad (3.3.6)$$

$$u^k_{p,q-1} = [I - h\frac{\partial}{\partial y} + \frac{h^2}{2!}\frac{\partial^2}{\partial y^2} - \frac{h^3}{3!}\frac{\partial^3}{\partial y^3}]u^k_{p,q} + O(h^4) \tag{3.3.7}$$

$$u^k_{p,q+1} = [I + h\frac{\partial}{\partial y} + \frac{h^2}{2!}\frac{\partial^2}{\partial y^2} + \frac{h^3}{3!}\frac{\partial^3}{\partial y^3}]u^k_{p,q} + O(h^4) \tag{3.3.8}$$

$$u^k_{p-1,q-1} = [I - h(\frac{\partial}{\partial x}+\frac{\partial}{\partial y}) + \frac{h^2}{2!}(\frac{\partial}{\partial x}+\frac{\partial}{\partial y})^2 - \frac{h^3}{3!}(\frac{\partial}{\partial x}+\frac{\partial}{\partial y})^3]u^k_{p,q} + O(h^4) \tag{3.3.9}$$

$$u^k_{p+1,q+1} = [I + h(\frac{\partial}{\partial x}+\frac{\partial}{\partial y}) + \frac{h^2}{2!}(\frac{\partial}{\partial x}+\frac{\partial}{\partial y})^2 + \frac{h^3}{3!}(\frac{\partial}{\partial x}+\frac{\partial}{\partial y})^3]u^k_{p,q} + O(h^4) \tag{3.3.10}$$

$$u^k_{p-1,q+1} = [I - h(\frac{\partial}{\partial x}-\frac{\partial}{\partial y}) + \frac{h^2}{2!}(\frac{\partial}{\partial x}-\frac{\partial}{\partial y})^2 - \frac{h^3}{3!}(\frac{\partial}{\partial x}-\frac{\partial}{\partial y})^3]u^k_{p,q} + O(h^4) \tag{3.3.11}$$

$$u^k_{p+1,q-1} = [I + h(\frac{\partial}{\partial x}-\frac{\partial}{\partial y}) + \frac{h^2}{2!}(\frac{\partial}{\partial x}-\frac{\partial}{\partial y})^2 + \frac{h^3}{3!}(\frac{\partial}{\partial x}-\frac{\partial}{\partial y})^3]u^k_{p,q} + O(h^4) \tag{3.3.12}$$

$$\begin{array}{ccc} & \mathrm{o} & \\ \mathrm{o} & + & \mathrm{o} \\ & \mathrm{o} & \end{array} \equiv u^k_{p-1,q} + u^k_{p+1,q} + u^k_{p,q-1} + u^k_{p,q+1} = (4I + h^2\nabla^2)u^k_{p,q} + O(h^4) \tag{3.3.13}$$

$$\begin{array}{ccc} \mathrm{o} & & \mathrm{o} \\ & \times & \\ \mathrm{o} & & \mathrm{o} \end{array} \equiv u^k_{p-1,q-1} + u^k_{p+1,q+1} + u^k_{p-1,q+1} + u^k_{p+1,q-1} = (4I + 2h^2\nabla^2)u^k_{p,q} + O(h^4) \tag{3.3.14}$$

Some numerical schemes can be set up using these expressions. The simplest scheme is usually an explicit scheme which advances a solution to the time k+1 from the only known values at time k.

Suppose the given diffusion equation is

$$\frac{\partial u}{\partial t} = b\nabla^2 u \tag{3.3.15}$$

Using forward differences to discretise the equation at time k and point (p,q) surrounded by four neighbours in the horizontal and vertical directions gives

$$\frac{1}{\tau}(u^{k+1}_{p,q} - u^k_{p,q}) = \frac{b}{h^2}(u^k_{p+1,q} - 2u^k_{p,q} + u^k_{p-1,q}) + \frac{b}{h^2}(u^k_{p,q+1} - 2u^k_{p,q} + u^k_{p,q-1}) + O(\tau+h^2) \tag{3.3.16}$$

the difference equation of which is

$$u^{k+1}_{p,q} = (1 - 4\frac{b\tau}{h^2})u^k_{p,q} + \frac{b\tau}{h^2}(u^k_{p-1,q} + u^k_{p+1,q} + u^k_{p,q-1} + u^k_{p,q+1}) \tag{3.3.17}$$

Letting

$$\beta \equiv \frac{b\tau}{h^2} \tag{3.3.18}$$

and using the symbols $\begin{array}{ccc} & \mathrm{o} & \\ \mathrm{o} & + & \mathrm{o} \\ & \mathrm{o} & \end{array}$ in (3.3.13), equation (3.3.17) can be rewritten as

$$u^{k+1}_{p,q} = (1-4\beta)\, u^k_{p,q} + \beta \begin{array}{ccc} & \mathrm{o} & \\ \mathrm{o} & + & \mathrm{o} \\ & \mathrm{o} & \end{array} \tag{3.3.19}$$

It is called the *normal cross DES* scheme since it can be denoted in the mask form in Figure 3.3.2:

	β	
β	$1-4\beta$	β
	β	

Figure 3.3.2. The normal cross DES mask.

Similar, discretising the diffusion equation with the other four neighbours in the oblique directions gives:

$$\frac{1}{\tau}(u_{p,q}^{k+1} - u_{p,q}^{k}) = \frac{b}{2h^2}(u_{p+1,q+1}^{k+1} - 2u_{p,q}^{k+1} + u_{p-1,q-1}^{k+1}) + \frac{b}{2h^2}(u_{p-1,q+1}^{k+1} - 2u_{p,q}^{k+1} + u_{p+1,q-1}^{k+1}) + O(\tau+h^2) \tag{3.3.20}$$

And the difference equation, accordingly called the *oblique cross DES* scheme, is

$$u_{p,q}^{k+1} = (1-2\beta)\, u_{p,q}^{k} + \frac{\beta}{2}\begin{array}{ccc} \mathrm{o} & & \mathrm{o} \\ & \times & \\ \mathrm{o} & & \mathrm{o} \end{array} \tag{3.3.21}$$

which can be represented in the mask form

$\frac{\beta}{2}$		$\frac{\beta}{2}$
	$1-2\beta$	
$\frac{\beta}{2}$		$\frac{\beta}{2}$

Figure 3.3.3. The oblique cross DES mask.

Taking the average combination of the above two DES schemes gives the *isotropic square DES* scheme including eight neighbouring points:

$$u_{p,q}^{k+1} = (1-3\beta)u_{p,q}^{k} + \frac{\beta}{2}\begin{array}{ccc} & \mathrm{o} & \\ \mathrm{o} & + & \mathrm{o} \\ & \mathrm{o} & \end{array} + \frac{\beta}{4}\begin{array}{ccc} \mathrm{o} & & \mathrm{o} \\ & \times & \\ \mathrm{o} & & \mathrm{o} \end{array} \tag{3.3.22}$$

and the relevant mask

$\frac{\beta}{4}$	$\frac{\beta}{2}$	$\frac{\beta}{4}$
$\frac{\beta}{2}$	$1-3\beta$	$\frac{\beta}{2}$
$\frac{\beta}{4}$	$\frac{\beta}{2}$	$\frac{\beta}{4}$

Figure 3 3 4 The isotropic square DES mask

Obviously, all the three schemes are compatible with the diffusion equation, *i.e.*, they will converge to the diffusion equation as τ and h in (3.3.16) and (3.3.20) tend to zero. They have the same truncation error order $O(\tau+h^2)$, but their numerical stability can be different.

B. Stability analysis of the explicit scheme DES

The diffusion equation has the property of stability [Courant and Hilbert 1953], so that any small initial perturbation will be reduced in the diffusion process, which leads to a unique result if the diffusion time is sufficiently long. Unfortunately, the stability of a partial differential equation itself does not automatically assure the stability of its discrete form, the difference equation, and an unstable numerical scheme usually yields a solution with so large an error that it is pointless to use. So, numerical stability analysis is an indispensable step once the diffusion equation has been discretised into a specific scheme.

Suppose $u^k_{p,q}$ is the exact solution of difference equation (3.3.22) at the point (p,q). Because of a computing error $\varepsilon^k_{p,q}$, including the round-off error or initial perturbation error, only the approximation $\tilde{u}^k_{p,q}$ can be obtained from (3.3 22), as denoted below:

$$\tilde{u}^k_{p,q} = u^k_{p,q} + \varepsilon^k_{p,q} \tag{3.3.23}$$

By substituting $\tilde{u}$ in (3.3.23) for u in (3.3 22), it can be found that the error $\varepsilon^k_{p,q}$ is subject to (3.3 22) also. Hence the stability (or the propagation of the computing error) of this linear scheme DES can be analysed by the Fourier separation method as follows:

Let

$$\varepsilon^k_{p,q} = \lambda^k e^{i(p\theta_1 + q\theta_2)} \tag{3.3.24}$$

where λ is the factor of error increase per time step τ, $i=\sqrt{-1}$ and θ_1 or θ_2 is the wave number per spatial step of the error in the form of simple harmonic motion with arbitrary frequency.

Substituting ε for u in the square DES scheme (3.3.22) gives:

$$\begin{aligned}\lambda^{k+1}e^{i(p\theta_1+q\theta_2)} = (1-3\beta)\lambda^k e^{i(p\theta_1+q\theta_2)} &+ \frac{\beta}{2}\left[\lambda^k e^{i[(p-1)\theta_1+q\theta_2]} + \lambda^k e^{i[(p+1)\theta_1+q\theta_2]} + \lambda^k e^{i[p\theta_1+(q-1)\theta_2]} + \lambda^k e^{i[p\theta_1+(q+1)\theta_2]}\right] \\ &+ \frac{\beta}{4}\left[\lambda^k e^{i[(p-1)\theta_1+(q-1)\theta_2]} + \lambda^k e^{i[(p+1)\theta_1+(q+1)\theta_2]} + \lambda^k e^{i[(p-1)\theta_1+(q+1)\theta_2]} + \lambda^k e^{i[(p+1)\theta_1+(q-1)\theta_2]}\right]\end{aligned} \tag{3.3.25}$$

that is,

$$\lambda = (1-3\beta) + \frac{\beta}{2}\left[e^{-i\theta_1} + e^{i\theta_1} + e^{-i\theta_2} + e^{i\theta_2}\right] + \frac{\beta}{4}\left[e^{-i(\theta_1+\theta_2)} + e^{i(\theta_1+\theta_2)} + e^{-i(\theta_1-\theta_2)} + e^{i(\theta_1-\theta_2)}\right] \tag{3.3.26}$$

or

$$\lambda = (1-3\beta) + \beta(\cos\theta_1 + \cos\theta_2 + \cos\theta_1 \cdot \cos\theta_2) \tag{3.3.27}$$

Since

$$-1 \le \cos\theta_1 + \cos\theta_2 + \cos\theta_1 \cdot \cos\theta_2 \le 3 \tag{3.3.28}$$

there will be

$$1-4\beta \le \lambda \le 1 \tag{3.3.29}$$

The scheme (3.3.22) will be stable only when $|\lambda| \le 1$ so that the computing error will not be amplified and may even be reduced in the process. It requires

$$0 < \frac{b\tau}{h^2} \equiv \beta \le \frac{1}{2} \tag{3.3.30}$$

or

$$0 < \tau \le \frac{h^2}{2b} \tag{3.3.31}$$

That is, for the square DES scheme, the choice of the time step τ is restricted by the choice of spatial step h.

With a similar analysis, it can be determined that the stability limit of the oblique cross DES scheme (3.3.21) is just the same as (3.3.31). But the stability limit of the normal cross DES scheme (3.3.19) is different, which is

$$0 < \tau \le \frac{h^2}{4b} \tag{3.3.32}$$

Remark 3.3: *The oblique cross or square DES scheme has a larger stability range with respect to time step than the normal cross DES scheme.*

C. General square form of the DES mask

Using a weighted combination of the normal and the oblique cross DES schemes gives a more general DES scheme, which is a difference approximation to the diffusion equation at point (p,q) surrounded by all 8-connected neighbours:

$$u^{k+1}_{p,q} = \omega\,[(1-4\beta)\,u^k_{p,q} + \beta \begin{array}{ccc} & o & \\ o & + & o \\ & o & \end{array}] + (1-\omega)\,[(1-2\beta)\,u^k_{p,q} + \frac{\beta}{2} \begin{array}{ccc} o & & o \\ & \times & \\ o & & o \end{array}] \tag{3.3.33}$$

or

$$u^{k+1}_{p,q} = [1-2(1+\omega)\beta]\,u^k_{p,q} + \omega\beta \begin{array}{ccc} & o & \\ o & + & o \\ & o & \end{array} + \frac{(1-\omega)\beta}{2} \begin{array}{ccc} o & & o \\ & \times & \\ o & & o \end{array} \tag{3.3.34}$$

together with a *general DES mask* form

$\frac{(1-\omega)\beta}{2}$	$\omega\beta$	$\frac{(1-\omega)\beta}{2}$
$\omega\beta$	$1-2(1+\omega)\beta$	$\omega\beta$
$\frac{(1-\omega)\beta}{2}$	$\omega\beta$	$\frac{(1-\omega)\beta}{2}$

Figure 3.3.5 The general DES mask.

where and hereafter coefficient $0 \le \omega \le 1$ will be called as the *balance coefficient* of the scheme or the mask. With this DES scheme and the general DES mask, most of the 3x3 square masks of the Gaussian filter can be derived from it.

Remark 3.4: *Gaussian filter's isotropy is implied in the general DES mask with a balance coefficient* $\omega = \frac{1}{2}$, *i.e., the isotropic square DES mask in Figure 3.3.4.*

3.3.4. Understanding the RA mask from DES

The equivalence between Gaussian convolution and a diffusion process has been shown. Note that Gaussian filtering can be implemented with the repeated averaging method which works on a *statistical* basis, whereas the diffusion explicit smoothing method works with a diffusion equation, which works on a *deterministic* basis. The difference between both methods' origins is so great that it is interesting to find some possible relationships between RA and DES, as that has been found between GS and DS. The following theorem makes explicit this relationship:

Theorem 3.2. *The repeated averaging (RA) is just a special case of the diffusion explicit smoothing (DES) with the equation*

$$\frac{\partial u}{\partial t} = \tfrac{1}{6}\nabla^2 u \tag{3.3.35}$$

Proof: Let h = 1 and τ = 1, then $\beta = \frac{b\tau}{h^2}$ = b. The square DES scheme (3.3.22) thus becomes:

$$u_{p,q}^{k+1} = (1-3b)u_{p,q}^{k} + \frac{b}{2}\begin{matrix} & o & \\ o & + & o \\ & o & \end{matrix} + \frac{b}{4}\begin{matrix} o & & o \\ & \times & \\ o & & o \end{matrix} \tag{3.3.36}$$

Note that RA works from only the values at the current time, so it is an explicit recursion, and the mask can be rewritten into a difference equation:

$$u_{p,q}^{k+1} = \frac{12}{24}u_{p,q}^{k} + \frac{2}{24}\begin{matrix} & o & \\ o & + & o \\ & o & \end{matrix} + \frac{1}{24}\begin{matrix} o & & o \\ & \times & \\ o & & o \end{matrix} \tag{3.3.37}$$

Comparing equation (3 3 37) with (3.3.36) directly gives b = $\frac{1}{6}$.

Q.E.D.

Brady *et al* [1985] stated that the RA method "appeals to the central-limit theorem and implements Gaussian filtering using repeated averaging with the 3x3 mask .... Iterating n times approximately corresponds to filtering with a Gaussian whose standard deviation is proportional to $\sqrt{n}$". However, the following theorem shows this scale proportion constant can be determined *without* any prior assumption in statistics or probability theory.

Theorem 3.3. *The RA's scale proportional constant should be* $\frac{1}{\sqrt{3}}$.

Proof: According to the relation between the smoothing scale σ and the diffusion time t:

$$\sigma = \sqrt{2bt} \tag{3.3.38}$$

When t = 1, 2, ..., n, $\sigma = \sqrt{2b\,1}$, $\sqrt{2b\,2}$, · , $\sqrt{2b\,n}$. So, after repeating n times, RA will approximate Gaussian smoothing at the scale level $\sigma = c\sqrt{n}$, where $c = \sqrt{2b}$ is the scale proportion constant. From Theorem 3.2, b = $\frac{1}{6}$, hence

$$c = \frac{1}{\sqrt{3}} \tag{3.3.39}$$

Q.E.D.

Now that Theorem 3.2 has shown that RA is a special case of DES, the RA mask's numerical stability limit can be derived from the corresponding DES scheme.

Theorem 3.4. *The numerical stability limit corresponding to the RA mask is max* $\tau = 3$.

Proof: From Theorem 3.2, the RA mask corresponds to the diffusion equation

$$\frac{\partial u}{\partial t} = b\,\nabla^2 u \tag{3.3.40}$$

Setting $b = \frac{1}{6}$ and $h = 1$ in the stability limit of the square DES scheme:

$$0 < \tau \leq \frac{h^2}{2b} \tag{3.3.41}$$

gives the upper bound $\tau_{max} = 3$.

Q.E.D.

The above theorem shows a criterion to identify the "filtering rate" of a given mask. Moreover, it gives:

Corollary 3.1: *There are other masks which achieve the same filtering with fewer iterations than the RA mask; but the number of iterations N required by the RA mask can, at most, be reduced to N/3.*

Proof: From Theorem 3.2, the diffusion coefficient $b = \frac{1}{6}$. Let the number of iterations required by the RA mask be N, where $\tau = 1$. The filtering rate will be raised with a larger time step τ. By choosing a larger $\tau = 2$, the number of iterations is reduced to $N/2$. Since $\beta \equiv \frac{b\tau}{h^2} = \frac{1}{3}$, the required square DES mask is:

$$\frac{1}{12} \times \begin{array}{|c|c|c|} \hline 1 & 2 & 1 \\ \hline 2 & 0 & 2 \\ \hline 1 & 2 & 1 \\ \hline \end{array}$$

Figure 3.3.6. The square mask corresponding to τ=2 reduces the iteration number down to a half.

By choosing the maximal $\tau = 3$, the number of iteration is reduced to $N/3$. This is the least number of iterations to achieve the same smoothing as the RA mask. Since $\beta \equiv \frac{1}{2}$, the required square DES mask is:

$$\frac{1}{8} \times \begin{array}{|c|c|c|} \hline 1 & 2 & 1 \\ \hline 2 & -4 & 2 \\ \hline 1 & 2 & 1 \\ \hline \end{array}$$

Figure 3.3.7. The square mask corresponding to τ=3 reduces the iteration number down to one third.

Q.E.D.

Some papers, such as [Davies 1987], argued that the mask of the Gaussian filter or the so-called *Gaussian-like* mask should be in a specific square form as shown in the following figure:

$$\begin{array}{|c|c|c|} \hline x^2 & xy & x^2 \\ \hline xy & y^2 & xy \\ \hline x^2 & xy & x^2 \\ \hline \end{array}$$

Figure 3.3.8. The Gaussian-like mask proposed by Davies.

The main reason is based on Gaussian's symmetry and separability, denoted in the following matrix-vector decomposition form:

$$\begin{bmatrix} x^2 & xy & x^2 \\ xy & y^2 & xy \\ x^2 & xy & x^2 \end{bmatrix} = \begin{bmatrix} x \\ y \\ x \end{bmatrix} [x \ y \ x] \tag{3.3.42}$$

However, the well-known RA mask (in Figure 3.2.1) does not conform with this claim, nor do those RA-like masks (in Figure 3.2.2). And some masks derived from the DES scheme, such as both the normal and the oblique cross DES masks, are even further away from this and yet maintain the symmetry property.

So, one logical inference is that the above Kronecker decomposition form (3.3.42) is not an inherent attribute of the Gaussian-related masks. All these masks mentioned above could be derived from a common mask which has the Gaussian's isotropy as an intrinsic property. In fact, Davies' Gaussian-like mask (in Figure 3.3.8) is just a special form of the RA-like mask; and both of them can be uniformly derived from another mask, the general DES mask (in Figure 3.3.5).

Theorem 3.5. *The RA-like mask can be derived from the general DES mask when the balance coefficient* $\omega \in (0, 1)$*, and the Gaussian-like mask can be derived from the RA-like mask with the further constraint* $\omega = 1-2\beta$.

Proof: The range of the balance coefficient ω is $[0, 1]$ in the general DES mask below:

$\frac{(1-\omega)\beta}{2}$	$\omega\beta$	$\frac{(1-\omega)\beta}{2}$
$\omega\beta$	$1-2(1+\omega)\beta$	$\omega\beta$
$\frac{(1-\omega)\beta}{2}$	$\omega\beta$	$\frac{(1-\omega)\beta}{2}$

Figure 3.3.9 The general DES mask (same to Figure 3.3.5).

$\omega = 0$ gives the oblique cross mask in Figure 3.3.3. And $\omega = 1$ gives the normal cross mask in Figure 3.3.2.

Apart from these two cases, masks are given in the RA-like form. Because $0 < \beta$ and $0 < \omega < 1$, all the 8-connected neighbouring grids in this mask are non-zero, thus the mask must be a 3x3 square. E.g., when $\omega = \frac{1}{2}$, it is the square DES mask (in Figure 3.3.4); setting $\beta = \frac{1}{6}$, it becomes Brady's RA mask.

To get Davies' Gaussian-like mask, a further constraint should be imposed on the RA-like mask, that is

$$\frac{(1-\omega)\beta}{2} [1-2(1+\omega)\beta] = (\omega\beta)^2 \tag{3.3.43}$$

Hence

$$\omega = 1 - 2\beta \tag{3.3.44}$$

and the general DES mask becomes:

β^2	$\beta\omega$	β^2
$\beta\omega$	ω^2	$\beta\omega$
β^2	$\beta\omega$	β^2

Figure 3.3.10. The Gaussian-like mask derived from the RA-like mask ($0<\omega<1$).

Q.E.D.

Remark 3.5: *Both RA-like masks in Figure 3.2.2 can be derived from the general DES mask in Figure 3.3 9 using* $\omega = \frac{1}{2}$, $\beta = \frac{1}{4}$ *and* $\omega = \frac{2}{3}$, $\beta = \frac{1}{5}$, *respectively.*

Corollary 3.2: *The Gaussian-like mask corresponds to the diffusion equation* $\frac{\partial u}{\partial t} = b\nabla^2 u$ *when* $0 < b < \frac{1}{2}$.

Proof: From Theorem 3.5, there are $0 < \omega < 1$ and $\omega = 1 - 2\beta$, that is $0 < \beta < \frac{1}{2}$.

For the Gaussian-like mask, $\tau = 1$ and $h = 1$. Hence $b = \frac{b\tau}{h^2} \equiv \beta$, and the conclusion follows.

Q.E.D.

3.3.5. Diffusion implicit smoothing schemes (DIS)

It has been seen that the time step of the explicit scheme is usually restricted by the stability limit. The time step of the implicit scheme will be shown to be free from the restriction owing to the unconditionally stable property. Thus for higher computational efficiency, some simple implicit schemes will be preferred since they can take a larger time step and have fewer terms to calculate.

A. Implicit numerical schemes of diffusion equation

One implicit form is

$$\frac{1}{\tau}(u_{p,q}^{k+1} - u_{p,q}^{k}) = \frac{b}{h^2}(u_{p-1,q}^{k+1} - 2u_{p,q}^{k+1} + u_{p+1,q}^{k+1}) + \frac{b}{h^2}(u_{p,q-1}^{k+1} - 2u_{p,q}^{k+1} + u_{p,q+1}^{k+1}) \qquad (3.3.45)$$

or its vector form:

$$\begin{bmatrix} 0 & -\beta & 0 & -\beta & 1{+}4\beta & -\beta & 0 & -\beta & 0 \end{bmatrix} \begin{bmatrix} u_{p-1,q-1} \\ u_{p-1,q} \\ u_{p-1,q+1} \\ u_{p,q-1} \\ u_{p,q} \\ u_{p,q+1} \\ u_{p+1,q-1} \\ u_{p+1,q} \\ u_{p+1,q+1} \end{bmatrix}^{k+1} = u_{p,q}^{k} \qquad (3.3.46)$$

where $p, q = 0, 1, 2, \ldots, M$, M is the width of a surface image.

It leads to a linear system with a tridiagonal-block coefficient matrix:

$$\begin{bmatrix} \mathbf{A}_0 & -\beta\mathbf{I} & & & & \\ -\beta\mathbf{I} & \mathbf{A}_1 & -\beta\mathbf{I} & & & \\ & \cdot & \cdot & \cdot & & \\ & & \cdot & \cdot & \cdot & \\ & & & \cdot & \cdot & \cdot \\ & & & -\beta\mathbf{I} & \mathbf{A}_{M-2} & -\beta\mathbf{I} \\ & & & & -\beta\mathbf{I} & \mathbf{A}_{M-1} \end{bmatrix} \begin{bmatrix} \mathbf{U}_0 \\ \mathbf{U}_1 \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{U}_{M-2} \\ \mathbf{U}_{M-1} \end{bmatrix}^{k+1} = \begin{bmatrix} \mathbf{U}_0 \\ \mathbf{U}_1 \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{U}_{M-2} \\ \mathbf{U}_{M-1} \end{bmatrix}^{k} \tag{3.3.47}$$

where sub-matrix $\mathbf{A}_p$ (p = 0, 1, ..., M-1) is an M×M tri-diagonal matrix, **I** is an M×M unit matrix, and $\mathbf{U}_p$ (p = 0, 1, ..., M-1) is an M-dimensional vector $\mathbf{U}_p = (u_{p,0}\ u_{p,1} \ldots u_{p,M-1})$ at time k or k+1, and M is the image width. Since solution of this system needs more computation than Gaussian convolution or repeated averaging, this scheme should be discarded.

With regards to the separability of Gaussian convolution and the well-known Crank-Nicolson scheme [Crank and Nicolson 1947][Douglas and Rachford 1956] [Douglas 1961] which alternatively applies implicit schemes in the x and y directions as below:

$$\frac{2}{\tau}(u_{p,q}^{k+\frac{1}{2}} - u_{p,q}^{k}) - \frac{b}{h^2}(u_{p+1,q}^{k+\frac{1}{2}} - 2u_{p,q}^{k+\frac{1}{2}} + u_{p-1,q}^{k+\frac{1}{2}}) - \frac{b}{h^2}(u_{p,q+1}^{k} - 2u_{p,q}^{k} + u_{p,q-1}^{k}) = 0 \tag{3.3.48x}$$

$$\frac{2}{\tau}(u_{p,q}^{k+1} - u_{p,q}^{k+\frac{1}{2}}) - \frac{b}{h^2}(u_{p+1,q}^{k+\frac{1}{2}} - 2u_{p,q}^{k+\frac{1}{2}} + u_{p-1,q}^{k+\frac{1}{2}}) - \frac{b}{h^2}(u_{p,q+1}^{k+1} - 2u_{p,q}^{k+1} + u_{p,q-1}^{k+1}) = 0 \tag{3.3.48y}$$

a simpler, split-type and fully implicit scheme from [Feng *et al* 1978] with the initial value condition is as follows:

$$\frac{1}{\tau}(u_{p,q}^{k+\frac{1}{2}} - u_{p,q}^{k}) - \frac{b}{h^2}(u_{p+1,q}^{k+\frac{1}{2}} - 2u_{p,q}^{k+\frac{1}{2}} + u_{p-1,q}^{k+\frac{1}{2}}) = 0 \tag{3.3.49x}$$

$$\frac{1}{\tau}(u_{p,q}^{k+1} - u_{p,q}^{k+\frac{1}{2}}) - \frac{b}{h^2}(u_{p,q+1}^{k+1} - 2u_{p,q}^{k+1} + u_{p,q-1}^{k+1}) = 0 \tag{3.3.49y}$$

$$u_{p,q}^{0} = f_{p,q} \qquad p,q = 0\ (1)\ M{-}1, \quad k = 0\ (1)\ \infty \tag{3.3.49i}$$

Since the above scheme will be used with a changeable time step τ during computing, it will be called the *DISCT* scheme (Diffusion Implicit Smoothing with Changeable Time step scheme).

Using the Taylor expansion, it is easy to show that as two separated difference equations, each of them is incompatible with the diffusion problem (3.3.1), but as a whole, the DISCT scheme is certainly compatible with (3.3.1). Adding (3.3.49.y) and (3.3.49.x) gives a simple verification.

The DISCT scheme can be represented in a matrix form:

$$\mathbf{A}\ \mathbf{U}_q^{k+\frac{1}{2}} = \mathbf{U}_q^{k} \tag{3.3.50x}$$

$$\mathbf{A}\ \mathbf{V}_p^{k+1} = \mathbf{V}_p^{k+\frac{1}{2}} \tag{3.3.50y}$$

$$\mathbf{U}_q^{0} = \mathbf{F}_q \tag{3.3.50i}$$

$$p,q = 0\ (1)\ M{-}1, \quad k = 0\ (1)\ \infty$$

where

$$\mathbf{A} = \begin{bmatrix} \alpha & -\beta & & & & \\ -\beta & \alpha & -\beta & & & \\ & \cdot & \cdot & \cdot & & \\ & & \cdot & \cdot & \cdot & \\ & & & \cdot & \cdot & \cdot \\ & & & -\beta & \alpha & -\beta \\ & & & & -\beta & \alpha \end{bmatrix} \tag{3.3.51}$$

$$\alpha = 1 + 2\beta, \qquad \beta = \frac{b\tau}{h^2} \tag{3.3.52}$$

$$U_q^k = \begin{bmatrix} u_{0,q}^k \\ u_{1,q}^k \\ \\ \\ \dot{u}_{M,q}^k \end{bmatrix} \qquad U_q^0 = \begin{bmatrix} f_{0,q} \\ f_{1,q} \\ f_{2,q} \\ \cdot \\ \cdot \\ \cdot \\ f_{M,q} \end{bmatrix} \equiv F_q \tag{3.3.53}$$

$$V_p^k = [u_{p,0}^k \quad u_{p,1}^k \qquad \cdot \quad u_{p,M}^k]^T \tag{3.3.54}$$

$$p,q = 0\ (1)\ M, \quad k = 0\ (\tfrac{1}{2})\ \infty$$

Note that A is a tridiagonal and diagonal-dominant matrix. Such a linear algebra system (3.3.50) can be easily solved with numerical stability.

B. Stability analysis of the implicit scheme DISCT

Koenderink [1984] and Gourlay [1985] have mentioned the numerical stability problem of diffusion smoothing. Here, DISCT is shown to be an *unconditionally* stable numerical scheme.

Because the DISCT scheme is also a linear system, the Fourier separation method can be applied to it, which is similar to the stability analysis of the explicit scheme.

Let

$$\varepsilon_{p,q}^k = \lambda^k e^{\iota(p\theta_1 + q\theta_2)} \tag{3.3.55}$$

and

$$\beta = \frac{b\tau}{h^2} > 0 \tag{3.3.56}$$

Substituting ε for u in (3.3.49x) of the DISCT scheme gives:

$$\lambda_1^{-\frac{1}{2}} = 1 - \beta(2\cos\theta_1 - 2) = 1 + 4\beta\sin^2\frac{\theta_1}{2} \tag{3.3.57}$$

that is

$$\lambda_1 = \left[1 + 4\beta\sin^2\frac{\theta_1}{2}\right]^{-2} \tag{3.3.58x}$$

Similarly, from (3.3.49y), there is

$$\lambda_2 = \left[1 + 4\beta\sin^2\frac{\theta_2}{2}\right]^{-2} \tag{3.3.58y}$$

So, the final result is:

$$\lambda = \lambda_1 \lambda_2 \leq 1 \tag{3.3.59}$$

Hence, computational error will not be amplified by operations whatever the value β takes in (3.3.58), so the DISCT scheme is unconditionally stable. A larger time step τ in computing is therefore permitted, which is expected to significantly reduce the computational complexity.

3.3.6. Complexity estimation of the DISCT scheme in scale space

Suppose the spatial step h = 1 is the same as that of GS, (*i e* the minimal step or the data sampling step) Let each time step $\tau = 1$. To smooth an M×M pixel image, the DISCT scheme needs to solve only M pairs of M order linear algebra systems in the x and y directions respectively

Note that solving one M order tridiagonal matrix equation needs only 5M multiplication and division operations, therefore, only $10M^2$ operations per time step are required.

Thus, corresponding to $\sigma = 0\ (1)\ K$, the volume of multiplication and division operations over $t = 0\ (1)\ \frac{K^2}{2b}$ will be *at most*

$$C_{DIS} = 10M^2 \times \frac{K^2}{2b} = \frac{5}{b}K^2M^2 = O(K^2M^2) \tag{3.3.60}$$

This conservative estimate can be further reduced if a changeable time step τ>1 is used in a relatively long time interval $[0, \frac{K^2}{2b}]$ with the assurance from the unconditional stability of the DISCT scheme. *E g* , when a changeable time step

$$\tau = \frac{2k+1}{6b} \tag{3.3.61}$$

is adopted in each time sub-interval

$$[\frac{k^2}{2b}, \frac{(k+1)^2}{2b}] \qquad 1 \leq k \leq K \tag{3.3.62}$$

the order of complexity in (3.3.60) is reduced to:

$$C_{DISCT} = 10(3K-2)M^2 = O(KM^2) \qquad K \geq 1 \tag{3.3.63}$$

A numerical comparison of complexity estimations of GS and DISCT in scale space can be found in Appendix 1. It will show that DISCT is faster than GS and the finite scale space produced by DISCT is denser than that of GS.

3.3.7. Boundary treatment of DS

So far, diffusion smoothing has been considered as an initial value problem of the diffusion equation and in this sense DS is equivalent to GS.

However, for many cases in depth data smoothing, especially, those with laser range data, the surface will have obvious boundaries in the image [Brady *et al* 1985]. A simple depth threshold, say, 5 units in depth is enough to separate surfaces from the background. As these boundaries should not be blurred while being smoothed, the diffusion problem is actually a boundary value problem. For this reason, a boundary condition should be imposed on the equation so as to preserve the distinction between object and background.

Parallel to boundary treatment in GS, where "any (computational) molecules across the boundary should be abandoned wholly" [Brady *et al* 1985] [Terzopoulos 1985], in DS any node outside the surface should make no contribution to the surface smoothing, and keep its current values. So, the boundary condition can be imposed on the diffusion equation using a simple modification to the coefficient matrix A in (3.3.52).

In detail, in computing along the x-direction, *i.e.*, for each $q = const$ and $p = 0\ (1)\ M$, the modification can be done in three cases as follows:

1) If (p,q) is a node outside the surface being smoothed, then the symmetric matrix A has:

$$a_{p,p-1} = 0, \qquad a_{p,p} = 1, \qquad a_{p,p+1} = 0. \tag{3.3.64a}$$

2) If (p,q) is a node on the surface and (p-1,q) is a node outside the surface, then

$$a_{p,p-1} = 0, \qquad a_{p,p} = \gamma, \qquad a_{p,p+1} = -\beta. \tag{3.3.64b}$$

3) If (p,q) is a node on the surface and (p+1,q) is a node outside the surface, then

$$a_{p,p-1} = -\beta, \qquad a_{p,p} = \gamma, \qquad a_{p,p+1} = 0 \tag{3.3.64c}$$

$$where \quad \gamma \equiv \alpha - \beta = 1 + \beta$$

And in computing along the y-direction, q takes the place of p in all revisions.

Hence, after imposition of the boundary condition, the former tri-diagonal matrix **A** becomes a block-diagonal matrix of the following form:

$$\bar{\mathbf{A}} = \begin{bmatrix} \mathbf{A}_1 & & & & \\ & \mathbf{A}_2 & & & \\ & & & & \\ & & & & \\ & & & \mathbf{A}_{s-1} & \\ & & & & \mathbf{A}_s \end{bmatrix} \tag{3.3.65}$$

where each sub-block $\mathbf{A}_s$ $(1 \le s \le M)$ takes the form of either a tridiagonal matrix or a unit matrix.

So, $\bar{\mathbf{A}}$ is still consistent with A. In fact, the imposition (3.3.64) can be derived from the scheme (3.3.49) itself if the surface margin has been extended one pixel outward in the horizontal direction [Cai 1987a], and this means the diffusion equation is solved with a "no leakage" boundary treatment model [Cai 1988a].

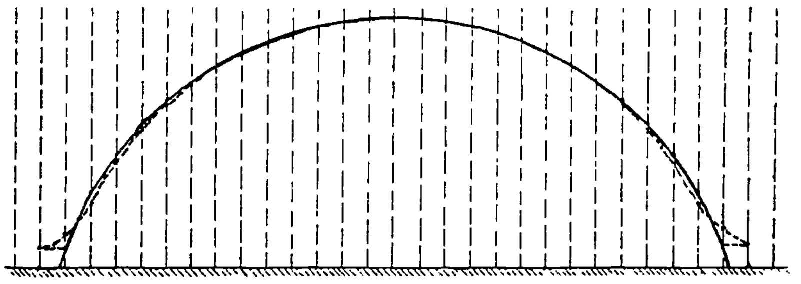


Figure 3.3.11. Smoothing a cylindrical surface with a "one pixel horizontal boundary extension".

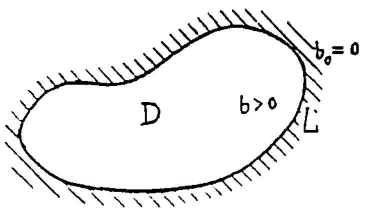


Figure 3.3.12. The "no leakage" heat conduction model, where the shading denotes heat-insulation.

Unfortunately, with such a boundary treatment, diffusion smoothing will also introduce some shape distortion just as with Gaussian smoothing where the signs of Gaussian curvature or mean curvature will change in a narrow strip along the surface boundary. The distortion is an embarrassing problem for multi-scale processing since it will aggravate and further propagate from the boundary into the internal area at larger scale levels, as shown in Figure 3.3.14. So, without a proper boundary treatment, it is difficult to implement scale space smoothing of surfaces.

Gaussian convolution or repeated averaging lacks means to preserve the curvature signs along the boundary. Ponce and Brady [1987] proposed estimating the 3-D normal at each point of the surface and moving the point a certain distance along its normal instead of smoothing the depth z. But the result of Gaussian smoothing is no longer orientation-dependent and it is rather expensive to compute.

With DS, this task seems much simpler than in GS. Cai [1988a] proposed a "small leakage" model for boundary treatment which allows a small amount of heat energy to leak out from the pixels at the slanting surface boundary by adjusting the diffusion coefficient from zero to a small positive value but maintains heat-insulation across the pixels at the horizontal surface boundary by keeping the diffusion coefficient zero.

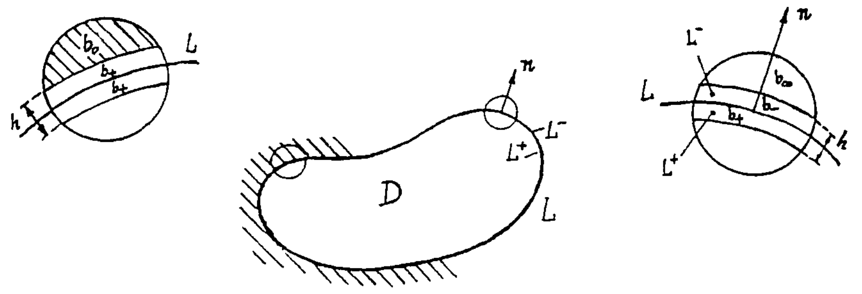


Figure 3.3.13. A "small leakage" heat conduction model, where the shading denotes heat-insulation.

In the above figure, D is the surface domain, L is the boundary of D, L^+ is the internal side and L^-, the external side; n is the outward normal vector at the boundary L; b is the heat conduction coefficient: b_+ for the surface material, e.g. 0.5, b_- for a thin heat-leak medium, about 10% of the b_+, e.g., 0.05, b_0 for heat-insulated medium, and $b_\infty = \infty$ for an ideal heat conductor so that the leaking heat energy will not accumulate outside the boundary L.

In fact, the "small leakage" model applies a "natural" boundary condition (the law of energy conservation):

$$\left[b_- \cdot \frac{\partial u}{\partial n} \right]_{L^-} = \left[b_+ \cdot \frac{\partial u}{\partial n} \right]_{L^+} \qquad (3.3.67)$$

to the diffusion equation:

$$\frac{\partial u}{\partial t} = b\nabla^2 u \tag{3.3.68}$$

In practice, to avoid computing the 2-D normal vector at each boundary pixel, this natural boundary condition is further simplified. It is split into the x and y directions as follows:

In the x-direction:

$$\left[b_- \cdot \frac{\partial u}{\partial x}\right]_{L^-} = \left[b_+ \cdot \frac{\partial u}{\partial x}\right]_{L^+} \tag{3.3.69x}$$

and in the y-direction:

$$\left[b_- \frac{\partial u}{\partial y}\right]_{L^-} = \left[b_+ \cdot \frac{\partial u}{\partial y}\right]_{L^+} \tag{3.3.69y}$$

Then discretising the x-directional gradient at the boundary pixel (p,q) at time k+$\frac{1}{2}$:

$$b_- \left(u_{p,q}^{k+\frac{1}{2}} - u_{p-1,q}^{k+\frac{1}{2}}\right) = b_+ \left(u_{p+1,q}^{k+\frac{1}{2}} - u_{p,q}^{k+\frac{1}{2}}\right) \tag{3.3.70x}$$

and the y-directional gradient at time k+1:

$$b_- \left(u_{p,q}^{k+1} - u_{p,q-1}^{k+1}\right) = b_+ \left(u_{p,q+1}^{k+1} - u_{p,q}^{k+1}\right) \tag{3.3.70y}$$

These can be expressed in the following vector forms:

$$\begin{bmatrix} b_- & -(b_-+b_+) & b_+ \end{bmatrix} \begin{bmatrix} u_{p-1,q}^{k+\frac{1}{2}} \\ u_{p,q}^{k+\frac{1}{2}} \\ u_{p+1,q}^{k+\frac{1}{2}} \end{bmatrix} = 0 \tag{3.3.71x}$$

and

$$\begin{bmatrix} b_- & -(b_-+b_+) & b_+ \end{bmatrix} \begin{bmatrix} u_{p,q-1}^{k+1} \\ u_{p,q}^{k+1} \\ u_{p,q+1}^{k+1} \end{bmatrix} = 0 \tag{3.3.71y}$$

Obviously, these coefficient vectors are compatible with the DISCT scheme's tridiagonal coefficient matrix (3.3.51). Hence, when the natural boundary condition is applied to a diffusion equation, there is only a small revision to matrix A corresponding to the slanting boundary pixels and it requires little computation. Matrix **A** holds its diagonal-dominant form, and the DISCT scheme's complexity estimate in subsection 3.3.6 still applies.

3.3.8. Some experimental results

Experimental results for two examples are shown in Figure 3.3.14 and 3.3.15.

The first example is a light bulb given in the real range data. The left half part is for DISCT smoothing, where the first column is the side-view of the light bulb surface from the bottom, the second column is the cosine-shading image of the surface and the third column is the KH sign image of the surface. The right three columns are for GS correspondingly. From the first row to the fourth row, smoothing scales are diffusion time $t = 0, 1, 4, 9$ corresponding to Gaussian scale $\sigma = 0, 1, 2, 3$.

From the KH sign images, it is easy to see that at $t = 9$ or $\sigma = 3$ most noise is reduced by DISCT or GS at the 1st and 4th columns respectively. However, the surface shape is obviously different in both processing. As shown in the 4th and 5th columns, GS introduces shape distortion which destroys the spheric shape of the bulb body, produces a narrow strip along the surface boundary and propagates to the internal area when the scale increases whereas the DISCT method preserves the surface shape quite well in the 1st and 3rd columns.

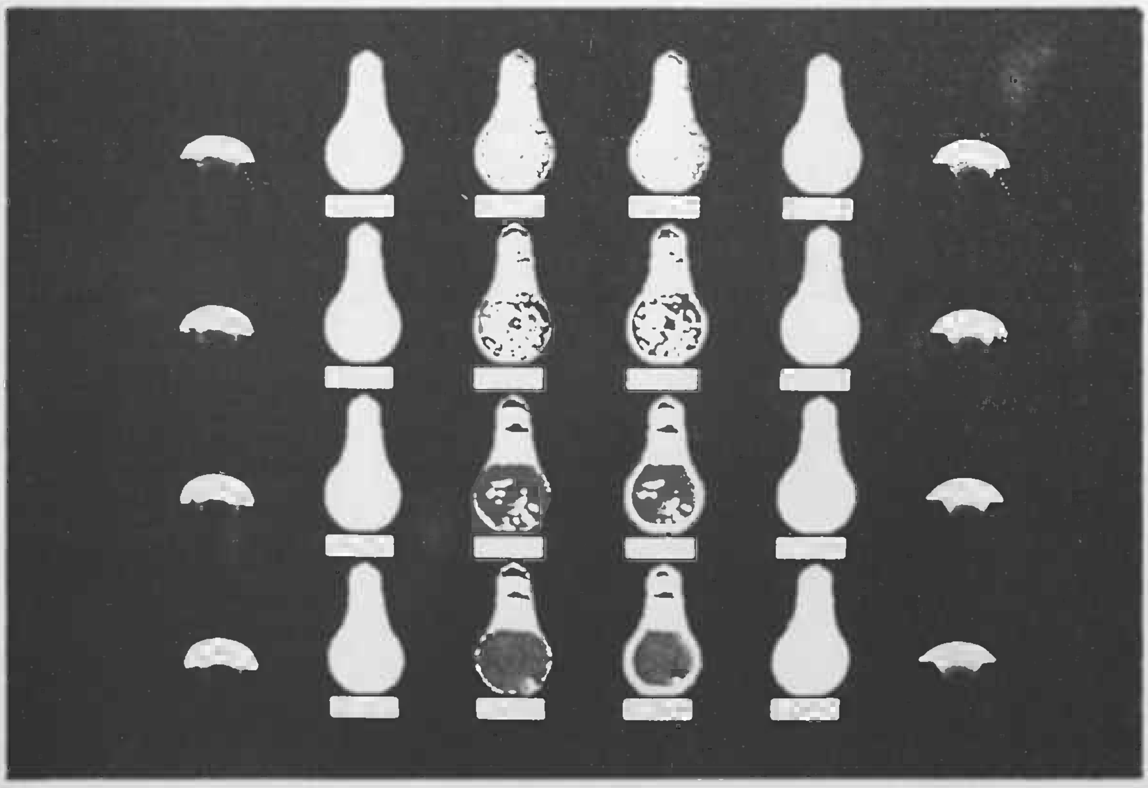


Figure 3.3.14. The surface of a light bulb smoothed by DISCT and GS respectively.

The second example is a dumb-bell shape given in the synthetic range data. The dumb-bell surface is smoothed by DISCT at scale levels $t = 0, 1, 4, 9$. From the KH sign images in Figure 3.3.15, it can be seen that both the boundary position and the surface shape of the dumb-bell are preserved well by DISCT hence there is little difference from the leftmost KH image to the rightmost one where both semi-spheres have not merged into one as described in other papers (*e.g.*, Babaud *et al* [1986]).

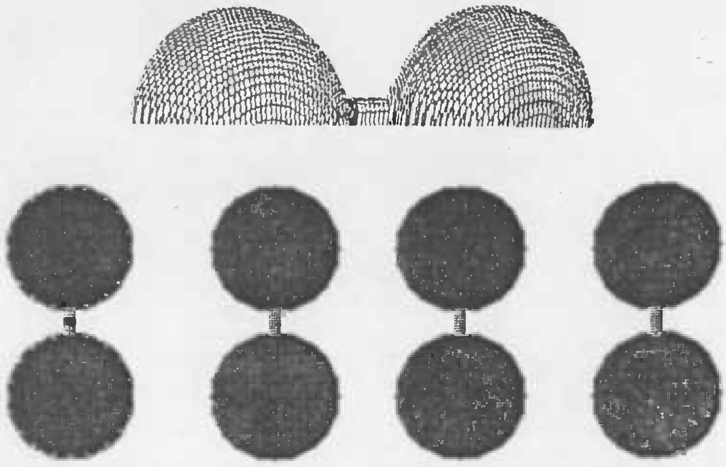


Figure 3.3.15. The KH sign image of the surface of a dumb-bell smoothed by DISCT.

3.4. Understanding Spline smoothing from Diffusion smoothing

The discussion on DS, GS and RA makes explicit their relationships, but it has not been exhaustive. It provides a more general tool to explain other popular smoothing methods from the viewpoint of DS even though they are quite different in appearance. In return, this implies that the DS theory can provide a uniform theoretical base for different smoothing methods. Spline smoothing (SS) is shown as an example.

3.4.1. Surface smoothing using a cubic B-spline approach

The spline surface approximation not only gives a smooth version of the surface but also preserves the convexity/concavity of the overall surface shape in the processing.

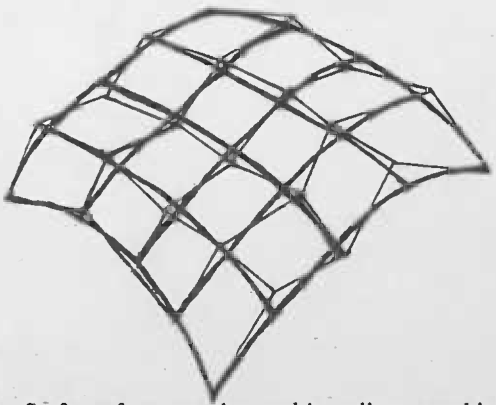


Figure 3.4.1. Surfaces from raw date and its spline smoothing approximation.

Speaking mathematically, given a surface $u(x,y) > 0$ in the form of discrete data $\{u_{p,q} \approx u(x_p,y_q) \,|_{p,q = 0\,(1)\,m-1}\}$ on a square mesh π: $[x_p = x_0 + ph,\ \ y_q = y_0 + qh \,|_{p,q = 0\,(1)\,m-1}]$, where m is the mesh width, $h>0$ is the mesh step and the intersection of the surface boundary L and the mesh π is a set $L_\pi = \{(s,r) \,|_{0 < u_{s,r}\,,\ (s,r) \in L\,,\ 0 \le s,r \le m-1}\}$.

First, along the boundary L_π, extending the data surface $\{u_{p,q} \,|_{p,q = 0\,(1)\,m-1}\}$ one node outwards from the surface gives:

$$u_{s-1,r} = 2u_{s,r} - u_{s+1,r} \qquad \text{if } u_{s-1,r} \,\overline{\in}\, L$$

$$u_{s+1,r} = 2u_{s,r} - u_{s-1,r} \qquad \text{if } u_{s+1,r} \,\overline{\in}\, L \qquad (3.4.1x)$$

$$u_{s,r-1} = 2u_{s,r} - u_{s,r+1} \qquad \text{if } u_{s,r-1} \,\overline{\in}\, L$$

$$u_{s,r+1} = 2u_{s,r} - u_{s,r-1} \qquad \text{if } u_{s,r+1} \,\overline{\in}\, L \qquad (3.4.1y)$$

Second, holding y constant in u(x,y) and approximating the surface u(x,y) in the x-direction produces the spline surface:

$$S_x u(x,y) = \sum_{p=-1}^{m} u(x_p,y)\ \phi_p(x) \qquad (3.4.2x)$$

Holding x constant in $S_x u(x,y)$ and approximating the surface $S_x u(x,y)$ in the y-direction produces the final spline surface.

$$S_y S_x u(x,y) = \sum_{q=-1}^{m} S_x u(x,y_q)\ \psi_q(y) = \sum_{q=-1}^{m} \sum_{p=-1}^{m} u(x_p,y_q)\ \phi_p(x)\ \psi_q(y) \qquad (3.4.2y)$$

Third, replacing $u(x_p,y_q)$ with $u_{p,q}$ gives discrete values of the spline approximation of the surface u(x,y):

$$\hat{u}(x,y) = \sum_{q=-1}^{m} \sum_{p=-1}^{m} u_{p,q}\ \phi_p(x)\ \psi_q(y) \qquad (3.4.3)$$

where $\phi_p(x)$ and $\psi_q(y)$ are in the forms of cubic B-spline function $\Omega_3(\cdot)$:

$$\phi_p(x) = \Omega_3\left(\frac{x-x_p}{h}\right) = \Omega_3\left(\frac{x-x_0}{h}-p\right) \qquad (3.4.4x)$$

$$\psi_q(y) = \Omega_3\left(\frac{y-y_q}{h}\right) = \Omega_3\left(\frac{y-y_0}{h}-q\right) \qquad (3.4.4y)$$

and the piecewise expressions of cubic B-spline function $\Omega_3(\cdot)$ are

$$\Omega_3(x) = \begin{cases} 0 & \text{if } |x| \geq 2 \\ \frac{1}{2}|x|^3 - x^2 + \frac{2}{3} & \text{if } |x| \leq 1 \\ -\frac{1}{6}|x|^3 + x^2 - 2|x| + \frac{4}{3} & \text{if } 1 < |x| < 2 \end{cases} \qquad (3.4.5)$$

So, after spline smoothing, the input data $\{u_{p,q}\,|_{p,q\,=\,0\ (1)\ m-1}\}$ will be changed to

$$\hat{u}_{p,q} = \sum_{i=-1}^{m} \sum_{j=-1}^{m} u_{i,j}\ \Omega_3(p-i)\ \Omega_3(q-j) \qquad (3.4.6)$$

Owing to the properties of the cubic B-spline functions [Ahlberg *et al* 1967][Li 1978][Reinsch 1967], a single spline smoothing has the following characteristics:

(1) The surface approximation is C^2 continuous and has $O(h^2)$ accuracy.

(2) The surface convexity and concavity are preserved.

(3) The computation can be done locally and in parallel within 3x3 windows.

For simplicity, only a cross-section of the surface in the x-direction along with its spline approximation is illustrated below:

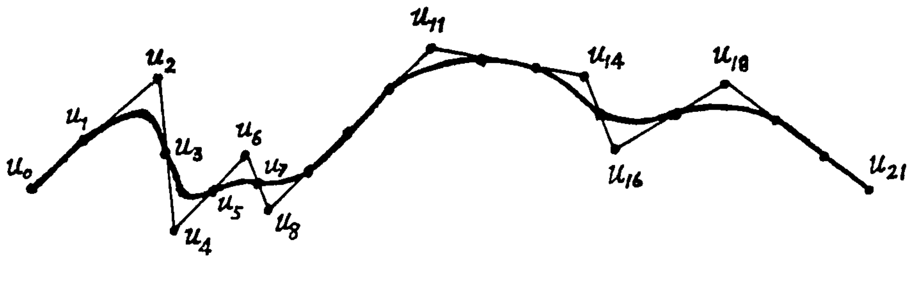


Figure 3 4 2 A cross section of the surface approximated by a spline function with convexity and concavity preservation

Repeatedly using this spline surface approach, the approximation will converge to a smooth surface with C^2 continuity that preserves the global shape of the surface. So, this spline surface approach can be seen as another surface smoothing method.

3.4.2. Spline smoothing, Gaussian smoothing and repeated averaging

It is easy to prove that surface spline smoothing is a special case of repeated averaging or discrete Gaussian smoothing. This result should not be surprising if noting that by using the regularisation theory Poggio *et al* [1984] derived an optimal smoothing operator which is just a spline Their conclusion is: "The solution to the variational problem $\sum_{k=1}^{n}(f_k - S(x_k))^2 + \lambda \int |S''(x)|^2 dx$ in the case of inexact data on a regular grid (and appropriate boundary conditions) can be obtained (a) by convolving the data with a filter, (b) which is a cubic spline, and (c) which is very similar to a Gaussian."

It is interesting to note that from the viewpoint of diffusion smoothing their result can be proved in the discrete case and in a simpler and more straightforward style.

Theorem 3.6. *Cubic B-spline surface smoothing is a special case of repeated averaging or discrete Gaussian convolution*

Proof: Due to the compactness of the cubic B-spline functions, for any mesh node $(p-i)$ or $(q-j)$ which is beyond the interval (-2, 2) there must be

$$\Omega_3(p-i) \equiv 0 \qquad \text{or} \qquad \Omega_3(q-j) \equiv 0 \tag{3.4.7}$$

Thus the spline approximation formula (3.4.6) can be rewritten as:

$$\hat{u}_{p,q} = \sum_{j=q-1}^{q+1} \sum_{i=p-1}^{p+1} u_{i,j}\, \Omega_3(p-i)\, \Omega_3(q-j) \tag{3.4.8}$$

Note that

$$\Omega_3(0) \equiv \frac{2}{3} \qquad \text{and} \qquad \Omega_3(\pm 1) \equiv \frac{1}{6} \tag{3.4.9}$$

A typical Gaussian-like mask form of the surface spline smoothing is:

$$\frac{1}{36} \times \begin{array}{|c|c|c|}\hline 1 & 4 & 1 \\ \hline 4 & 16 & 4 \\ \hline 1 & 4 & 1 \\ \hline \end{array}$$

Figure 3.4.3. Surface spline smoothing mask.

So, spline smoothing is a special case of RA. Repeatedly using this mask approximately corresponds to filtering with a Gaussian.

Q.E.D.

3.4.3. Spline smoothing and diffusion explicit smoothing

Spline smoothing is now further interpreted as the explicit diffusion smoothing associated with a specific boundary treatment.

Theorem 3.7. *Surface spline smoothing is a diffusion process with a "convex corner cling" boundary condition:*

$$\frac{\partial u}{\partial t} = \frac{1}{6}\nabla^2 u \tag{3.4.10e}$$

$$u\big|_{t=0} = f \tag{3.4.10i}$$

$$u\big|_{L_\pi^o} = f\big|_{L_\pi^o} \tag{3.4.10b}$$

where f is the input data surface given in the initial time of the diffusion process, L is the boundary of the surface domain D and $L_\pi^o = \{(s,r)\,\big|_{(s,r)\in L,\ 0 < u_{s,r}^k \equiv u_{s,r}^o,\ 0\le s,r\le m-1,\ 0\le k<\infty}\}$ *is the set of convex corner*[1] *boundary nodes, which is a subset of the boundary node set* L_π.

Proof: From Theorem 3.5, repeated averaging is a special case of explicit diffusion smoothing. Comparing the general DES mask (in Figure 3.3.9) with the spline smoothing mask (in Figure 3.4.3) gives $\omega = \frac{2}{3}$ and $\beta = \frac{1}{6}$. It yields the diffusion coefficient $b = \frac{1}{6}$ when $\tau = h = 1$. So spline smoothing is a special case of explicit diffusion smoothing.

Because the boundary treatment is performed with extension formula (3.4.1), surface values at some boundary points are invariant during the smoothing process. These points are just the convex corner nodes $L_\pi^o \subset L_\pi$ shown in Figure 3.4.4.

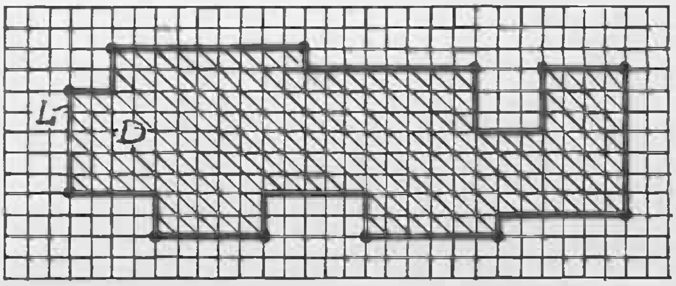


Figure 3.4.4. Convex corner node set L_π^o (blackened dot nodes of L_π) in spline smoothing.

This means that under such a boundary treatment, the spline surface must cling to the input data surface on the node set L_π^o throughout the diffusion process. In return, it is named as the treatment with the "convex corner cling" boundary condition.

Q.E.D.

---

[1] Moving along the boundary L and linking two neighbouring nodes, for example, l_{i-1} and l_{i+1} around a corner node l_i of L_π, if the line segment $\overline{l_{i-1}l_{i+1}}$ falls into the domain D surrounded by L, this corner node is called convex corner boundary node.

Corollary 3.3. *Repeatedly using the surface spline approach n times approximately corresponds to filtering with a Gaussian whose standard deviation is* $\sqrt{\frac{n}{3}}$.

Proof: It is similar to the proof of Theorem 3.3.

Q.E.D.

Remark 3.6: *While the spline smoothing mask and Brady's RA mask share the same diffusion equation* $\frac{\partial u}{\partial t} = \frac{1}{6}\nabla^2 u$, *they do not share the same balance coefficient* ω *in the general DES mask*

Spline smoothing is attractive because of three characteristics listed before. However, there is no intention of replacing diffusion smoothing with spline smoothing because of the difference in scale space behaviour at the surface boundary and computational complexity.

First, cubic B-spline smoothing is an explicit diffusion smoothing with stability condition $\beta = \frac{1}{6}$ which promises a much lower computational efficiency than DISCT in scaled space.

Second, spline smoothing's surface boundary treatment leads to the spline surface cling to the input data surface at convex corner nodes along the boundary. When the data is noise-free and approximated by one-step iteration, where the surface shape has not been distorted, this adherence would not be a problem, even an advantage. However, when the data is noisy and smoothed repeatedly, the adherence might provide false information about the surface tendency at the boundary which cannot be corrected throughout the whole process. As a comparison, the surface boundary treatment is better when using diffusion smoothing with the "small leakage" model.

In addition, the balance coefficient $\omega = \frac{2}{3}$ suggests spline smoothing is not so good a filter from the viewpoint of the isotropy as will be discussed in the next section.

3.5. Other Applications

The transition from Gaussian convolution to diffusion smoothing not only gives an alternative way to implement Gaussian filtering but also supplies a basis for possible applications in other domains.

3.5.1. Optimal smoothing mask within a smallest neighbourhood

Davies [1987] proposed criteria for optimisation in the design of Gaussian operators in small neighbourhoods. They are 1) the isotropy, 2) the overall absolute error of a Gaussian-like mask that approximates the Gaussian kernel and 3) the optimal scale is obtained when the anisotropy or the overall absolute error reaches the minimum. He set up a model for accuracy optimisation, but for isotropy optimisation he mentioned "we have not found it possible to attempt a purely analytic solution to this problem; instead, we made a numerical experiment ...". And he gave two optimal fractional value masks together with their closest integer value masks and the concerned optimal scales as the optimal approximations in accuracy and isotropy as shown in Figures 3.5.1 and 3.5.2

respectively.

$$\frac{1}{8.02}\times\begin{array}{|c|c|c|}\hline 0.45 & 1.00 & 0.45 \\ \hline 1.00 & 2.22 & 1.00 \\ \hline 0.45 & 1.00 & 0.45 \\ \hline\end{array}\qquad \frac{1}{16}\times\begin{array}{|c|c|c|}\hline 1 & 2 & 1 \\ \hline 2 & 4 & 2 \\ \hline 1 & 2 & 1 \\ \hline\end{array}\qquad \sigma_{opt}\approx 0.735$$

Figure 3.5.1. The optimal fractional and nearby integer value masks and scale in overall accuracy approximation.

$$\frac{1}{8.05}\times\begin{array}{|c|c|c|}\hline 0.43 & 1.00 & 0.43 \\ \hline 1.00 & 2.33 & 1.00 \\ \hline 0.43 & 1.00 & 0.43 \\ \hline\end{array}\qquad \frac{1}{56}\times\begin{array}{|c|c|c|}\hline 3 & 7 & 3 \\ \hline 7 & 16 & 7 \\ \hline 3 & 7 & 3 \\ \hline\end{array}\qquad \sigma_{opt}\approx 0.710$$

Figure 3.5.2. The optimal fractional and nearby integer value masks and scale in overall isotropy approximation.

However, his formulation is questionable. From the viewpoint of diffusion smoothing, as the Laplacian ∇^2 itself is an isotropic operator, it is not correct to derive the optimal mask using separated isotropy and accuracy criteria. The discretisation of the diffusion equation must be consistent with the isotropy of the Gaussian kernel. Otherwise, the overall accuracy in (3.3.13) and (3.3.14) will be reduced.

Based on the analysis of the Gaussian-like mask in Theorem 3.5, a conclusion can be easily drawn without appealing to numerical experiment and thus any computing error or empirical assumption can be avoided.

Theorem 3.8. *In the sense of the overall isotropy and accuracy, the optimal* 3×3 *Gaussian-like mask in a one-step iteration with a local truncation error* $O(\tau+h^2)$ *should be*:

$$\frac{1}{16}\times\begin{array}{|c|c|c|}\hline 1 & 2 & 1 \\ \hline 2 & 4 & 2 \\ \hline 1 & 2 & 1 \\ \hline\end{array}\qquad \sigma_{opt}=\frac{1}{\sqrt{2}}\approx 0.707$$

Figure 3.5.3. The optimal 3x3 Gaussian-like mask and scale in overall accuracy/isotropy approximation in one-step iteration.

Proof: Once the 3×3 form has been given for the isotropic mask, all 8-connected neighbouring nodes are involved. Thus the order of local truncation error is virtually determined. Now, only the weighting coefficients in the mask are unknown, but they are restricted by the Gaussian-like form. So the deduction of optimal mask can start from the isotropic square DES mask below (cf. Figure 3.3.4), satisfying the isotropy requirement:

$$\begin{array}{|c|c|c|}\hline \frac{\beta}{4} & \frac{\beta}{2} & \frac{\beta}{4} \\ \hline \frac{\beta}{2} & 1-3\beta & \frac{\beta}{2} \\ \hline \frac{\beta}{4} & \frac{\beta}{2} & \frac{\beta}{4} \\ \hline\end{array}$$

Figure 3.5.4. The isotropic square DES mask.

To be a Gaussian-like form, the above mask must satisfy $\frac{\beta}{4} \cdot (1-3\beta) = \frac{\beta}{2} \cdot \frac{\beta}{2}$, so there is $\frac{b\tau}{h^2} \equiv \beta = \frac{1}{4}$. This gives the mask in Figure 3.5.3.

And for a one-step iteration, $t = \tau = 1$ and $h = 1$, so $b = \beta = \frac{1}{4}$, which leads to the scale $\sigma = \sqrt{2bt} = \frac{1}{\sqrt{2}} \approx 0.707$.

From the construction of the DES schemes, this one-step iteration 3x3 Gaussian-like mask approximates the Gaussian filter with a local truncation error $O(\tau+h^2)$.

From Remark 3.4, the square DES mask is the only isotropic 3x3 mask. So the mask in Figure 3.5.3 is the unique isotropic 3x3 Gaussian-like mask. Any variation from the coefficients will violate either isotropy or accuracy of this Gaussian-like mask. Hence, it must be the optimal mask.

Q.E.D.

Remark 3.7: *When the Gaussian-like form is imposed on an isotropic mask, there is virtually no accuracy optimisation problem*

3.5.2. Smoothing a moving object

When a rigid object drifting with a constant velocity, such as an industrial part passing down a conveying belt, is measured, influence from motion will be contained in the raw data. Motion influence is, in nature, quite different from the influence of random noise. The application of direct smoothing (or sharpening) would *not* be satisfactory due to the motion term. However, image smoothing (or sharpening) can be performed with the method below which takes explicit account of the motion.

By means of the physical analogy to heat conduction-convection [Cai 1988b][Scott 1988], a convection term can be introduced to set up a diffusion-wave equation to smooth this drifting object.

First, observe the 1-D convection case which is subject to a wave equation with an initial value condition

$$\frac{\partial u}{\partial t} = -a\frac{\partial u}{\partial x} \qquad (3.5.1e)$$

$$u(x,0) = f(x) \qquad (3.5.1i)$$

where a is the convection coefficient and for simplicity set $a > 0$.

The solution of equation (3.5.1) in the D'Alembert form is:

$$u(x,t) = f(x-at) \qquad (3.5.2)$$

This general form means that the wave-front will remain invariant along the characteristic $x - at = \text{const}$ shown in Figure 3.5.5:

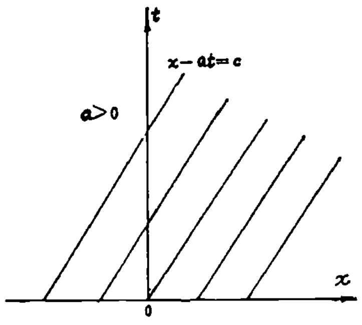


Figure 3.5.5. Characteristic for the 1-D wave equation.

Letting the initial value distribution $f(x)$ be a triangular wave, it will float with a constant speed along the x-direction at each time t = 1, 2, · as shown in Figure 3.5.6.

Then observe the 1-D conduction case which is subject to the diffusion equation

$$\frac{\partial u}{\partial t} = b\frac{\partial^2 u}{\partial x^2} \tag{3.5.3}$$

where $b > 0$ is the diffusion coefficient.

The analytical solution of equation (3.5.3) is

$$u(x,t)=\frac{1}{\sqrt{4\pi bt}}\int_{-\infty}^{\infty} f(\xi) exp[-\frac{(\xi-x)^2}{4bt}]\, d\xi \tag{3.5.4}$$

The initial triangular wave keeps stationary at the original position but its wave-form becomes more and more smoothed when time t increases as shown in Figure 3.5.7.

Now observe the 1-D conduction-convection case which is subject to the diffusion-wave equation

$$\frac{\partial u}{\partial t} = -a\frac{\partial u}{\partial x} + b\frac{\partial^2 u}{\partial x^2} \tag{3.5.5}$$

where $a \geq 0$ is the convection coefficient, and $b > 0$ is the diffusion coefficient.

The analytical solution of equation (3.5.5) is

$$u(x,t)=\frac{1}{\sqrt{4\pi bt}}\int_{-\infty}^{\infty} f(\xi) exp[-\frac{(\xi-x-at)^2}{4bt}]\, d\xi \tag{3.5.6}$$

The wave still floats with the same speed as in the convection case in Figure 3.5.8. However, its shape becomes more and more smoothed, similar to the pure heat conduction process. That is, due to the diffusion the wave-front no longer remains invariant in motion.

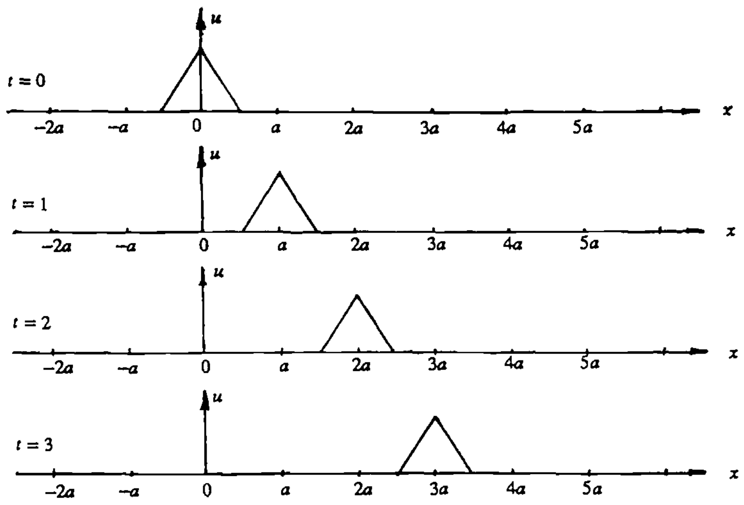


Figure 3.5.6 Behaviour of the triangular wave in 1-D heat convection.

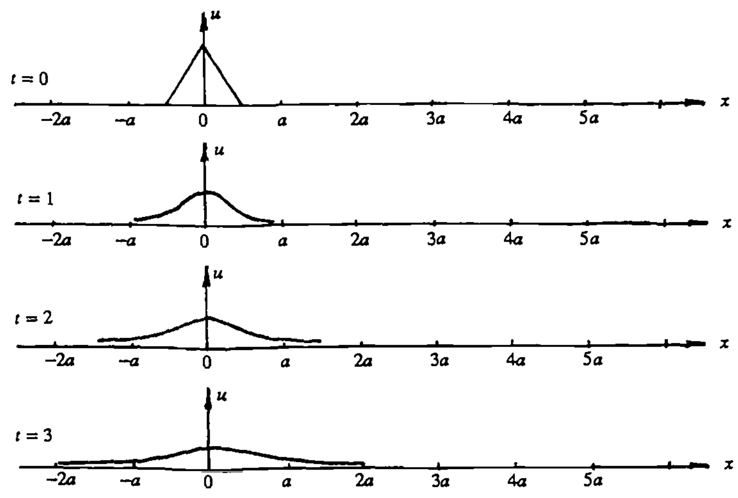


Figure 3.5.7. Behaviour of the triangular wave in 1-D heat conduction.

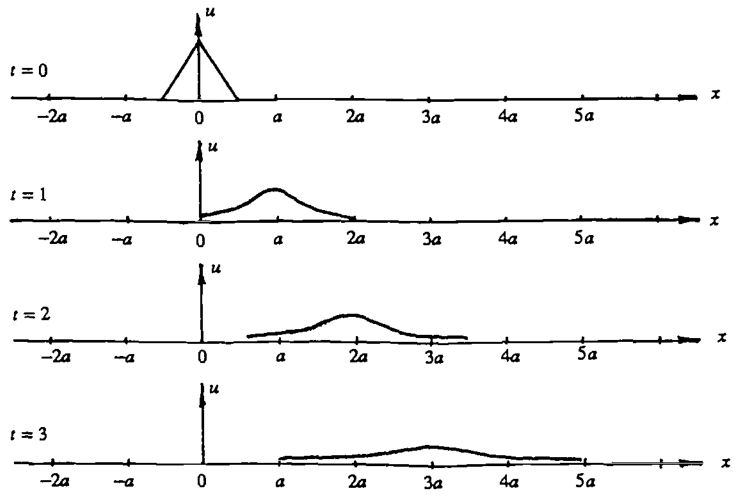


Figure 3.5.8. Behaviour of the triangular wave in 1-D heat conduction-convection.

By comparison, it is seen that the motion influence can be filtered out within the wave-diffusion smoothing process, which leads to the wave being held at the original position. As the motion is of a constant velocity, the convection coefficient a in equation

(3.5.5) is easy to determine by measurement.

For example, an object in horizontal motion can be smoothed using a mixed 2-D diffusion/1-D wave equation:

$$\frac{\partial u}{\partial t} = -a\frac{\partial u}{\partial x} + b\nabla^2 u \qquad (3.5.7)$$

where $a \geq 0$ is the convection coefficient or the constant velocity of the object, and $b > 0$ is the diffusion coefficient.

Using the central difference to discretise the convection term gives an explicit scheme with its mask form below:

$$u_{p,q}^{k+1} = -\frac{\rho}{2}(u_{p+1,q}^{k} - u_{p-1,q}^{k}) + (1 - 4\beta)u_{p,q}^{k} + \beta(u_{p-1,q}^{k} + u_{p+1,q}^{k} + u_{p,q-1}^{k} + u_{p,q+1}^{k}) \qquad (3.5.8e)$$

$$u_{p,q}^{0} = f_{p,q} \qquad p,q = 0\ (1)\ M-1, \quad k = 0\ (1)\ \infty \qquad (3.5.8i)$$

	β	
$\beta+\frac{\rho}{2}$	$1\text{-}4\beta$	$\beta-\frac{\rho}{2}$
	β	

Figure 3 5.9. The normal cross DES mask of the convection-conduction process.

where

$$0 \leq \rho = \frac{a\tau}{h} \quad \text{and} \quad 0 < \beta = \frac{b\tau}{h^2} \qquad (3.5.9)$$

The stability analysis of scheme (3.5.8) gives an error increase factor:

$$\lambda = -i\rho\sin\theta_1 + (1-4\beta) + 4\beta\cos\theta_1\cos\theta_2 \qquad (3.5.10)$$

To ensure $|\lambda| \leq 1$, the stability limit is:

$$\rho^2 \leq 4\beta \leq 1 \qquad (3.5.11)$$

or

$$\tau \leq min(\frac{h^2}{4b}, \frac{4b}{a^2}) \qquad (3.5.12)$$

An implicit scheme can be set up by introducing a convection term only in the x-direction in DISCT:

$$\frac{1}{\tau}(u_{p,j}^{k+\frac{1}{2}} - u_{p,j}^{k}) + \frac{a}{h}(u_{p+1,j}^{k+\frac{1}{2}} - u_{p-1,j}^{k+\frac{1}{2}}) - \frac{b}{h^2}(u_{p-1,j}^{k+\frac{1}{2}} - 2u_{p,j}^{k+\frac{1}{2}} + u_{p+1,j}^{k+\frac{1}{2}}) = 0 \qquad (3.5.13x)$$

$$\frac{1}{\tau}(u_{p,j}^{k+1} - u_{p,j}^{k+\frac{1}{2}}) - \frac{b}{h^2}(u_{p,j-1}^{k+1} - 2u_{p,j}^{k+1} - u_{p,j+1}^{k+1}) = 0 \qquad (3.5.13y)$$

$$u_{p,j}^{0} = f_{p,j} \qquad p,j = 0\ (1)\ M-1, \quad k = 0\ (1)\ \infty \qquad (3.5.13i)$$

The stability analysis shows a similar result to the DISCT scheme:

$$\lambda = \lambda_1\lambda_2 = (1+2\beta-2\beta\cos\theta_1+i\rho\sin\theta_1)^{-1}(1+4\beta\sin^2\frac{\theta_2}{2})^{-1} \qquad (3.5.14)$$

So, the error increase factor $|\lambda| \leq 1$ will always be true without any restriction to β and ρ. The scheme is therefore unconditionally stable.

Note that the vector form of (3.5.13x) is

$$\begin{bmatrix} -\rho-\beta & 1+2\beta & \rho-\beta \end{bmatrix} \begin{bmatrix} u_{p-1\,j}^{k+\frac{1}{2}} \\ u_{p\,j}^{k+\frac{1}{2}} \\ u_{p+1\,j}^{k+\frac{1}{2}} \end{bmatrix} = u_{p\,j}^{k} \tag{3.5.15}$$

The tri-diagonal matrix of this scheme is still diagonal-dominant. Hence all the advantages from the DISCT scheme are still available.

3.6. Summary

In this chapter, the diffusion smoothing theory is developed. The main results are:

1) The relationships between Diffusion Smoothing (DS), Gaussian Convolution (GC), Repeated Averaging (RA) and Spline Smoothing (SS) have been made explicit: (a) GC is equivalent to the initial value problem of DS, (b) RA is a special case of the above explicit DS; (c) SS is a special case of the explicit DS with a "convex corner cling" boundary condition. Hence, DS provides a uniform theoretical base for interpretation of GC, RA and SS

2) DS provides a convenient and more reasonable way to treat the boundary condition. A "small leakage" diffusion model is proposed which efficiently preserves surface shape at the boundary. With this improvement, it becomes practicable to get good quality curvature sign images in scale space.

3) Both the diffusion explicit smoothing (DES) and diffusion implicit smoothing (DIS) schemes are scale incremental methods, they produce denser intermediate results without extra computation. Especially, DIS is an unconditionally stable scheme, thus errors will be automatically reduced in computation, which leads to a lower computational complexity than GC, RS, SS and DES, thus makes scale space processing feasible.

4) DS can also be applied to intensity data processing, drifting object smoothing and symmetry axis eliciting, *etc.*

In this way, DS theory provides a feasible, convenient and efficient tool for scale space filtering. Within this framework the surface segmentation and description will be implemented at each scale level, leading to a scale space representation suitable for object recognition.

As an overview, the relationships amongst DS, DES, DIS, GS, GC, RS and SS are illustrated in figure 3.6.1:

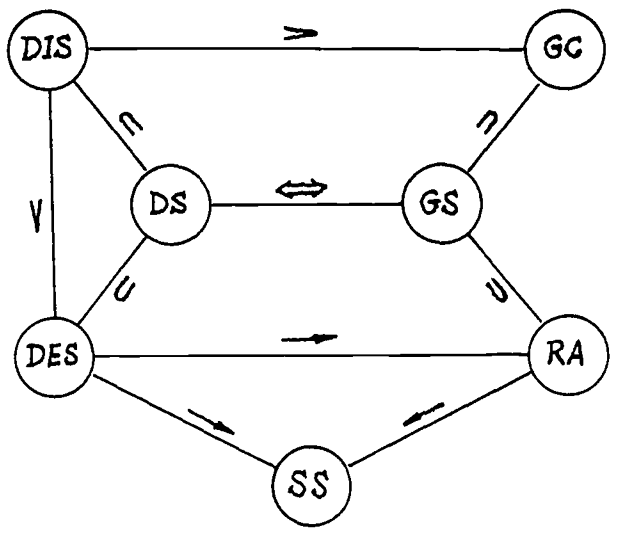


Figure 3.6.1. Relationships between DS, DES, DIS, GS, GC, RA and SS.

where, the meanings of symbols are:

"$x \Longleftrightarrow y$"	denotes that	"x is equivalent to y",
"$x \leftarrow y$" or "$y \rightarrow x$"	denotes that	"x is a special case of y",
"$x > y$" or "$y < x$"	denotes that	"x is better than y", and
"$x \subset y$" or "$y \supset x$"	denotes that	"x belongs to y".

Chapter 4

Scale space behaviour of surface curvature features

In Chapter 3, diffusion smoothing theory supplied a framework for subsequent scale space processing. Now scale space behaviour of surface features will be discussed in terms of surface curvatures within this framework.

4.1. Introduction

4.1.1. Scale effects

Surface understanding depends on the spatial scales at which the surface is measured and analysed. The problem of scale effects in sampling or filtering is thus raised. When considering the analysis of a *given* image, the discussion of the sampling scale effects may be omitted, and attention will concentrate on the structure of the given image. Thus the term scale effects used in this study actually means the smoothing scale effects alone.

Scale effects may be observed in three aspects. First, significant features will appear and disappear at different scale levels in human perception For example, a pencil looks like a prism in front of your face and appears a circular cylinder when viewed from beyond four meters Hence, the key features vary at different scale levels, as does the importance of any individual surface feature

Second, the raw image data is usually corrupted by noise, which has to be suppressed. However, without some prior knowledge about the small features, it is difficult to distinguish the behaviour of a small feature from that of noise; thus small features may be filtered along with noise as the scale increases.

Third, filtering up to a large scale can remove most noise, but will gradually introduce surface distortion as well When the surface becomes smoother or even flat, the signal to noise ratio of the surface depth first rises over a certain range of scales due to the noise reduction, but eventually decreases due to the distortion of the surface. The variation of the signal to noise ratio in multiple scale filtering affects the results of visual processing, which vary from scale to scale. Therefore a visual processing method which succeeds at one scale is not necessarily or not wholly applicable at multiple scale levels.

However, it must be noted that the variation of the results over the scale range is not random or totally unpredicatable In the 1-D signal case, Yuille [1983] described the relevant rule as the fingerprint theorem, Witkin [1983] described the variation with a tree-like structure and Koenderink [1984] discussed the structure of an image as a one-parameter family using the diffusion equation. All three lead to the theorem of "uniqueness of the Gaussian kernel for scale space filtering" as in [Babaud *et al* 1986]. The key ideas in [Witkin 1983] are to retain the useful information from tracking the movement of significant features in scale space and establishing a stable description over a large range of scales. For the 3-D surface case, the situation is more complicated, but it is still possible to formulate some qualitative rules, such as the curvature variation rules or the "Shoe flattening" rules as they are termed in Chapter 6 of this thesis, by analogy with the reverse of the process in which initially flat leather is turned into the curved form of a shoe.

4.1.2. Organisation of the chapter

The problem of scale effects has been raised in the earlier part of this introduction to the issue of scale space behaviour. Discussion will be developed in the subsequent six sections. Section 4.2 gives the classification of surface shape primitives and its colour denotation for display. Section 4.3 will discuss the stability of these surface shape primitives, or the behaviour of large features which are in a weak state. Section 4.4 will discuss scale space behaviour of small surface features on KH sign images in terms of a small feature on a large background patch which is in a strong state (both "weak" and "strong" states are defined in Section 4 3); and also discuss the structure of their interacting behaviour. Section 4.5 applies this behaviour theory to explain an odd phenomenon occurring in surface detection with the curvature sign images. Section 4.6 is the summary of this chapter

4.2. The classification of surface shape primitives

4.2.1. Surface shape primitives based on the curvature signs

"Human visual system is remarkably sensitive to curvature" [Watt and Andrews 1982]. There are two main reasons for qualitative information (*i e*, signs) rather than quantitative information (*i e*, magnitudes) about the curvatures being used to extract the shape features from surface curvatures. First, "people are mostly capable of making qualitative judgements of surface shape" [Stevens 1981], since human perception is sensitive to surface convexity/concavity and their alternation, but insensitive to continuous or small variations of curvature magnitudes along a simple convex (or concave) curve. Second, in the presence of noise the sign of the estimated curvature is usually robuster than the magnitude since the sign is invariant on both sides of the critical point zero whereas the magnitude varies everywhere.

In Chapter 2, it was shown that at each point of a smooth surface there is a pair of principal curvatures C_1 and C_2 associated with its principal directions ϕ_1 and ϕ_2 respectively. And using the signs of the Gaussian curvature $K = C_1 C_2$ and the mean curvature $H = \frac{1}{2}(C_1+C_2)$ surface shapes can be described in eight types (see Figure 2.6.3). Note that the signs of K and H convolve both sign and magnitude information of principal curvatures C_1 and C_2. The (K,H) sign category is adopted to classify surfaces and its eight elements will be chosen as shape primitives in this research. From the uniqueness theorems of Gaussian curvature and mean curvature, this categorisation promises to analyse and capture surface features in a straightforward and simple style.

4.2.2. A colour denotation of surface primitives

For the graphic display of a surface using the above (K,H) sign category, a colour denotation is introduced as shown in Figure 4.2.1, and its comparison with Figure 2.6.3 is shown in Figure 4.4.1. This denotation gives each surface type a unique colour and each surface a unique "coloured" KH sign image (at a specific scale level).

K \ H	–	0	+
–	saddle ridge	minimal	saddle valley
0	ridge	flat	valley
+	peak	(none)	pit

K \ H	–	0	+
–	RED	INDIGO	VIOLET
0	ORANGE	WHITE	YELLOW
+	BLUE	(BLACK)	GREEN

Figure 4 2 1 Colour denotation for surface shapes from the signs of K and H

The "colour" will be frequently used in description since it is more comprehensible than a set of numbers when showing the region's homogeneity, intersections, convexity or concavity *etc.* It is especially suitable in those cases where only the sign of the result is significant for the task. For example, based on the correspondence between colours and (K,H) sign pairs, each pixel can be described with a set of attributes: {(x,y,z), shape (colour), flatness, .. } When a group of pixels of the same colour are merged into one patch, it must be a uniform region of the same shape and this can be verified just at a glance without reading any numbers.

For convenience in discussion, a surface shape (colour) whose K or H sign is zero will be called an *interim shape (colour)*, such as, ridge, valley, minimal and flat (and accordingly ORANGE, YELLOW, INDIGO and WHITE) And a pair of surface shapes (colours) will be referred to as the *conjugate shape* of each other when they have the same K sign but inverse H sign or the same H sign but inverse K sign. They are combined as a *conjugate shape pair*, such as, <saddle ridge, peak>, <saddle ridge, saddle valley>, <peak, pit>, <pit, saddle valley> and <ridge, valley> Further, if a KH sign image is obtained from another one by replacing a shape (colour) with its conjugate shape (colour), this image can be called a *conjugate image* to the original one with respect to the conjugate shape pair <· , ·>, and this relationship is called the *conjugacy* between them. This also applies to two series of KH sign images in a one-to-one conjugacy correspondence.

4.3. The stability of the (K,H) sign category in the presence of noise

The (K,H) sign category is to be used to classify surface shapes, but the question arises: What is the significance of the (K,H) sign categorisation? Is it merely a mathematically neat classification of surfaces or can it also provide a clue to explain some puzzling phenomena in visual processing? To answer these questions the stability of (K,H) sign types in the *weak* state is discussed The term "weak state" here means a state where $K \approx 0$ or $H \approx 0$ thus the sign of K or H tends to or is easy to change under a small perturbation from noise. The term "strong state" appearing later means the opposite case, where K and H are far away from zero As surface patches of these types are large features on the surface, this is also a discussion on the scale space behaviour of large shape features under small perturbation.

4.3.1. The perception of hyperbolic surface types in noisy data

Surfaces of negative Gaussian curvature K are hyperbolic surfaces (saddles). Unlike the (C_1,C_2) sign category which gives only two saddle types, the (K,H) sign categorisation provides an extra type to describe saddle surfaces. The minimal surface (for which H= 0) is further separated out as an independent type (cf. Section 2.6).

EX BIBL. UNIV. EDINBURGEN.

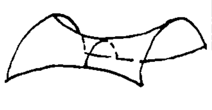
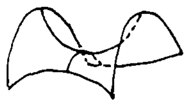
H	−	o	+
$K<0$			

Figure 4.3.1. Hyperbolic surfaces ($K<0$) saddles categorised by the H sign (+,o,−)

Unfortunately, it is difficult for human perception to tell a minimal surface from the other saddle surfaces, especially, when their values of H are near zero. Moreover, this small difference may vanish in scale space filtering due to scale effects. Hence, the minimal surface type has only a limited importance in visual perception of the real world though it does have a great significance in theoretical classification and physical research.

In computer analysis of surface data, the detection of a minimal surface point depends on the 2nd order numerical derivatives; hence it is rather sensitive to the noise, similarly for the minimal surface region (as a uniform set of the minimal surface points). Therefore, it is rare to detect a minimal region in raw data.

Filtering to reduce noise will *not* directly improve this situation, as surface distortion always accompanies noise suppression during scale space filtering. It is possible that there is a specific though unknown scale at which more minimal points appear than at the zero scale (*i.e.*, raw data case). However, at most scale levels, more minimal points disappear instead of remaining due to noise perturbation or surface distortion. All these factors suggest that the minimal surface type is unstable to small perturbations in H, and that it could be preferable to ignore it in categorising and merge it into the other two related types, *i.e.*, to be the (C_1,C_2) category, or even merge all three saddle types into one as in the (C_1,C_2) *sign* category as some researchers did (*e.g.*, [Fisher 1989]), where the curvature magnitude information of C_1 and C_2 is completely ignored.

However, it is preferable to keep the minimal surface type in the category. The aim of doing so is not to detect this type directly but to use it as an interim state to keep apart the other states $H>0$ and $H<0$ (saddle ridge and saddle valley), because the difference between them does have significance in some cases, such as surface segmentation of a Renault automobile part as will be shown later (cf. Figure 5.3.5, t = 9 and 16).

To realise this idea, it is necessary to stabilise the minimal type as a state. A feasible method is to give the mean curvature $H=0$ an appropriate zero threshold ε_H as half the width of a tolerance band. If a hyperbolic region's mean curvature H falls into this band, this region is seen as a minimal surface. So this band must be small enough for states on both sides that the minimal surface is just an interim state between them but it must be broad enough to inhibit crossing in the presence of noise. Once such a tolerance band has been set up, the $H=0$ state will be robuster to both noise and distortion than before, as will both the $H>0$ and $H<0$ states.

4.3.2. The perception of ellipsoidal surface types in noisy data

When $K \equiv C_1 C_2 > 0$, surfaces are ellipsoidal surfaces. As the case $H \equiv \frac{1}{2}(C_1+C_2)=0$ is impossible in theory, there exist only two ellipsoidal surface types, peak and pit, standing separately as shown in Figure 4.3.2.

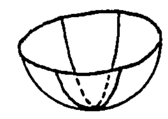
H	–	o	+
$K>0$			

Figure 4.3 2. Ellipsoidal surfaces ($K>0$) peak and pit surfaces categorised by the H sign (+,o,–).

When a small perturbation occurs in H, they cannot directly change into each other. So, they are stable types in the categorisation, especially, when they are in the strong states, *i e* , H and K are far away from the zero. Thus, in scale space, a large (BLUE) peak region or a (GREEN) pit region will usually persist on the KH sign image over a broad scale range.

What is of interest here is not their behaviour in the strong state but in the weak state, *i e* , when $H \approx 0$ thus $K \approx 0$ (due to $K > 0$).

It seems easy for humans to see the difference between a peak and a pit even though they are in a weak state, so they are still perceptually stable However, for a computer, things are not so simple. A "pure" zero is rarely met. A small number will be seen as zero in the computation if it is less than a given zero threshold (cf. subsection 4.3.1). However, these small numbers may occur frequently, so the zero thresholds of ε_K and ε_H cannot be fixed casually, or disastrous unforeseen results may occur. When both principal curvatures are small, *e g* , $C_1 = C_2 = 9\times10^{-4}$, as K is their product, it is much smaller than H as their average. Hence, it seems reasonable to set a much smaller ε_K than ε_H, such as, $\varepsilon_K = 10^{-7}$ and $\varepsilon_H = 10^{-3}$. However, this casual choice leads to $H = 9\times10^{-4} < \varepsilon_H$ and $K = 8.1\times10^{-7} > \varepsilon_K$, and the "phantom case" of $H = 0$ and $K > 0$ occurs!

Note that $K > 0$ means C_1 and C_2 have the *same* sign. The following result can be derived from the inequality between the arithmetic average and the geometric average:

$$K = C_1\,C_2 = |C_1|\,|C_2| \leq \left[\frac{|C_1|+|C_2|}{2}\right]^2 = \left[\frac{C_1+C_2}{2}\right]^2 = H^2 \tag{4.3.1}$$

thus, ε_H and ε_K must satisfy the following inequality to guard against the small perturbation:

$$\varepsilon_K \geq \varepsilon_H^2 \tag{4.3.2}$$

If so, computational practice will be consistent with the classification theory. There will be no "phantom type" as the intermediate type between the peak and the pit, hence the contrast between them is fairly stable since both types still cannot exchange when a small perturbation occurs.

Remark 4.1: *Unfortunately, a peak or a pit patch created by significant noise instead of an object in the raw data might enjoy the same stability too, thus it might not be eliminated easily in some cases during the smoothing as will be shown in later sections.*

4.3.3. The perception of parabolic surface types in noisy data

Surfaces of the Gaussian curvature $K = 0$ are parabolic surfaces. Similarly to the minimal type, the planar or "flat" type is an interim between the ridge and the valley, so it is discussed first.

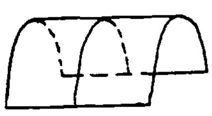
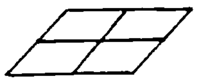
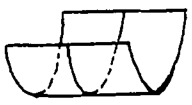
H	–	o	+
K=0			

Figure 4 3 3 Parabolic surfaces (K=0) ridge, plane and valley categorised by the H sign (+,o,–).

A point or a patch is the flat type when both $K = 0$ and $H = 0$, so it might be "doubly" sensitive to noise and difficult to detect in raw data using the curvature signs in terms of the KH sign image

Interestingly, the situation of the flat type seems perceptually better than the minimal type in scale space. Human perception seems to work with several scale levels at the same time It tends to even out small perturbations of the shape to handle the global tendency of a surface. So, on the one hand, perception is acute to the local variation around the smooth joint of a straight line and a circle as well as of a plane and a cylinder; on the other hand, perception seems more tolerant to small perturbations of a globally planar surface. This capability of observing an object at different scales causes the flat type to be slightly more perceptually stable than the minimal type.

Again, the flat type as an interim state can be consolidated by imposing a zero threshold ε_H as a half band-width on it. Notice that where both K=0 and H=0, the value of ε_H should be greater than that in the minimal type case where only H=0. However, it cannot go too far as the threshold ε_H will be applied uniformly to both flat and minimal surfaces. So the band must not be too broad, otherwise it will be against the aim that the minimal type is just an interim state used to keep apart the saddle ridge and saddle valley and not a state being directly detected. Once a proper zero threshold ε_H has been set up, this interim state is consolidated into an intermediate type.

The discussion so far has mainly focused on the stability against a perturbation around the H=0 cases. Attention is now turned to the cases where K is near zero.

Obviously, a ridge type, as the intermediate type between the peak (K>0) and the saddle ridge (K<0), always has zero Gaussian curvature. Consequently, in noisy circumstances, it is not easy to detect a ridge patch in raw data using the KH sign image. A threshold ε_K is thus given as the half band-width to consolidate the state of Gaussian curvature K=0. If a patch's Gaussian curvature K falls into this band, the patch will be seen as a ridge surface. Note that a ridge patch can be seen as a valley patch put in the inverse direction. The situation is similar for the valley type between the pit type and the valley saddle type and the same threshold will be applied to it also.

From the Gaussian curvature uniqueness theorem, variations in Gaussian curvature K are more significant than that in the mean curvature H. Surfaces of different Gaussian curvatures cannot be isometric, *e.g.*, a sphere cannot be mapped to a plane through an isometric transformation; whereas surfaces of different mean curvatures may be isometric, *e.g.*, a cone or a cylinder can be developed into a plane. Thus, setting a band ε_K here has a different purpose from before: it aims to preserve the ridge/valley region itself, such as a cylinder, not just to separate the peak/pit and the saddle types.

To achieve this aim, the current value of ε_K should be larger than the value of ε_K chosen in subsection 4.3.2. That is, ε_K here should be much larger than ε_H^2. 4.3.2. Such a thresholding enables the ridge (or valley) state to be robuster to both noise perturbation and surface distortion, and improves the detection of ridges (or valleys).

It is unclear at the moment that whether the above approach is proper or not, of which an estimation will be made soon. But it seems that this approach has support in human visual perception, since a ridge patch, once it has existed, is a more perceptually stable state than the minimal patch. Generally, it is easier to see the difference between a cylinder and an ellipsoid (or a saddle) than that between a minimal surface and a saddle, and it is much easier to distinguish a saddle from a peak. Hence, compared with ε_H^2 it is reasonable to give a broader zero threshold ε_K to detect or preserve a ridge patch against small perturbations A large value for ε_K would also be appropriate for the flat type since its only possible interchange, that between a flat patch and a minimal patch, seems rather unlikely.

Once both the zero thresholds ε_H and ε_K have been properly set up, the whole (K,H) sign category is consolidated. This also gives four contrast pairs relevant to the types at four corners on the table of the (K,H) sign category in Figure 2.6 3. They are <peak, pit>, <peak, saddle(s)>, <pit, saddle(s)> and <saddle ridge, saddle valley>. The strength of contrast between two types in each pair is different and depends on the zero thresholds ε_H and ε_K. Roughly speaking, they can be listed in descending order of the strength of contrast:

<peak, pit> → <peak/pit, saddles> → <saddle ridge, saddle valley>

Further, eight shape primitives themselves have different stability in the "normal" cases. They can be grouped into several sets in a descending order of the stability:

{peak, pit} → {saddle ridge, saddle valley} → {ridge, valley} → {flat} → {minimal}

This will be helpful later to explain some puzzling phenomena occurring in region detection using the KH sign image

4.3.4. The relationship between zero thresholds ε_K and ε_H

So far, to consolidate the stability of the weak (K,H) types against perturbation, the key point is to give *proper* zero thresholds around $K = 0$ and $H = 0$ respectively, and these zero thresholds must be compatible so as to prevent the "phantom case" from occurring. Hence, it is natural to think about: *Is there any inter-relationship between ε_K and ε_H themselves?* This is an open problem so far

Surely, there has been one expression as in (4.3.2), but it is too simple to be seen as the *inter*-relationship. It considers only the planar surface case and ignores any effect of scales. Recall that [Besl and Jain 1986] and [Besl 1986] once used the following zero thresholding in their experiment:

$$\varepsilon_k = 0.015\ , \qquad \varepsilon_H = 0.03 \qquad (4.3.3a)$$

and

$$\varepsilon_K = 0.015\ , \qquad \varepsilon_H = 0.06 \tag{4.3.3b}$$

The inequality (4.3.2) actually cannot give a satisfactory explanation for these thresholdings, where not only $\varepsilon_K >> \varepsilon_H^2$ but also $\varepsilon_K = O(\varepsilon_H)$. Therefore, the zero thresholding problem looks too flexible to handle.

Fortunately, a more certain and appropriate relationship can be found between the zero thresholds ε_H and ε_K for all eight surface types. Note that K is the product of C_1, C_2 and H is the mean of C_1, C_2. It can be obtained by simply introducing a small noise perturbation to principal curvatures and observing the behaviour of the Gaussian and mean curvatures in this circumstance.

Suppose that the principal curvatures have a small perturbation ξ_1 in C_1 and ξ_2 in C_2. They introduce perturbations E_H in H and E_K in K respectively:

$$E_H = \frac{(C_1+\xi_1) + (C_2+\xi_2)}{2} - \frac{C_1+C_2}{2} = \frac{\xi_1 + \xi_2}{2} \tag{4.3.4}$$

$$E_K = (C_1+\xi_1)(C_2+\xi_2) - C_1C_2 = (C_1\xi_2 + C_2\xi_1) + \xi_1\xi_2 \tag{4.3.5}$$

Note that the noise perturbation is from the same source and the magnitude of the perturbation is the main concern here. A further simplification can be made.

Let $\xi_1 = \xi_2 = \xi$, therefore

$$E_H = \xi \tag{4.3.6}$$

$$E_K = (C_1+C_2)\xi + \xi^2 \tag{4.3.7}$$

Now it can be seen that E_K and E_H are explicitly related via H as below[1]:

$$E_K = 2H\,E_H + E_H^2 \tag{4.3.8}$$

Set $|E_H| = \varepsilon_H$, there is

$$|E_K| \leq 2|H|\cdot\varepsilon_H + \varepsilon_H^2 \tag{4.3.9}$$

Hence, to guard against the perturbation the zero threshold of the Gaussian curvature must satisfy:

$$\varepsilon_K \geq \ \sup\, |E_K|$$

Otherwise, for example, an exact $K = 0$ plus the perturbation E_K will be beyond the zero band $[-\varepsilon_K\,, \varepsilon_K]$ thus leading to incorrect classification due to the incompatible choice of zero thresholds.

Therefore, the following inequality is obtained for consistent zero thresholding of the Gaussian and mean curvatures:

Inequality 4.1 (*Consistent zero thresholding*):

$$\varepsilon_K \geq 2|H|\,\varepsilon_H + \varepsilon_H^2 \tag{4.3.10}$$

---

[1] This implies that whatever K is, the perturbation introduced in K might have a larger amplitude than that in H when $H > \frac{1}{2}$, and a much larger amplitude when $H > 5$.

Comparing with (4.3.2), the formula (4.3.10) has a term containing the mean curvature H. This term shows evidence that the zero thresholds of K and H are not irrelevant but related via mean curvature H, which brings the scale effects into zero thresholdings over different scale levels. When $H \equiv 0$, (4.3.10) degenerates to $\varepsilon_K \geq \varepsilon_H^2$ just as in (4.3.2). And when $|H| = \varepsilon_H$, (4.3.10) becomes

$$\varepsilon_K \geq 3\varepsilon_H^2 \tag{4.3.11}$$

So (4.3.10) provides a more general formula than (4.3.2).

Setting up the value of ε_K for the whole surface S gives

$$\varepsilon_K = \varepsilon_H^2 + 2 \max_S |H| \, \varepsilon_H \tag{4.3.12}$$

It is now no surprise to get the following result:

$$\varepsilon_K = O(\varepsilon_H) \qquad \text{when} \quad \max_S |H| > 0.1 \tag{4.3.13}$$

$$\varepsilon_K > \varepsilon_H \qquad \text{when} \quad \max_S |H| > 0.5 \tag{4.3.14}$$

$$\varepsilon_K > > \varepsilon_H \qquad \text{when} \quad \max_S |H| > 5.0 \tag{4.3.15}$$

This can be used to interpret why the zero threshold imposed on the Gaussian curvature is usually close to and sometimes even larger than the zero threshold of mean curvature. It also gives a better interpretation to zero thresholds (4.3.3) used by [Besl 1986] and [Besl and Jain, 1986].

In scale space, when the smoothing scale increases, the surface becomes more and more smooth. Note that the high curvature points on a smooth surface are always few even though some false high curvature points produced by noise are included. Also note that the aim of setting ε_K and ε_H is to treat those zero curvature points properly. Therefore, using the average value of H over the whole surface may be feasible in the scale space processing. The following formula is thus suggested:

$$\varepsilon_K = \varepsilon_H^2 + 2 \underset{S}{Average} |H| \, \varepsilon_H \tag{4.3.13}$$

Remark 4.2: *Once a proper zero threshold ε_H has been chosen, the zero threshold ε_K will be given automatically by the above formula, which is particularly convenient for scale space processing, where any zero threshold empirically imposed on K and H at a certain scale may not be appropriate at other scales*

The above discussion on the stability of (K,H) sign shape types has shown the behaviour of large surface features around the $K = 0$ and/or $H = 0$ on the (K,H) sign image in the presence of a small perturbation (*e.g* , noise) Proper zero thresholds are required to stabilise some types. For this purpose, the relationship between the zero thresholds ε_K and ε_H has been formulated as the consistent zero thresholding inequality, which makes explicit that ε_K and ε_H are not irrelevant but related via the mean curvature H, which brings the scale effects into zero thresholdings over different scale levels. The next section will discuss the behaviour of small shape features or noise on the (K,H) sign image assuming each surface patch, which hosts the small features, to be in a strong state.

4.4. The scale space behaviour of small features in the KH sign image

A very smooth surface may have a perfect KH sign image which is separated into several large uniformly coloured patches. Once the surface is corrupted by noise, the KH sign image may become covered in spots. By applying scale space filtering to the noisy surface, the noise is reduced, but surface distortion is also introduced. The interaction between these scale effects could lead to a more complicated situation than might be naively foreseen.

When the small surface features are mixed with the large features in the KH sign image, the situation is the same. Even though many small features are larger than noisy spots, compared with the large features or the global surface shape, they might be seen as noise in effect. In this sense, the investigation of the behaviour of a small feature can be replaced by that of noise. This gives a convenient model for the shapes of small features.

What will be highlighted is that the scale space behaviour of a small feature in the KH sign image depends not only on its own shape but also on the global shape of the larger background surface patch, in which it is embedded.

For simplicity, assume the small surface feature to have a convex or concave spherical shape, hereafter termed a "convex" or "concave" (shape) feature; and that the background surface patches are termed "background" (patches) which will be classified into eight KH sign types: flat, peak, pit, saddle ridge, saddle valley, minimal, ridge and valley respectively. The table below gives the correspondences between the KH sign categories, their colour denotations on the KH sign image and the grey levels in the monochrome images shown in this section.

$K>0,H<0$	$K=0,H<0$	$K=0,H>0$	$K<0,H<0$	$K<0,H>0$	$K<0,H=0$	$K>0,H>0$	$K=0,H=0$
peak	ridge	valley	saddle ridge	saddle valley	minimal	pit	flat
BLUE	ORANGE	YELLOW	RED	VIOLET	INDIGO	GREEN	WHITE
▓	▓	▓	▓	▓	▒	░	□

Figure 4.4.1. Correspondences between KH sign categories and their colours (grey levels) in images.

4.4.1. The scale space behaviour of small features on a flat background

A. Convex features

The first observed case is that of a convex feature on a flat background. The evolution of the surface in scale space filtering will be illustrated in Figure 4.4.2, where hereafter the scale increases from top to bottom and from left to right; Gaussian scale σ and diffusion time t hold a relation as $\sigma = t^2$ and $\sigma = 1, 2, 3, \cdots$.

At the beginning, the spherical (shape) feature on the KH sign image is a BLUE spot surrounded by a VIOLET ring due to the local saddle valley shape around the intersection of the sphere and the (WHITE) flat background. (The whole intersection is actually a shape like the lower part of a torus.)

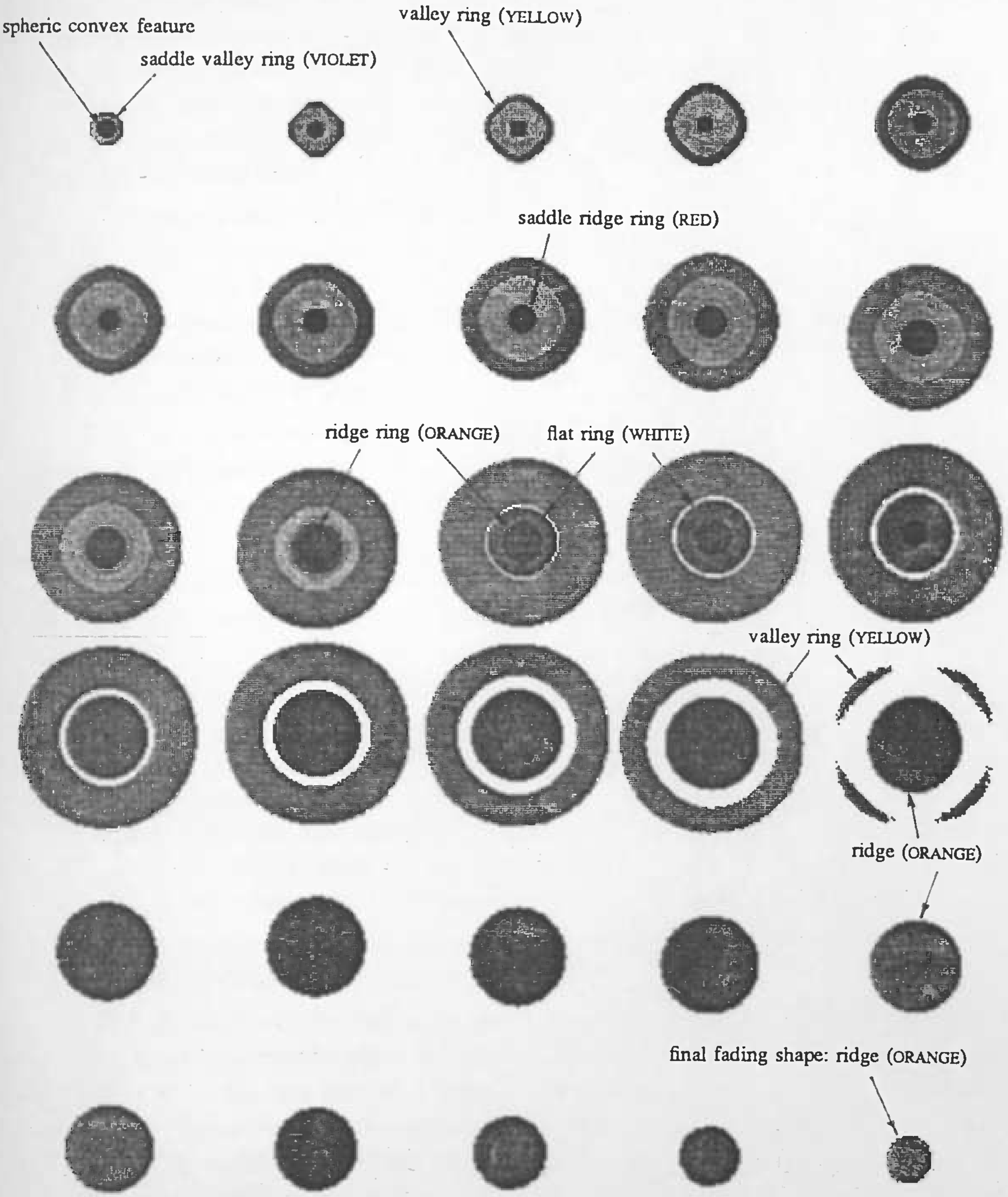


Figure 4.4.2. Scale space behaviour of a convex feature on a flat background patch ($\sigma = 1, 2, ..., 30$).

During scale space filtering, most of the area of the original flat background remains flat (WHITE) and the central area of the noisy sphere stays convex (BLUE) until it becomes nearly flat. However, when the VIOLET ring part gradually expands in size and changes in shape (colour), a YELLOW ring of the valley shape appears outside the VIOLET ring of the saddle valley shape, and a RED ring of the saddle ridge shape appears inside.

Later, while still being a convex patch, the peak patch's flatness quantity (cf. formula (2.8.10) in Chapter 2) is more and more close to the neighbouring (RED) saddle ridge ring's, the BLUE spot of the peak shape begins to expand. An ORANGE ring of the ridge shape appears as the interim between the BLUE spot of the peak shape and the RED ring of the saddle ridge shape. It further expands and covers both rings of the (RED) saddle ridge shape and (VIOLET) saddle valley shape up to the border of the YELLOW ring of the expanding valley shape.

After this, a WHITE ring of the flat shape appears as the interim zone between the YELLOW ring of the valley shape and the ORANGE ring of the ridge shape when the BLUE spot of peak shape begins to vanish under pressure from the ORANGE ring of the ridge shape. As a result, on the (WHITE) flat background, there is a large ORANGE spot of the ridge shape surrounded by a WHITE ring of the flat shape that is further surrounded by the YELLOW ring of the valley shape outside.

Finally, first the YELLOW ring of the valley shape disappears then the ORANGE spot vanishes under pressure from the WHITE ring of the flat shape. The whole image becomes a single WHITE colour — when all the noise has been filtered out from the planar background.

The change of colours (shapes) on the KH sign image shows the scale space behaviour of a small convex sphere feature under interference from large planar background patch.

In general, *all* 2½-D surfaces will tend towards flatness in scale space filtering, thus the colour KH sign image will eventually be washed WHITE! As observed above, this evolution may take a long period to complete. And events happened during the process raise some interesting questions:

- Why a small spot can have such a large range of effects in the scale space smoothing?
- Which interim shape (colour) will appear, where and when?
- What is the *final fading shape* (colour) before the KH sign image is washed out?

These questions are closely related to the scale space behaviour referred to as the *remaining, creation and fading* of shapes in this thesis.

For simplicity in the following discussion, suppose each iteration of Gaussian smoothing is implemented using a 3×3 square mask, the smallest size square mask. Also, a specific term, the "top part of a bulge", is introduced here. The word "bulge" here denotes a convex protrusion of a surface from a large flat surround and the "top part" is a neighbourhood of the peak of the bulge and the neighbourhood radius on the support plane is r. Although the top part is a relatively small part of the whole bulge, it has the same convexity with the bulge. And as long as the top part remains, the bulge is "alive" regardless of changes at the boundary part. So a bulge can be represented by its top part in contexts of the surface shape.

Now three theorems are given as the answers to the above questions.

Theorem 4.1. *(Convexity remaining:) The convexity at the top part of a convex positive function f will not be changed by a single Gaussian convolution.*

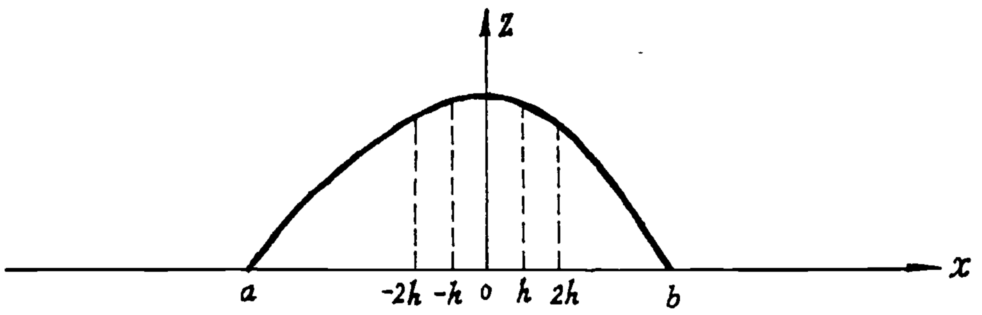


Figure 4 4.3 A convex function $f(x) > 0$ defined over the interval $[a, b]$ with its top part over $[-r, r]$.

Proof: By the symmetry of the Gaussian function, it is only necessary to prove the theorem in the 1-D case.

Note that a convolution is invariant to shifting. Without lose of generality, suppose the positive convex function $f(x)$ is over an interval $[a, b]$ with its peak at the origin $x = 0$ and the top part over an interval $[-r, r]$, where $a < 0 < b$ and $0 < 2r < \min[|a|, |b|]$.

For any t and r satisfying

$$-r \le t \le r, \qquad 0 < 2r < min[|a|, |b|] \tag{4.4.1}$$

there will be

$$0 \in [-r-t, r-t] \subseteq [-2r, 2r] \subset [a, b] \tag{4.4.2}$$

As the function $f(x)$ is positive over $[a, b]$, its convolution with the Gaussian $g(x) > 0$ remains positive over the same interval $[a, b]$. That is,

$$G*f(x) = \int_{[a,b]} g(t)f(x-t)\, dt > 0 \qquad \forall x \in [-2r, 2r] \subseteq [a, b] \tag{4.4.3}$$

For any $x \in [-r-t, r-t]$ satisfying (4.4.1) and (4.4.2), as $f(x)$ is convex over $[a, b]$, there will be

$$f(x) \equiv f(\lambda(-r-t)+(1-\lambda)(r-t)) \ge \lambda f(-r-t) + (1-\lambda)f(r-t) \qquad \lambda \in [0, 1] \tag{4.4.4}$$

Thus, applying the Gaussian convolution to $f(x)$ over $R=[-r, r]$ gives

$$\begin{aligned} G*f(x) = \int_R g(t)f(x-t)dt &= \int_R g(t)f(\lambda(-r)+(1-\lambda)r-t)dt \\ &= \int_R g(t)f(\lambda(-r-t)+(1-\lambda)(r-t))dt \\ &\ge \int_R g(t)[\lambda f(-r-t)+(1-\lambda)f(r-t)]dt \\ &= \lambda\int_R g(t)f(-r-t)dt + (1-\lambda)\int_R g(t)f(r-t)dt \\ &= \lambda G*f(-r) + (1-\lambda)G*f(r) \end{aligned} \tag{4.4.5}$$

Hence, after a single Gaussian convolution the function $F(x)=G*f(x)$ remains convex as well as positive over the top part $R=[-r, r] \subset [a, b]$.

Q.E.D.

Certainly, behaviour of the convex feature described by this theorem is only *locally* available because the convexity may be lost after sufficiently many iterations of Gaussian convolutions which may introduce significant influence from areas beyond $[a, b]$. An estimate is made in the following corollary, which gives the inferior band of the smoothing iterations where the convexity remains over the top part or at the peak of a bulge.

Corollary 4.1: *Suppose the neighbourhood radius* r *of the top part satisfies* $\min\{-2r-a, b-2r\} \geq n > 0$ *(*n *is an integer), the convexity at the peak will be preserved for, at least,* $n+r$ *iterations of smoothing.*

Proof: From Theorem 4.1, convexity over the top part will not be changed by a single Gaussian convolution. According to the square mask propagation theorem in Chapter 3, when the Gaussian convolution is implemented using a 3×3 square smoothing mask, the influence from areas beyond $[a, b]$ towards the top part will propagate only one step at a time. So the convexity over the top part $[-r, r]$ will be prevented from the influence for, *at least*[1], n iterations of smoothing. As the peak centres at the interval $[-r, r]$, its convexity will be preserved for, *at least,* $n+r$ iterations of smoothing.

Q.E.D.

Even when the peak itself disappears, its effect has not yet. Some other shapes stay as the residual products of the interaction between the bulge and the plane and can only be removed by further smoothing. This can explain why the effects of a small spot may remain for a *long* time on the KH sign image when the scale increases from fine to large.

Remark 4.3: *The above theorem is still available if "convexity" is replaced by "concavity" or "KH sign"*

Corollary 4.2: *Replacing the Gaussian kernel with other symmetric, positive functions, the convexity remaining theorem is still available.*

Proof: In the proof of the theorem, there is no restriction to $g(x)$ except that it is symmetric and positive. This implies that the conclusion actually applies to convolutions using any symmetric and positive function as the kernel.

Q.E.D.

Recalling the characteristics of spline smoothing and its relationship with Gaussian smoothing (cf. Theorem 3.6) in Chapter 3, the convexity remaining theorem has actually been implied there. It is very trivial to make it explicit as in the above corollary.

Remark 4.4: *Even in corollary 4 2 the restriction to* $g(x)$ *is still rather strong A weaker condition is actually implied in (4 4 3), where* $g(x)$ *is not necessarily positive everywhere over* I *as long as the integral sum or the convolution with* $f(x)$ *keeps positive over* I *This allows a broad range of functions to be chosen as the convolution kernels The explicit diffusion mask in Figure 3 3 7 shows such an example, where the central weighting coefficient is negative, but the sum of all coefficients of the mask is positive.*

---

[1] The above corollary gives only an inferior band. Usually, the convexity at the peak will remain more iterations than this band In subsection 4.4.2, the "response rule" will show a circumstance where the "at most" case appears.

Remark 4.5: In the proofs of Theorem 4.1 and Corollary 4.1 and 4.2, simply replacing interval neighbourhoods with plate neighbourhoods and adopting vectors as variables will give a proof for the 2-D case. So, the symmetry property of the smoothing kernal $g(x)$ is also not essential for convexity remaining.

Theorem 4.2. *(Interim shape creation.) Interim shapes will appear at the intersection of those conjugate shapes when the magnitude of curvature there becomes less than the zero thresholds given to K and/or H.*

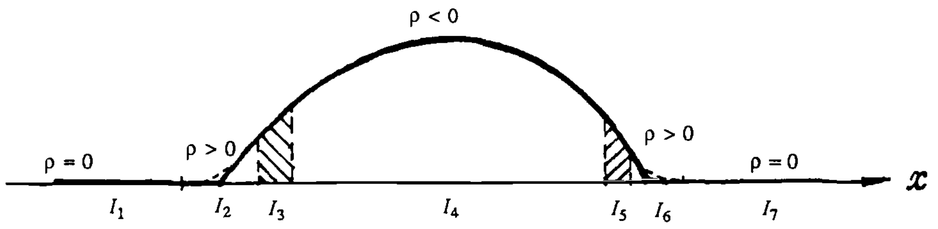


Figure 4 4 4 Interim shapes appear around the inflection points of the curve.

Proof: As principal curvatures can be derived from the Gaussian and mean curvatures, the threshold of principal curvatures, ε_ρ can be determined from the thresholds of the Gaussian and mean curvatures. Hence, for simplicity without lose of generality, the proof starts in the normal section of the surface, which goes along one of the principal directions in this case

As the convex spherical patch is on a flat background, $f(x)$, the surface contour in the normal section is composed of three connected segments: a convex segment and two straight lines in both sides. In the discrete case, curvature signs along the contour $f(x)$ can be divided into seven sub-intervals as shown in Figure 4.4.4, where curvatures of both straight lines at intervals I_1 and I_7 are zero, curvatures at I_2 and I_6, the intervals around the concave corners of both lines and the convex segment, are positive, curvatures of the convex top part at I_4 are negative, and curvatures at the remaining intervals I_3 and I_5, *i.e.*, the intervals between the top and both corner parts, are of interest.

From the convexity remaining theorem, the contour $f(x)$ remains convex at the top part I_4 after a single Gaussian convolution, thus the curvature of $f(x)$ remains negative there:

$$\rho(x) < 0 \qquad\qquad x \in I_4 \tag{4.4.6}$$

whereas $f(x)$ remains concave at either corners in I_2 and I_6, hence the curvature $\rho(x)$ remains positive there:

$$\rho(x) > 0 \qquad\qquad x \in I_2 \cup I_6 \tag{4.4.7}$$

During the scale space filtering, the order of continuity of the curve $f(x)$ and its curvature $\rho(x)$ rise, meanwhile the maximum of curvature magnitude decreases monotonically. According to the mean value theorem, curvature $\rho(x)$ over $I_3 \cup I_5$, will get intermediate values around zero and become smaller in magnitude as well:

$$|\rho(x)| < \varepsilon \qquad\qquad x \in I_3 \cup I_5 \tag{4.4.8}$$

When the magnitude of curvature at an intermediate point $x_0 \in I_3 \cup I_5$ becomes less than the given zero threshold ε_ρ:

$$|\rho(x_0)| < \varepsilon_\rho \qquad\qquad x_0 \in I_3 \cup I_5 \tag{4.4.9}$$

and its curvature sign changes to zero, an interim curve shape is created there.

Note that Gaussian curvature K is the product of both principal curvatures and the mean curvature is the average of both principal curvatures. When the magnitude of the other principal curvature is greater than ε_ρ, the above magnitude change leads to the sign change of K at the point x_0. Otherwise, it leads to the sign change of H at the point x_0. In either case, an interim surface shape will be created, of which the sign of K must be zero. Hence, it will be one of the three types: ridge, valley or flat; but will never be minimal.

In the case of a surface point of arbitrary shape type, consider curvature signs of the neighbouring points in both sides along either principal direction, the mean value theorem applies when the signs in both sides are different. Thus the conclusion is still available.

Q.E.D.

Remark 4.6: *For the same reason, a created interim shape may disappear in due time during scale space filtering*

Theorem 4.3. *(Final fading shape:) When a non-minimal surface changes to flat through successive Gaussian smoothing, for any zero thresholds of K and H satisfying the consistent zero thresholding inequality (4.3.10), the final fading non-flat patch must be cylindrical: a ridge shape for a globally convex surface and a valley shape for a globally concave surface*

Proof: During the course of a patch changing to flat, if its principal curvatures satisfy

$$C_1 \equiv 0 \qquad\qquad C_2 \neq 0 \tag{4.4.10}$$

it is a cylindrical patch, the conclusion follows. Otherwise, for a non-minimal surface patch its principal curvatures can be expressed as

$$C_1 = O(c)\ , \quad C_2 = O(\mu c) \qquad \text{when} \quad c \to 0 \tag{4.4.11}$$

where $\mu = const \neq 0$, So, the Gaussian curvature K and the mean curvature H satisfy

$$K \equiv C_1 C_2 = O(c^2)\ , \quad H \equiv \tfrac{1}{2}(C_1 + C_2) = O(c) \qquad \text{when} \quad c \to 0 \tag{4.4.12}$$

that is

$$K = O(H^2) \qquad \text{when} \quad c \to 0 \tag{4.4.13}$$

Hence, for any zero threshold ε_K and ε_H satisfying the consistent zero thresholding inequality (4.3.10):

$$\varepsilon_K \geq 2|H| \cdot \varepsilon_H + \varepsilon_H^2$$

there must exist a small $c = O(\varepsilon_H) < <0.5$, where

$$|K| < \varepsilon_K < \varepsilon_H < |H| \tag{4.4.14}$$

which leads to the patch falling into the (*K,H*) category: (o,–) for a ridge or (o,+) for a valley. Both are cylindrical shape.

From the convexity remaining theorem, the top part of a globally convex patch maintains H negative before becoming flat, so the *final patch shape* should be a ridge; and for a globally concave patch, the *final patch shape* should be a valley.

Q.E.D.

Remark 4.7: *This theorem seems somewhat counter-intuitive if the effect of zero thresholds is neglected.*

In the case of minimal surface patch, (4.4.13) is invalid and the final patch shape remains minimal in theory. However, such a minimal surface patch seldom occurs in discrete data, especially in the presence of noise; and it may lose the attribute (principal curvatures $C_1 = -C_2$) when $|C_1| = |C_2|$ is close to ε_H due to the sensitivity to noise perturbation. So, it might be ignored in practice. A theoretical exception for the final fading shape also exists in the case of using (C_1,C_2) categorising, where the top point on a hemisphere patch will remain a peak before the whole patch becomes flat while both C_1 and C_2 are below the zero threshold together. However, such a case also seldom occurs in discrete data and is sensitive to noise perturbation. So, for ordinary cases, the final fading shape theorem applies to the (C_1,C_2) category as well as to the (*K,H*) category.

Corollary 4.3: *Given zero thresholds for (K,H) categorising in scale space filtering as in the consistent zero thresholding inequality (4 3 10), most non-flat surface patches will become flat via cylindrical types*

Proof: Before completely becoming flat, the remaining part of a non-flat patch will seldom be non-minimal. So, in most of the cases, the final fading theorem is available.

Q.E.D.

The above three theorems explain the local behaviour (remaining, creation and fading) of simple shapes during the scale space filtering. Among them the convexity remaining theorem is the central part.

Remark 4.8: *Recalling Yuille's conclusion that "there are interactions between the smoothing in the two directions of curvature" and his conjecture that these interactions "will probably be true for many smoothing methods", the above theorems verified Yuille's conjecture, at least, for smoothing with a symmetric, positive function as the convolution kernel In fact, Yuille's conjecture can be further verified for more smoothing methods because Remark 4.4 and 4 5 have pointed out that the essential condition for convexity remaining is (4 4 3), where the convolution kernal needs to keep the integral sum positive*

B. Concave features

When a feature has a concave spherical shape, its scale space behaviour can be illustrated as a series of *conjugate images* to the images in Figure 4.4.2 with respect to these conjugate shape pairs: <peak, pit>, <saddle valley, saddle ridge> and <ridge, valley>. (cf.

subsection 4.2.2 for definitions). To make it easy to observe the conjugacy between both series, Figure 4.4.5 shows the scale space behaviour of both concave and convex features on the same flat background.

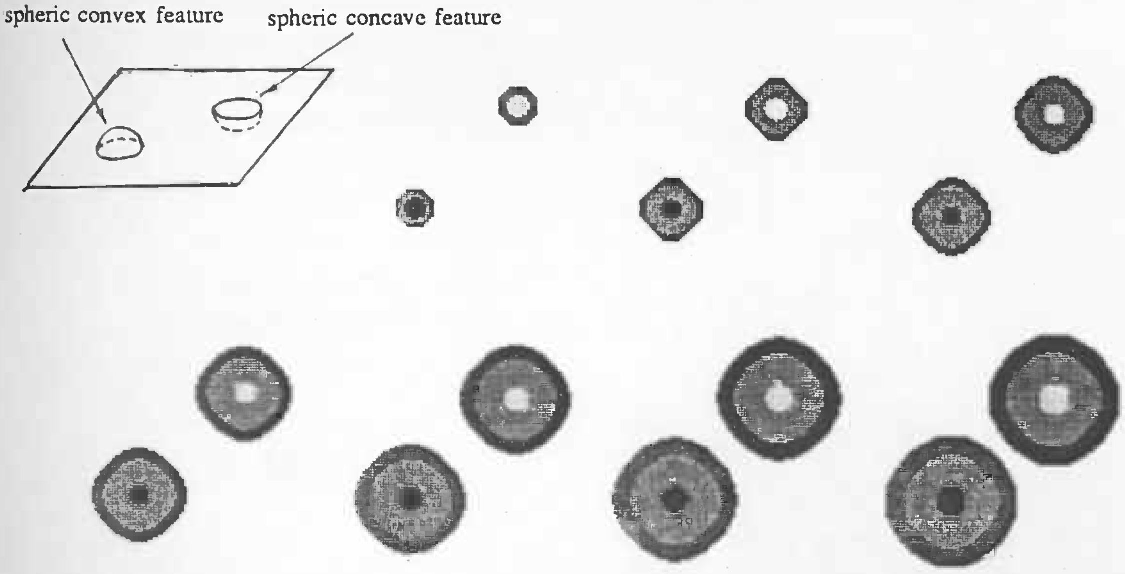


Figure 4.4.5. The conjugacy of a concave feature's scale space behaviour with the convex feature's on the same flat background ($\sigma = 1, 2, ..., 8$).

4.4.2. The scale space behaviour of small features on a peak background

A. Convex features

In the case of a convex feature on a peak background as shown in Figure 4.4.6, both the background and the feature spot have the same (BLUE) shape and around their intersection there is a VIOLET ring of the saddle valley shape accompanied by two RED rings of saddle ridge shape at both sides. During smoothing, both RED rings of saddle ridges expand and cover the VIOLET ring of saddle valley to join as one, then this RED ring contracts and disappears under pressure from the expanding BLUE spot.

The behaviour of features here is different from that in the last case in two aspects. First, the BLUE spot of the peak feature *expands* until it merges with the BLUE (peak) background. Second, until then no YELLOW colour (valley) or ORANGE colour (ridge) appears. The reason is that, in spite of the surface distortion introduced around the intersection between them, both the small feature and the background have the same shape type, thus the global shape is so strongly controlled by the peak patch that a finer scale increment is required to observe a locally flat area created near the intersection in the process. In fact, an important rule can be derived from the first aspect of the feature's behaviour:

Rule 4.1 *(Response): If a feature has the same shape type as that of the background patch in which it is embedded, the feature will not lose its shape attribute in scale space filtering; otherwise, the feature will lose it when the background patch strongly dominates the global shape.*

This rule shows in which circumstance the "*at most*" case of the convexity remaining appears, that is, the local convexity remains through the scale space. This rule also could be used in an active form by passing several known shapes as the test signals over an unknown shape surface, then observing the behaviours of these signals in their local areas (called testing areas). If a signal spot maintains its shape attribute, this is a strong clue that the unknown background surface has the same shape type as the testing signal's. To track the responses from the surface of an unknown object in this way, scale space filtering is only needed over every covering area (*i.e.*, every area which covers a testing area). Although each covering area must be sufficiently large to contains the testing area so as to protect it from scale effects from outside areas[1], the covering areas themselves are still much smaller than the whole unknown surface. And the computation of KH sign image can even be limited in the testing areas only. So, this is a local testing. It may save a lot of computation.

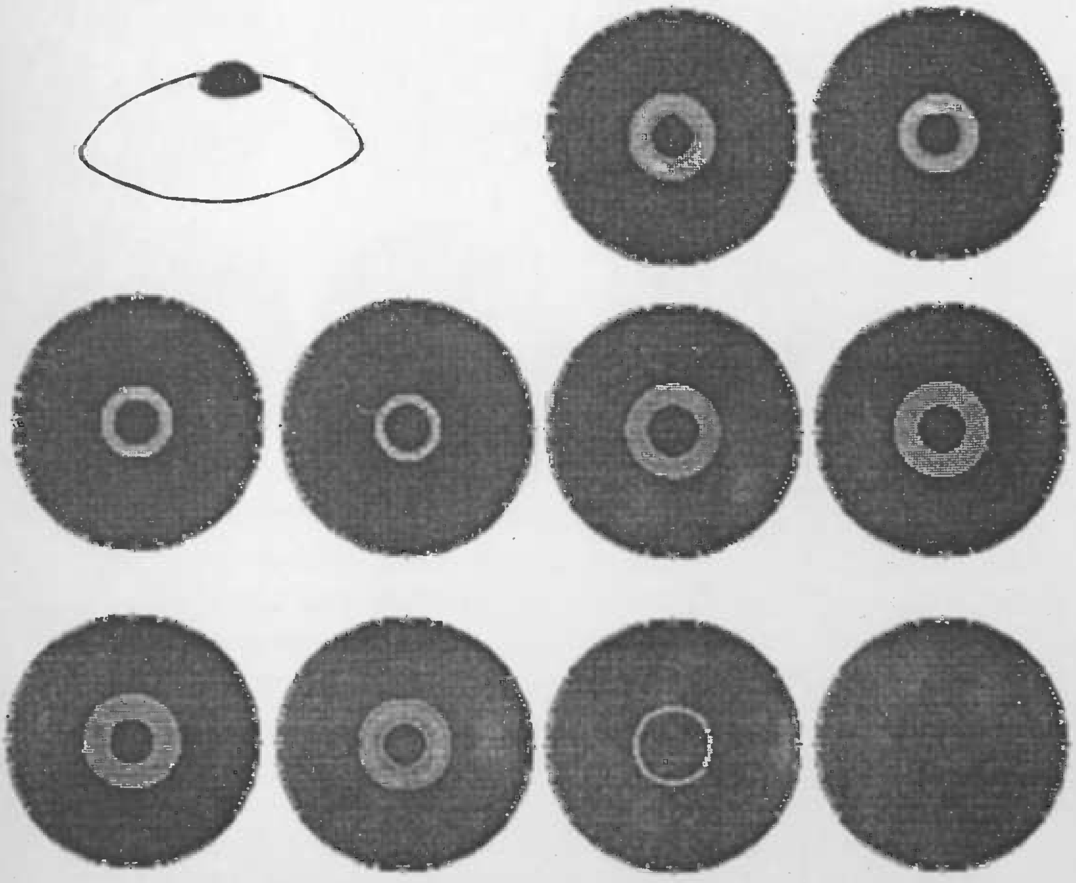


Figure 4.4.6. Scale space behaviour of the convex feature on a peak background patch ($\sigma = 1, 2, \ldots, 10$).

B. Concave features

When the feature has a concave spherical shape as shown in Figure 4.4.7, the feature spot will be a GREEN spot with a RED ring of the saddle ridge shape around the intersection with the BLUE background. During the filtering, the RED ring of the saddle ridge expands outwards while the GREEN spot changes little. In the middle stages a thin VIOLET ring of the saddle valley appears as the interim zone between the RED ring of the saddle valley and the GREEN spot of the pit, then disappears when the RED ring of the saddle

---

[1] cf. the "structure rule" in subsection 4.4.9.

ridge and the GREEN spot of the pit begin to contract. After that the saddle ridge and the pit spot vanish quickly from the peak patch which dominates the global shape throughout the filtering.

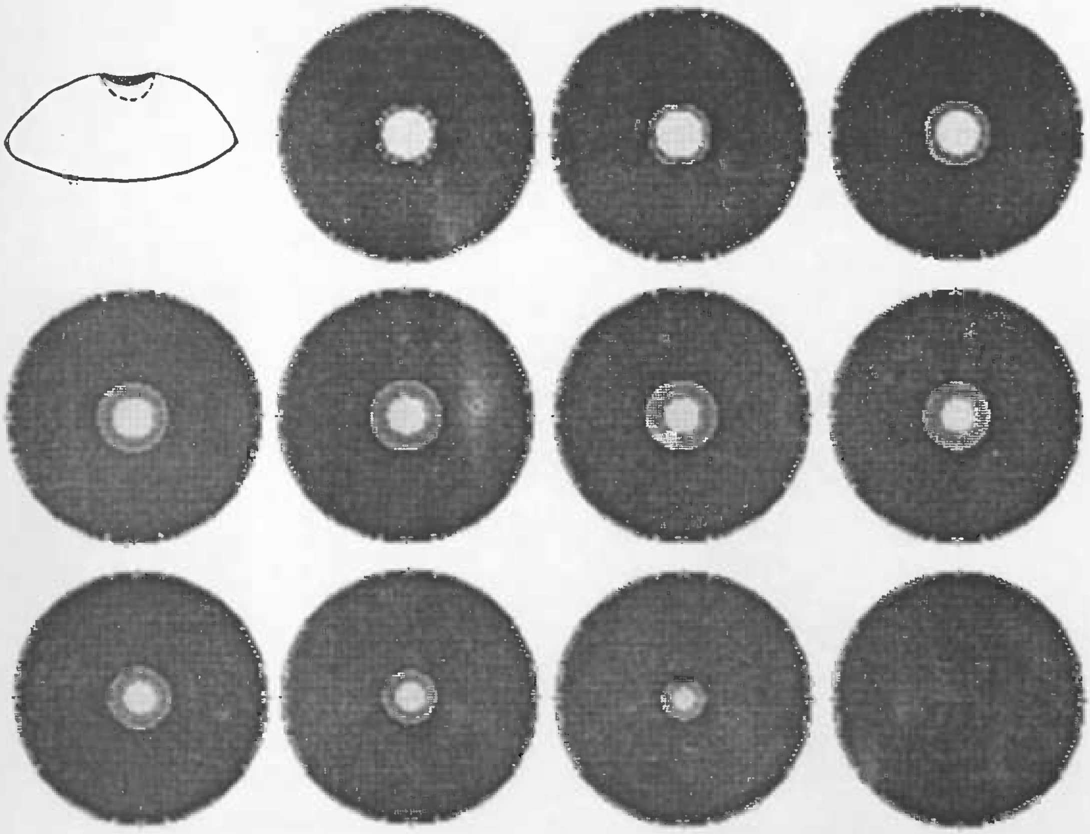


Figure 4.4.7. Scale space behaviour of the concave feature on a peak background patch ($\sigma = 1, 2, ..., 11$).

Although a concave feature has the conjugate shape to the convex feature, its behaviour on a peak background, as have been shown above, does *not* conjugate to the convex feature's as happened in the case of a planar background because the global shape has now been changed from being flat to being a peak. This is an obvious indication that a small feature changes its behaviour on different background patches. Even though the small feature changes the shape, and thus changes the behaviour of the background in a local area and over a certain range of scales, the background feature retains dominance in the interaction with the small feature. Hence, it is likely that *conjugacy will occur when both backgrounds have conjugate shapes*[1]*. This hypothesis will be checked in later observations.*

4.4.3. The scale space behaviour of small features on a pit background patch

This case is now conjugate to the case of small features on a pit background with respect to these conjugate shape pairs: <peak, pit>, <saddle ridge, saddle valley>. It confirms the conjugacy condition mentioned in subsection 4.4.2.

4.4.4. The scale space behaviour of small features on a saddle ridge background

---

[1]: The flat shape can be seen as a "self-conjugate" shape type.

A. Convex features

In the case of a convex feature on a saddle ridge background, the process is as shown in Figure 4.4.8.

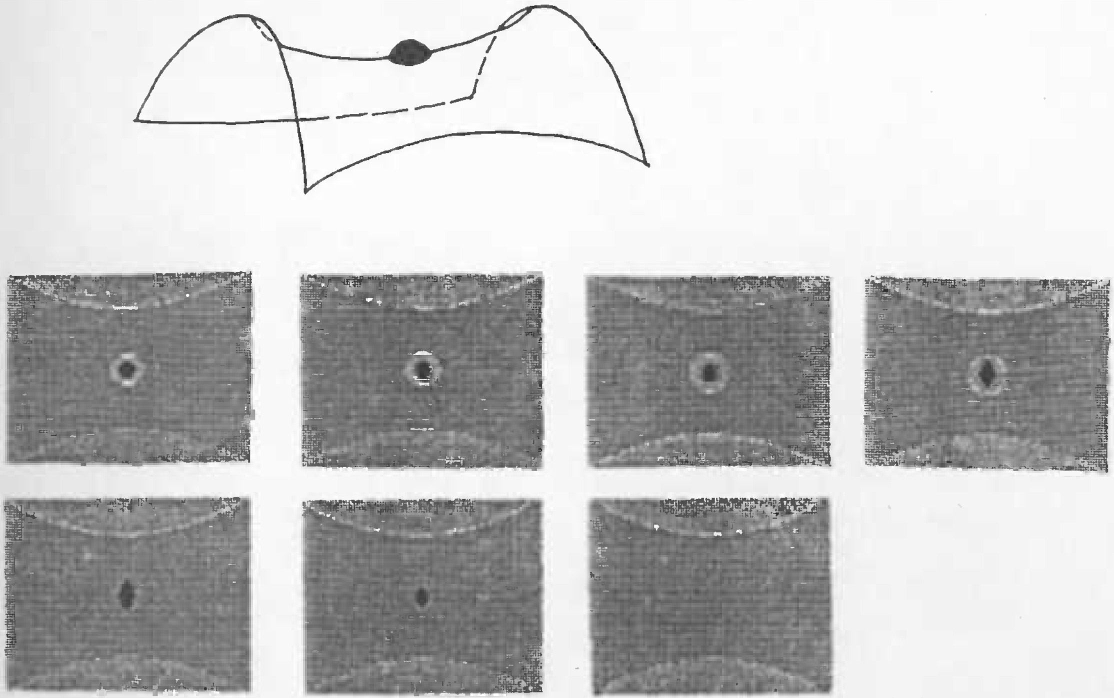


Figure 4.4.8. Scale space behaviour of a convex feature on a saddle ridge background patch ($\sigma = 1, 2, ..., 7$).

In the KH sign image, the saddle ridge background is RED, the small peak feature is a BLUE spot surrounded by a mixed VIOLET/GREEN ring of saddle valley and pit shape around the intersection, where the saddle valley shape coincides with the principal direction ϕ_1, and the pit shape coincides with the principal direction of ϕ_2. During the filtering, a RED ring of the saddle ridge shape appears between the BLUE spot and the VIOLET/GREEN ring. It expands outwards accompanied by the BLUE spot, and merges with the RED background of the saddle ridge shape when the VIOLET/GREEN ring of the saddle valley and pit shape vanishes, then the BLUE spot contracts to nothing.

Remark 4.9: *The "response rule" is confirmed in the way that an interim shape (the RED ring of the saddle ridge) maintains its attribute through the filtering since it has the same shape type as the background.*

B. Concave features

When the small feature has a concave shape, it is a GREEN spot on the RED background, there is no complete ring but two BLUE spots as a pair of local bulges at the intersection of the spot and the background. During the filtering, both BLUE peak spots only expand and contract gently and for a slight while; once a VIOLET ring of saddle valley shape appears around the intersection, both BLUE and GREEN spots begin to vanish. The saddle shape covers the place where the concave feature stood, then it also disappears. These are shown as in Figure 4.4.9.

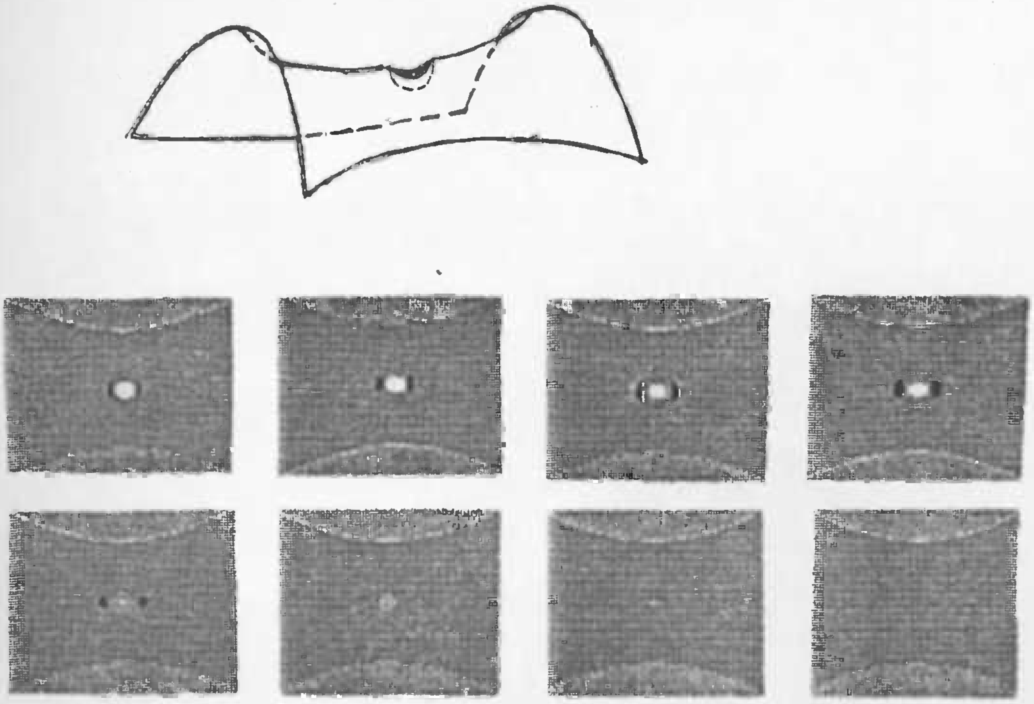


Figure 4.4.9. Scale space behaviour of a concave feature on a saddle ridge background ($\sigma = 1, 2, \dots, 8$).

Remark 4.10: *In this case all local shapes related to the concave feature disappear since none of them has the same shape with the background, so there is no response to the background shape.*

4.4.5. The scale space behaviour of a small feature on a saddle valley background

When the background is a saddle valley, the scale space behaviour of a small feature conjugates its behaviour on the saddle ridge background with respect to the conjugate shape pairs: <saddle ridge, saddle valley> and <pit, saddle valley>. The conjugacy hypothesis is confirmed in this case again.

4.4.6. The scale space behaviour of a small feature on a minimal background

When a convex feature is at the centre of a minimal background as shown in Figure 4.4.10, the KH sign image of the background itself is usually divided into four regions of saddle ridge shape (RED) and saddle valley shape (VIOLET) being symmetric to the centre, at which the peak shape feature is a BLUE spot that is surrounded by a VIOLET ring of the saddle valley shape attached by a pair of small GREEN spots of pit shapes inside the VIOLET regions. During the filtering, the small GREEN spots vanish, the VIOLET ring of saddle shape is replaced by the (RED) saddle ridge shape, the BLUE feature spot contracts to nothing, then only two colours belonging to the minimal patch are left. The influence of the small feature might still exist in the depth image but it is so weak that it is hard to see in the KH sign image.

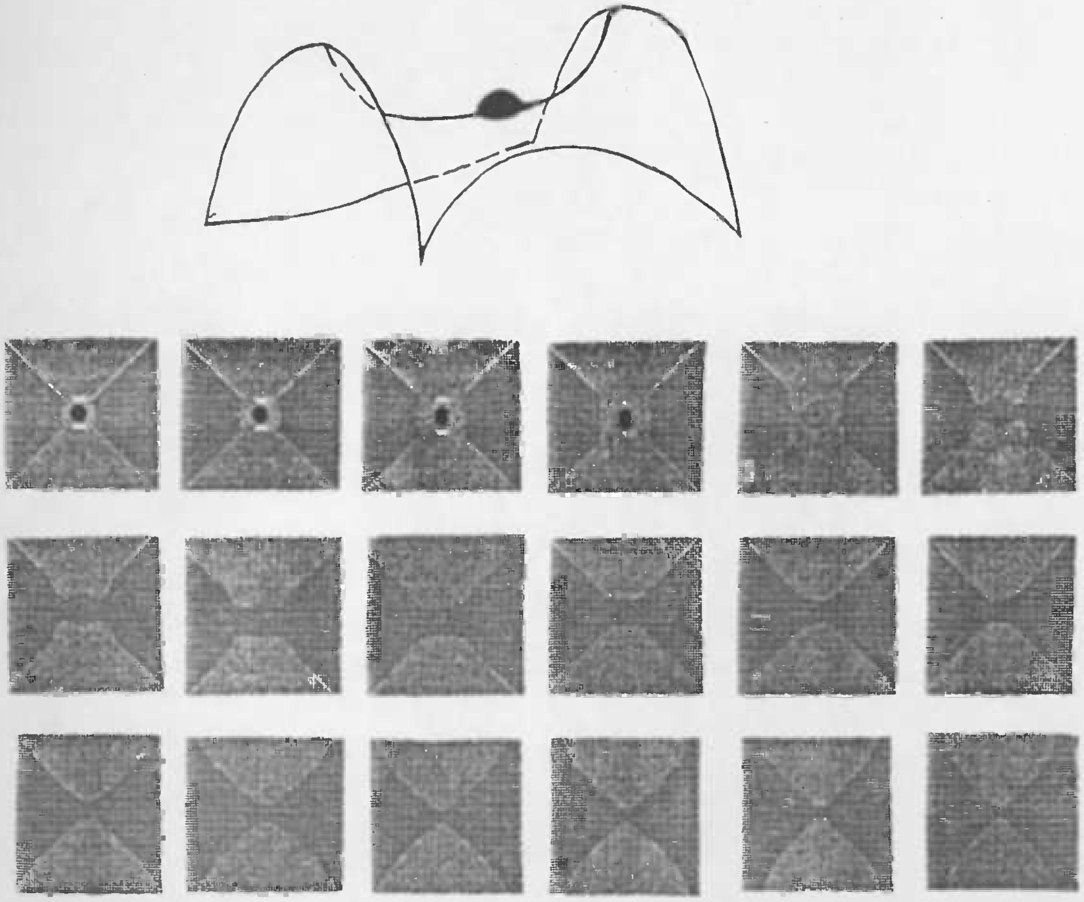


Figure 4.4.10. Scale space behaviour of a convex feature on a minimal background patch ($\sigma = 1, 2, \ldots, 18$).

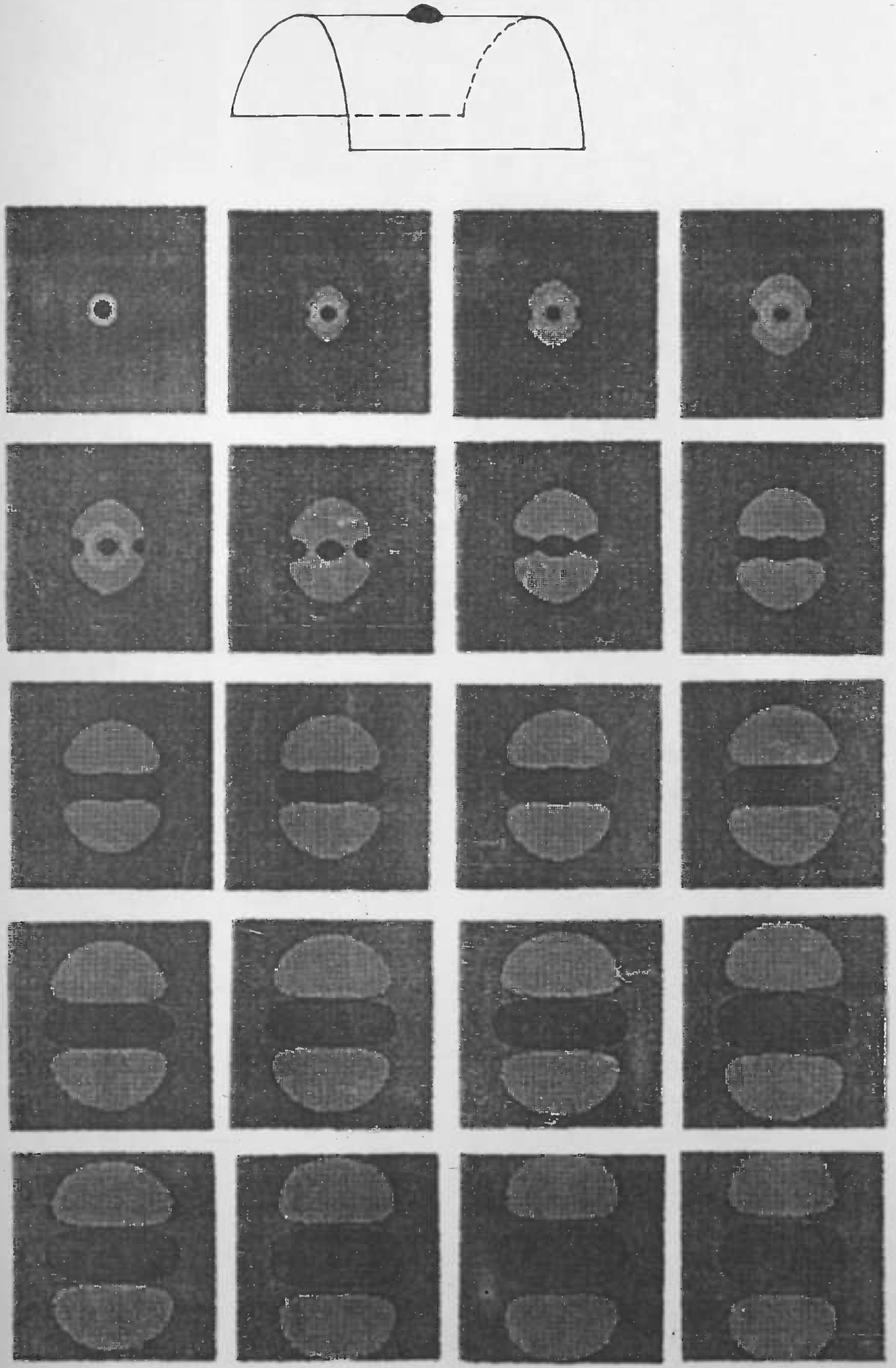
The KH sign image of a concave feature on the minimal background conjugates to that of the convex feature, because the minimal shape is also of the "self-conjugacy" like the flat shape. So once the background feature becomes self-conjugate, image conjugacy is determined by the small feature.

4.4.7. The scale space behaviour of a small feature on a ridge background

A. Convex features

When a small convex feature is on a ridge background, the convex feature is a BLUE spot surrounded by a VIOLET ring of saddle valley shape on the ORANGE background. During filtering, RED rings of the saddle ridge shape appear on both sides of the VIOLET ring of the saddle valley shape, then expand and cover the VIOLET ring. Meanwhile, parallel to the axis direction of the ridge background two small BLUE spots appear and grow outside the RED ring. Later, it cuts the RED ring into two parts to join with the BLUE feature spot in the centre. These connected RED, BLUE and RED spots expand together until a large scale level where they eventually begin to contract into three separated large spots and all disappear from the background. These are shown in Figure 4.4.11.

Remark 4.11: *The interim shape between peak and saddle ridge in the KH category is ridge, which is the same shape (colour) to the background patch here. The separation of the connected RED, BLUE and RED spots is due to this identity, as is the case of the concave feature on a saddle ridge background, where there is no complete ring at the initial time.*



(continued)

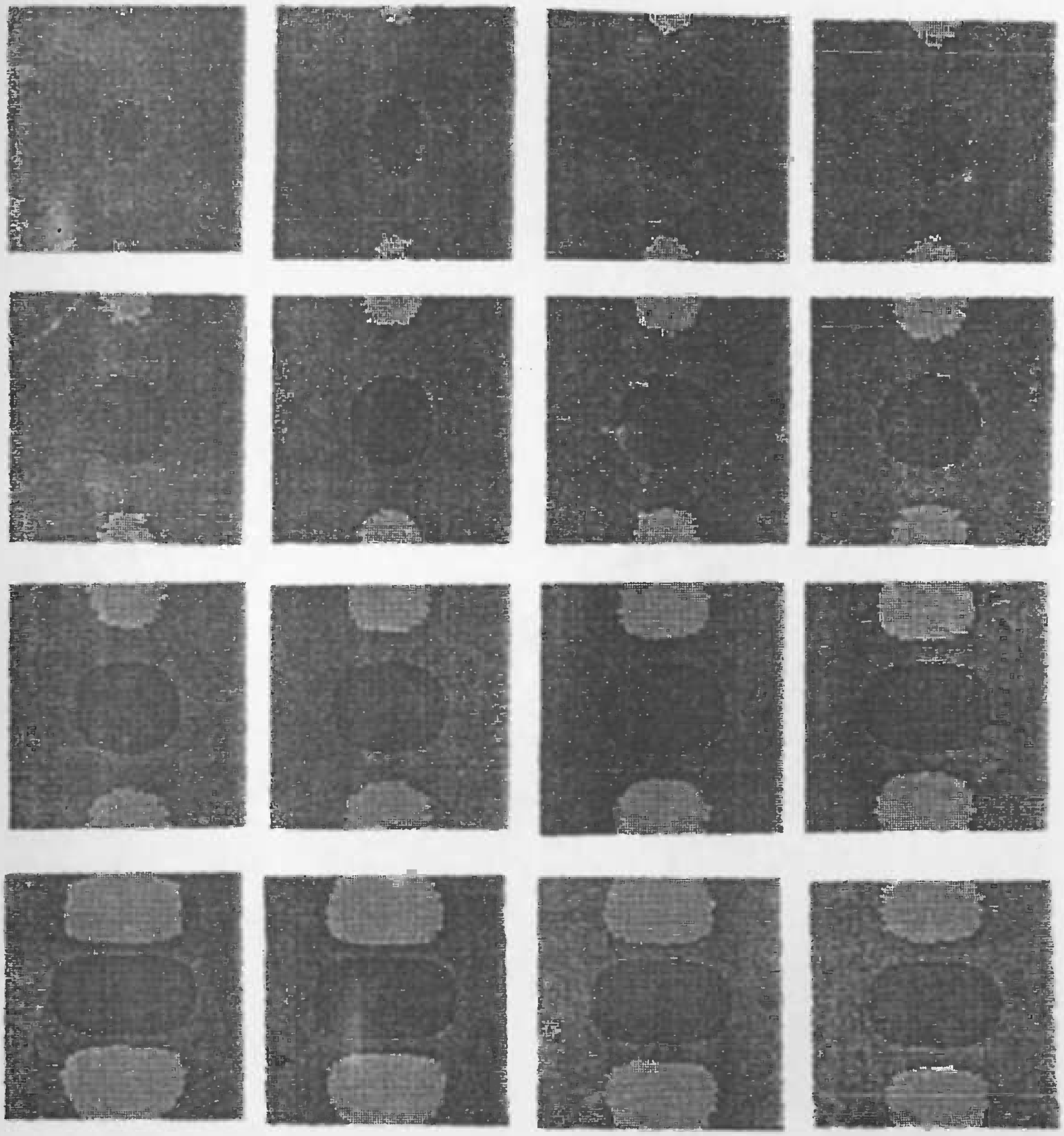


Figure 4.4.11. Scale space behaviour of a convex feature on a ridge background patch (σ = 1, 2, ..., 36).

b. Concave features

When the small feature is concave in shape, as shown in Figure 4.4.12 its KH sign image is a GREEN spot of pit shape surrounded by a RED ring of saddle ridge shape attached by two BLUE small spots of peak shape along the axis direction of the ridge patch. During filtering, the RED ring of saddle ridge shape and both BLUE spots of peak shape expand whereas the GREEN spot of pit shape contracts only a little. Later, a VIOLET ring of saddle valley shape appears and grows between the RED ring and the GREEN spot. It first covers the GREEN spot then quickly disappears. Only three colour spots now remain on the (ORANGE) ridge background: one RED spot for saddle ridge shape and a pair of BLUE spots for peak shape. Afterwards the behaviour of the concave feature conjugates to that of the convex feature with respect to the conjugate shape pair <RED, BLUE>.

Remark 4.12: *In contrast to all previous cases, the conjugacy of behaviour can occur between a concave feature and a convex feature on a ridge background, but it occurs only at the late stages of scale space filtering, where local area of the ridge background is near flat, i.e., it is almost self-conjugate.*

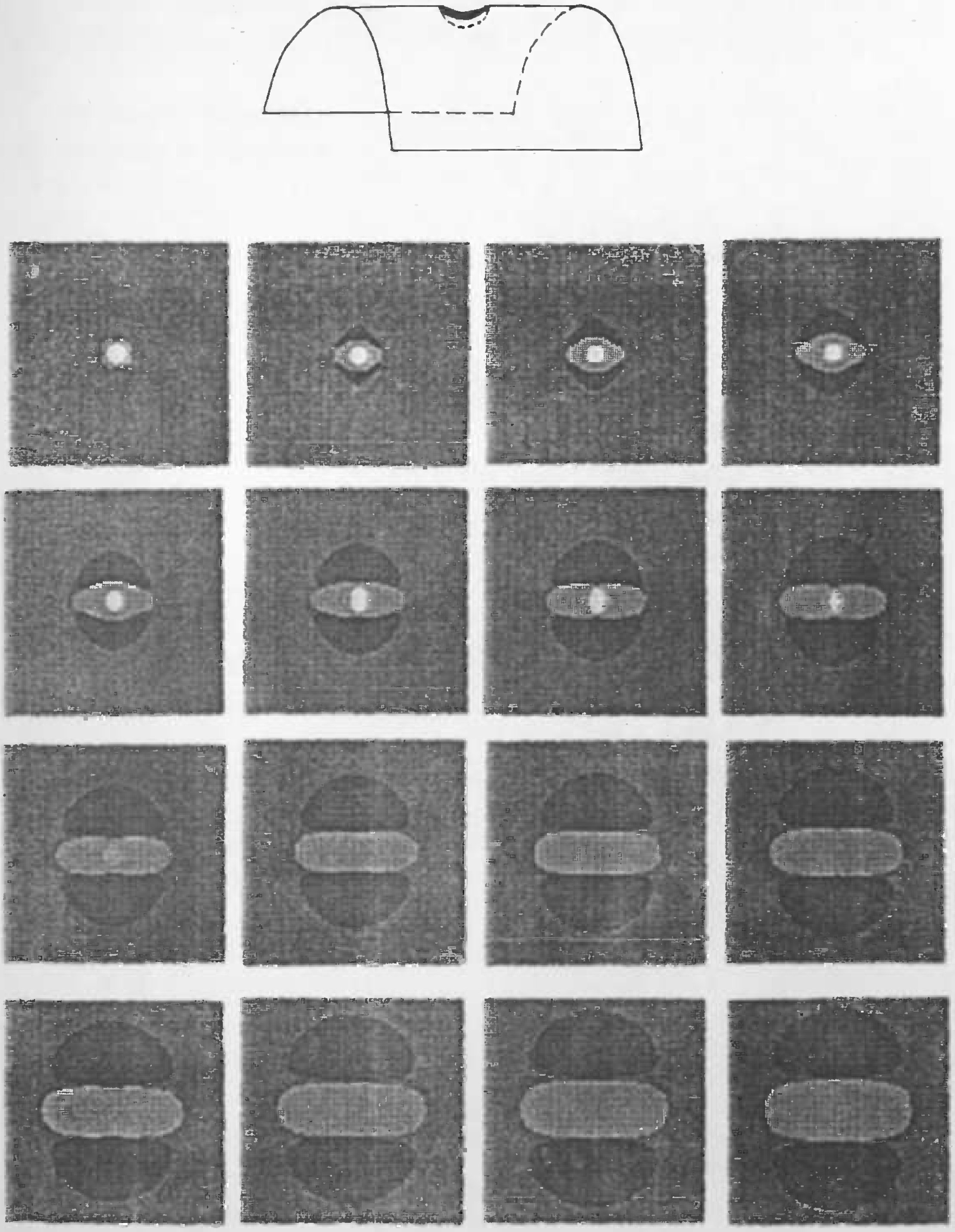


Figure 4.4.12. Scale space behaviour of a concave feature on a ridge background ($\sigma = 1, 2, \ldots, 16$, the radii of the spherical feature and the cylindrical background are 5 and 60 respectively).

Remark 4.13: *Similar to the case of a small feature on the flat background, the effect of a small feature on the cylindrical background is also dramatic: cylindrical features will be hard to extract because large amounts of smoothing will be needed to overcome small amounts of perturbation.*

4.4.8. The scale space behaviour of a small feature on a valley background

This case conjugates to the case of features on a ridge background with respect to these conjugate shape pairs: <peak, pit>, <saddle ridge, saddle valley> and <ridge, valley>.

The term of "conjugacy" was initially used as an abbreviation of the relationship between images or image series. By reviewing the observations in the above eight cases, it is easy to see that the "conjugacy" virtually reflects the interaction of behaviour between the small features and background features in scale space smoothing, where the background feature always plays a dominant role, and rules can be listed as the following:

Rule 4.2 *(Conjugacy)· Conjugacy occurs only when both background features have conjugate shapes, except the following two cases*

1) The background feature is a self-conjugate shape type (i e , flat or minimal), where conjugacy occurs when small features have conjugate shapes

2) The background feature is of a cylindrical shape, where the conjugacy occurs when small features have conjugate shapes, and it happens only at the late stages of the process when both backgrounds become nearly self-conjugate in a local area

Remark 4.14: *The conjugacy rule and the response rule show that the most important attributes of a patch are probably the curvature signs and the patch area. This gives important clues for tracking the movement of features in scale space*

4.4.9. The scale space behaviour of a small feature on a nest-structured surface

In the previous eight cases, the structure of a surface composed of the small feature and the background is quite simple It has been shown that the scale space behaviour of a small feature not only depends on its own shape but also on the shape of the background patch in which it embeds. When the background patch itself is a part of another larger patch which controls the global shape, question is raised: *In which way will the largest patch affect the others?* Matters now become complicated due to the nested structure that consists of the small, the intermediate and the largest patches.

Fortunately, when the intermediate patch is in a strong state, the influence from the largest feature is passed onto the small feature via the intermediate feature. That is:

Rule 4.3 *(Structure)· Nest-structured patches organise their scale space behaviours within a nest structure on the KH sign image*

This means that corresponding to their geometrical distribution, different features behave in a nested way at each scale level, and in a hierarchical structure over all scale levels. As behaviour of a feature is a local action, it can be isolated temporarily. This largely eases the tracking of the movement of features in scale space because small features can be temporarily neglected when its background patch is undergoing tracking as a feature along with a much larger shape feature.

This rule also gives a qualitative interpretation to the *convexity remaining* and the *response rule* from a global point of view.

As an example, observe the behaviour of a concave feature on a peak background patch on a convex spherical surface and compare it with behaviour of both concave and convex features on a simple peak background patch (see Figure 4.4.13). In general, as long as the surface curvatures change gradually and locally in a process like the Gaussian or diffusion smoothing, the structure of behaviour will be nested.

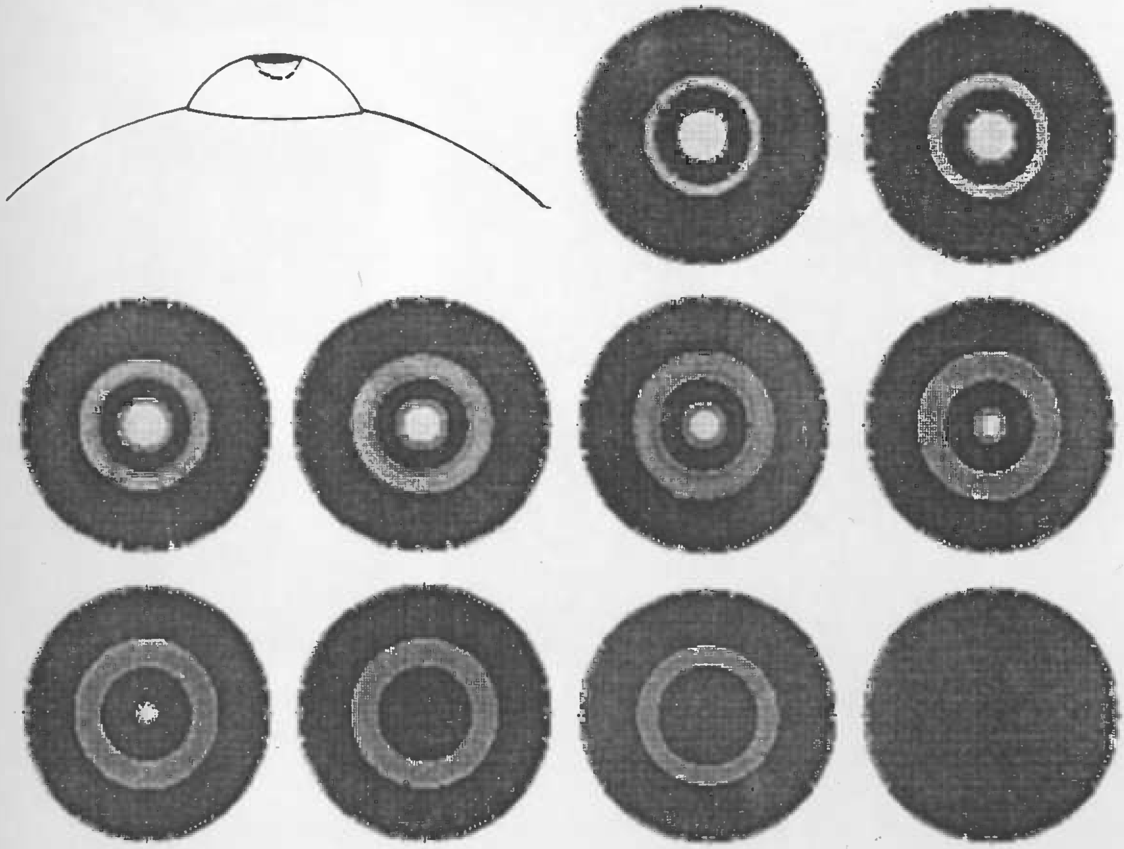


Figure 4.4.13. Scale space behaviours of features in a nested-structure, where a concave feature is on a larger convex feature which is on a convex spherical background ($\sigma = 1, 2, ..., 10$).

Another example is behaviour of features of a human face as shown in Figure 4.4.14. Pay more attention to the patches around the nose and the eyes. The (GREEN) pit patches gradually contract inside the (VIOLET) saddle valley patches, and the (VIOLET) saddle patches contract inside the (RED) saddle ridge patches which eventually disappear in the (BLUE) peak surface. And the last image in Figure 4.4.14 shows the global shape of a human face smoothed at a large scale level. During this process, a larger feature will remain for longer time on the KH sign image, as does a stronger feature of the same shape (colour).

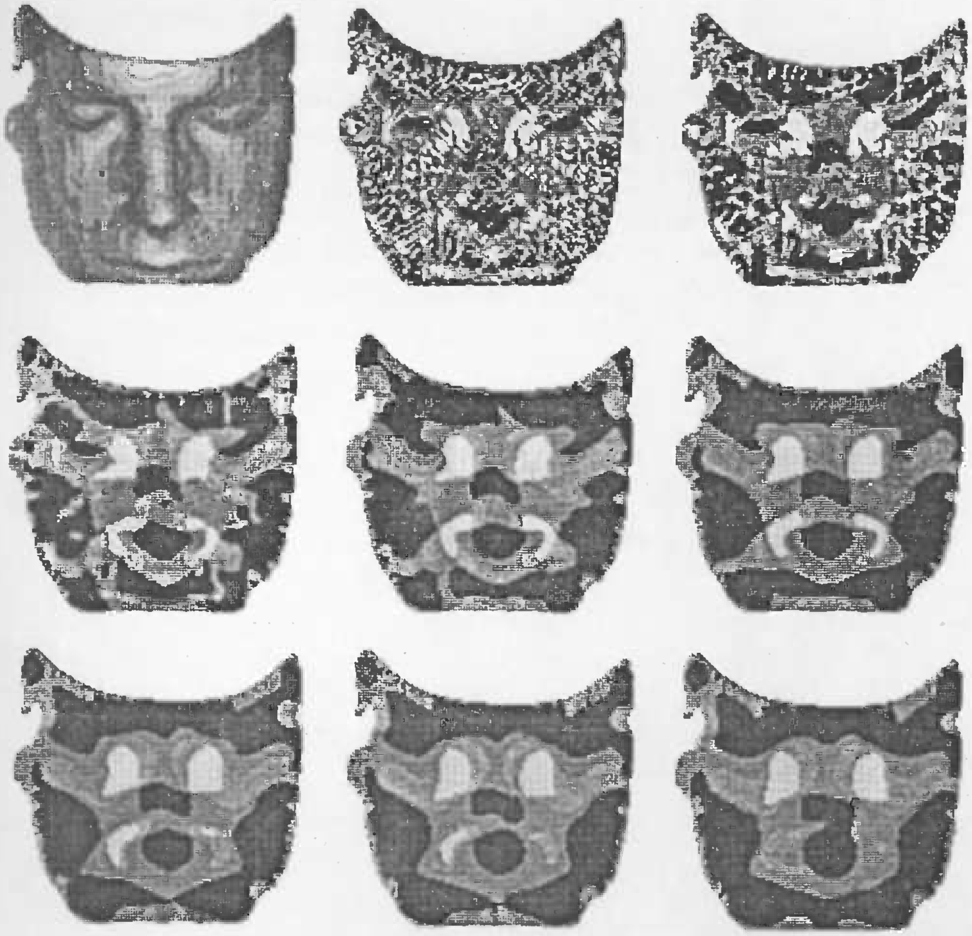


Figure 4.4.14. Scale space behaviours of features on a human face ($\sigma = 1, 2, ..., 8$).

4.5. Scale space behaviour analysis and shape feature extraction

Processing in scale space is quite different from processing at one single scale. The scale space filtering will suppress both small signals and noise without distinction. This kind of scale effect is hard to avoid. It makes the KH sign image more complex than might be expected and leads to techniques that are successful at a specific scale being not necessarily successful or wholly appropriate at other scales.

For instance, in the KH sign image, a pair of intersecting planes will produce not one or two patches but three patches! (*i.e.*, two planar patches and one ridge/valley patch at the intersection). Another example is the problem of "object melting into the background" as mentioned by Ponce and Brady [1987]. There are also other difficult problems in feature detection using the KH sign image. All these force research back to investigate the capability of the KH sign image itself.

4.5.1. The greatest advantage of KH sign image in surface feature extraction

Although the KH sign image is rather insensitive to the change of curvature at tangent join points due to the zero-drift effects (cf. Appendix 3), it is fairly sensitive to the sign change of curvature at the reflection points. Hence, the KH sign image can be relied on in detecting an orientation change or a C^2 discontinuity at the intersection of a ridge patch and a valley patch as shown in Figure 4.5.1, where for simplicity, only a cross-section of the surface is observed. The reason that a joint point p can be located very well by the KH sign image is that according to the convexity remaining theorem both convex and concave segments will maintain for a long time during the scale space filtering, so does the inflection point p as the zero curvature point between both stable segments of opposite curvature signs. Hence, this good performance will be maintained over a large range of scales. Therefore, it can be expected that a smooth surface composed of saddles and peaks or pits as shown in Figure 4.5.2 can be easily segmented using the KH sign image of the surface. It is probably the most significant advantage of the KH sign image as a tool in detecting surface shape features in scale space.

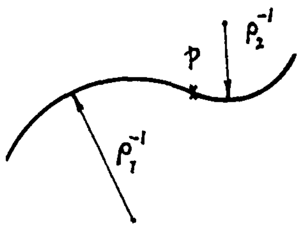


Figure 4.5.1. The cross-section of a ridge/valley surface, where C^2 discontinuity point p detected by the KH sign image.

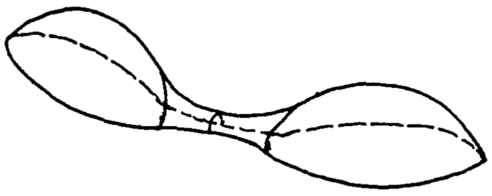


Figure 4.5 2 Smoothly connected peaks and saddle can be easily segmented by the KH sign image.

However, the curvature sign image may fail to detect a pair of convex (or concave) segments with a smooth join and with slight different curvature as shown in 4.5.3.

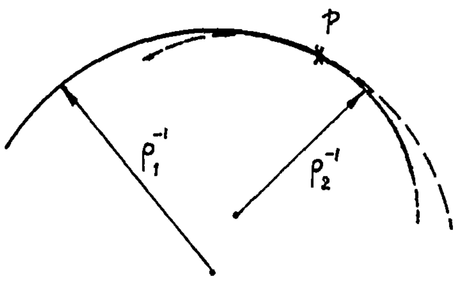


Figure 4.5.3. C^2 discontinuity point p missed from the KH sign image.

As a result, an elliptic arc will be identified as a uniform curve unless both curvatures are quite different as shown in the left of Figure 4.5.4, where a C^2 discontinuity might be detected at the point p; and sometimes, *e.g.*, when the sampling scale is improper (as shown in the right of Figure 4 5.4), it could be detected as a C^1 discontinuity (roof) instead of a C^2 discontinuity. Similar phenomenon will occur in detection of elliptic or generalised cylindric surface patches using the KH sign image. Nevertheless, these results are not too bad and can be accepted as long as they do not seriously influence the later recognition stage.

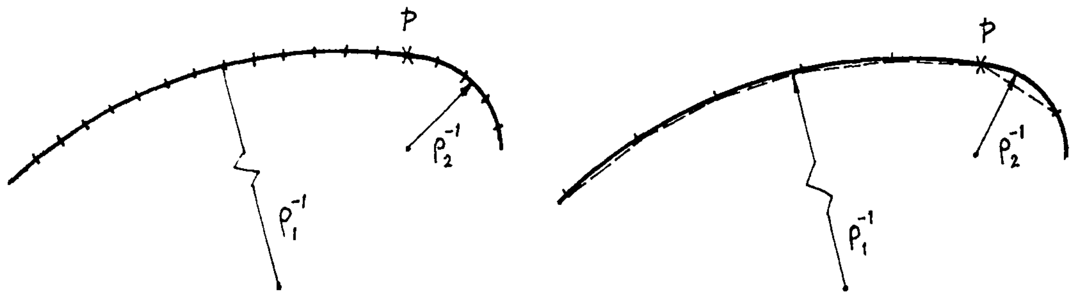
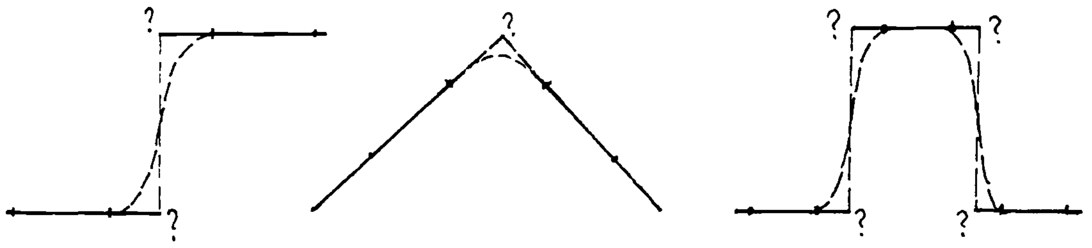


Figure 4 5 4. C^2 discontinuity at the point p detected as C^1 discontinuity by the KH sign image.

The other aspect is that the KH sign image gives only vague evidence for C^0 discontinuity around the edges of a step and for C^1 discontinuity around a roof as shown in Figure 4 5 5. Earlier, Link and Zucker [1985] have pointed out this similarly. These edges and the roof behave not like lines but like narrow strips at fine scales, they expand to broad strips even to patches in scale space. This is partially due to the discrete form and partially due to the scale effects.

Some reasoning is therefore needed for boundary location. As surface boundaries can be defined as the intersections of patches, the situation of a step is not so bad because the edges of a step always appear in pairs and have opposite curvature signs. Therefore when the scale is not too large they form a conjugate shape pair in the KH sign image. This gives an indication of the surface boundary. However, there is no clear indication to distinguish a roof from an conic arc around the peak point. The situation is similar for the bar shape, which can be thought of as a pair of steps. The KH sign image gives only vague evidence for segmentation.



step roof bar

Figure 4 5 5. C^0 and C^1 signs may provide only vague evidence about curve shapes.

Hence, when applying the KH sign image to detect discontinuities on a surface in scale space filtering, an interesting and plausible phenomenon happens: the response from a free-form surface, which is rich in alternations of large convex and concave patches, will correspond so well to the expected surface distribution that only KH sign image detection seems sufficient to directly segment the curved surface in the presence of noise; whereas the response from a simple surface, composed of some developable patches (mainly cylinders, cones and planes), is rather vague and not as expected. It thus appears that a complex curved surface is easier to segment than a simple one! In the next chapter, a discussion will be made to interpret why a complex curved surface, such as human faces, can be well segmented using only the KH sign image resulting from the diffusion smoothing.

4.5.2. The limit of the KH sign image in shape feature extraction

Although the KH sign image is strong in partitioning complex curved surfaces, its power is not unlimited. As an example, some KH sign images of an oil bottle at different scale levels is illustrated from left to right in Figure 4.5.6, and attention is drawn to the question why some simple surfaces are not so easy to segment using its KH sign image.

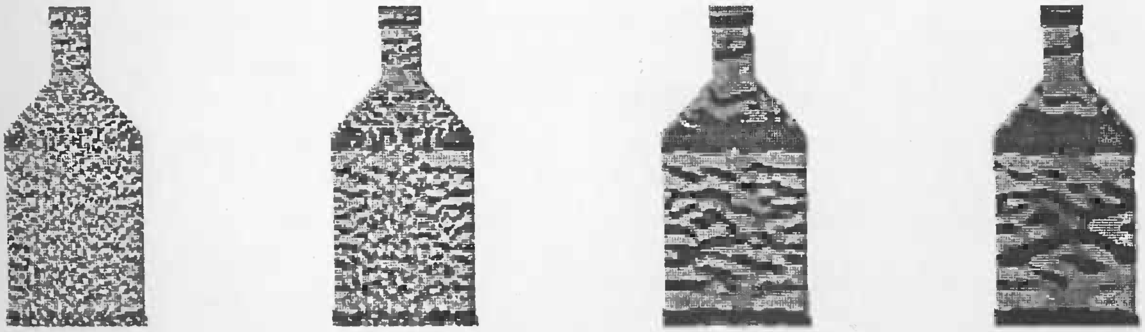


Figure 4.5.6. Odd textures stick on the smoothed surface due to the noise.

In the figure, the intersection patch between the conic shoulder and the cylindric body expands gradually as the scale increases, meanwhile some odd textures stick on the cylindrical body across scales. In fact, those textures are related to the extrema and zero-crossings of surface curvature [Ponce and Brady 1987]. The expansion of the intersection patch can be easily foreseen as surface distortion due to scale effects, and the roof position can be determined by a method proposed in [Ponce and Brady 1987] using the extrema of the principal curvature associated with its direction. However, the odd texture sticking phenomenon is a puzzle so far, because the surface appears to have been smoothed a lot in the reflection (cosine shading) image. Certainly, a careful observation has shown that the depth data of this oil bottle is seriously corrupted by noise, and the situation may be improved by replacing it with better quality data. Nevertheless, the odd texture sticking phenomenon still appears. An interpretation to this phenomenon is thus given based on the discussions on the stability of KH category and the behaviour of the small features.

The Gaussian curvature of a cylindrical surface is zero, so its sign is sensitive to small perturbation around the zero value. As the global shape is dominated by the revolution surface, concave noise will distort the ridge (cylinder) shape to locally form a small patch of the saddle ridge/valley shape, and convex noise distorts the surface to form a small patch of the peak shape.

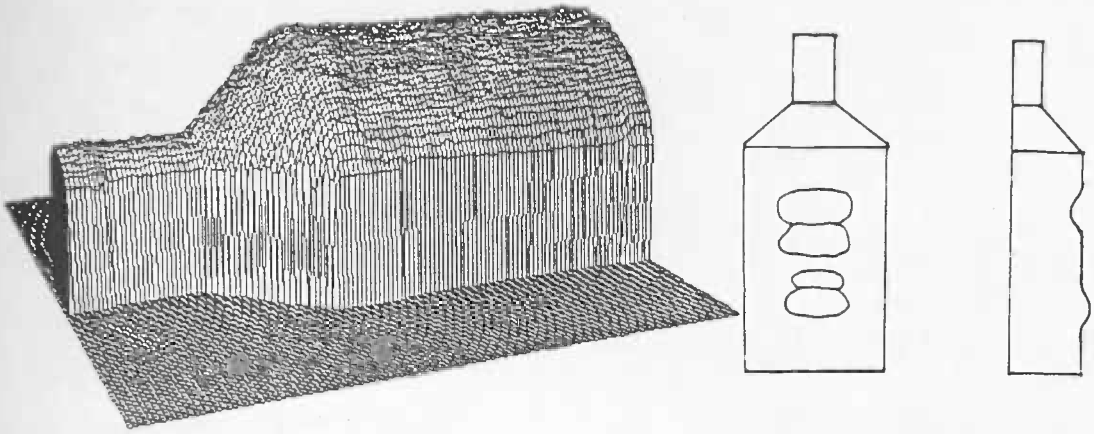


Figure 4.5.7. Ridge/valley saddle or peak patches on the cylindrical surface being distorted by noise.

In the side-view, a saddle valley is a deeper cave, a saddle ridge is a shallow cave and a peak is a bulge. As the global shape is cylindrical, any peak must appear together with a saddle, and vice versa, regardless of whether the noise is convex or concave in shape. According to the behaviour of small features on a ridge patch, noise influence will remain on the KH sign image over a quite large range of scales. Due to the behaviour of noise, mainly shapes of saddle valley (VIOLET), saddle ridge (RED) and peak (BLUE) appear in the earlier stage, afterwards saddle ridge (RED) and peak (BLUE) shapes remain no matter whether the noise is convex or concave in shape. When the same colour spots meet, they group into a larger spot or a strip. The odd textures in these shapes (colours) thus appear and stick on the KH sign image. Such a result also coincides with what has been known as the stability of the (K,H) category, where the peak and saddle ridge types have higher stability than the ridge type. This means that *even using large scale filtering, the noise spots are still hard to clear up from the KH image of a cylindrical surface.* In such a situation, a global operation is helpful to extract 3-D shape features, which will be discussed in the next chapter.

4.6. Summary

In this chapter, the scale space behaviour of large and small shape features is discussed in terms of the stability of large shape features and the creation, remaining and fading of small shape features. The consistent zero thresholding inequality is formulated as an estimate of the inter-relationship between the zero thresholds of Gaussian and mean curvatures, which governs the behaviours of large (primitive) features in the weak states. The theorems of "convexity remaining", "interim shape creation" and "final fading shape" are proven, which interpret the scale space behaviour of small features up to very large scales. Three rules are observed. The "conjugacy rule" reflects the dominant role of large features in the interaction with small features. In turn, the "response rule" shows it is possible to extract information about large features from behaviour of small features. Finally, the "structure rule" gives an overview of the structure of the behaviours of both simple and nested features in the KH sign image. These results summarises the behaviour of surface shape features.

Unlike the last chapter, the results obtained in this chapter are, basically, qualitative ones. Together with the diffusion smoothing theory, they supply a basis for tracking the movements of surface shapes and producing a stable shape description in scale space.

Chapter 5

Surface segmentation in scale space

5.1. Introduction

In the last two chapters a series of theoretical results were obtained for diffusion smoothing and the behaviour of surface features in scale space. They provide a framework for scale space processing and guidance in the dynamic processing of surface shape features. These results are now applied to surface segmentation.

The surface segmentation problem is related to the psychological phenomenon of figure-ground perception. This is the natural and unconscious ability of human to see objects as unities and to delineate their boundaries. Surface segmentation is also an essential stage in computer vision processing. By partitioning surfaces into uniform regions, it is possible to set up a symbolic description which is convenient for recognition.

There are many segmentation methods to approach this aim Traditionally, they are divided into two opposing types: edge-based methods and region-based methods. Section 5.2 will have a brief review of both methods. However, the goal of the review is not to contrast them, nor to claim which one is better for processing the range data, but to show that both methods are complementary from the viewpoint of surface curvature. This gives a way to unify them under one flag of the curvature-based segmentation method. Some features of such a curvature-based method are therefore discussed.

It is then natural to discuss the roles of surface curvature in surface segmentation in Section 5.3. Note that surface segmentation in this research is based on the surface curvatures in terms of the KH sign image which is used as an indication of the dynamic change of surface shape. Hence, getting a good KH sign image in scale space is crucial for the curvature-based segmentation method. Also, surface segmentation based on the KH sign image is posed as an under-determined mathematical problem of approximating surfaces up to the degree of curvature signs using the depth data alone. This shows the direction along which efforts should be made to obtain desired solutions. For these reasons several treatments in favour of shape preservation are therefore addressed. And answers are given to two questions: why a sculptured surface can be segmented using the KH sign image resulting from the diffusion smoothing of depth data, and why surface fitting is an indispensable global operation for Besl's method.

Section 5.4 points out the difference between surface fitting for region growing and for generating a KH sign image, then discusses the feasibility of generating a KH sign image via surface fitting as another solution of the above under-determined problem, where restriction is imposed on the polynomial basis fitting functions to ensure a proper extraction of curvature signs, and the surface fitting process is decomposed into 1-D polygonal approximation associated with parabolic modification.

In contrast to the global fitting operations, Section 5.5 shows that the principal directions, resulting from local operations, can be used directly to detect cylindrical patches. Therefore, the image of principal directions provides a means compatible with the KH sign image in segmentation using local operations. General surfaces may be segmented in scale space by using KH sign images, principal direction images and roof detections synthetically. This is the main version of the curvature-based segmentation methods advocated in this research.

An additional benefit from the above discussion is that it is also helpful in extending Besl's method from single scale processing to multiple scale processing. Some comments are made on such an extension in Section 5.6.

Section 5.7 is the summary of the chapter.

5.2. What is the curvature-based segmentation method?

In the literature of surface segmentation, for either intensity or depth images, methods are sorted into two opposing types: edge-based methods or region-based methods. A high proportion of the published papers are concerned with edge-based methods but more papers on region-based methods appeared in recent years due to the range striper being widely applied in the vision area.

5.2.1. Edge-based segmentation

Early work [Rosenfeld and Kak 1976],[Pratt 1978],[Maclead 1972] primarily used first and second order directional derivative operators to perform the 2-D differentiation so as to find the discontinuities where edges are located.

Marr and Hildreth [1980], Hildreth [1983] explored the combination of zero-crossing descriptions that arise from convolving an image with $\nabla^2 G$ of different sizes. They advocated the use of the spatial coincidence of zero-crossings across scales as a means to indicate the presence of real edges in a scene.

Canny [1983], Torre and Poggio [1984] proposed using directional edge finders and gave three criteria for estimating the performance of a good edge detector: 1) a good detection ability; 2) a good localisation ability; and 3) the uniqueness of the edge detected. The Canny detector itself satisfies these conditions and has been widely used. The optimisation of the three criteria used a linear combination of four exponentials which can be approximated closely by the first derivatives of a Gaussian.

Ponce and Brady [1987] proposed the Surface Primal Sketch method. A surface is first smoothed at a set of scales, and the principal curvatures and directions are computed everywhere; then the zero-crossings of the Gaussian curvatures and the directional extrema of the dominant curvatures are marked; finally, surface descriptions are matched at different scales to find points that lie on the step, roof, shoulder and smooth joint surface discontinuities. The results are good. Scale effects are recognised and a nice scale-related threshold is given to the curvature extrema test, but neither a thresholding of zero-crossing nor a convenient boundary treatment in scale space was discussed.

In short, the edge-based method emphasises the contrast among image pixels. It detects all surface edges as boundaries of different surface patches. Thus, the edge-based method requires that all detected edges form a set of closed curves.

5.2.2. Region-based segmentation

On the other hand, the region-based method emphasises the similarity among image pixels. It detects every patch using surface shape properties. The edges are the borders of patches.

Pavlidis [1972] posed the surface segmentation problem as an optimisation problem of functional analysis. He proposed to divide the image into thin strips so that some methods effective in waveform segmentation could be applied, such as the 1-D piecewise linear function approximation. Adjacent segments in different strips are merged if their slopes are sufficiently close to one another. The merging continues until no more merging is possible. It is certainly possible to get good results on a certain set of images by using this method which attempts to solve an inherently 2-D problem by decomposing it into 1-D problems that can be solved separately. But he did not make explicit what kinds of surfaces the method suits or what conditions the decomposition is subject to.

Haralick and Watson [1981] developed a facet model for image data. A facet is a connected region of image pixels with a local surface fitted to the pixel values in that region. By a convergent relaxation procedure that partitions an arbitrary image into facet regions in an iterative and parallel fashion, an ideal noise-cleaned version of the image is eventually created The borders of the facet regions form a segmentation of the image. From the example of a house scene shown in the paper, the slope-facet model gets good results in interior regions, but does not get good results at the edges, especially, with slanted, high contrast edges Similar ideas are involved in [Pong *et al* 1984] using a facet model region grower based on property vectors. Good results were obtained on a complex aerial scene.

Faugeras *et al* [1983] developed a region-growing method for segmenting range data into small numbers of planar and quadric patches within some fitting error thresholds. During the region growing, adjacent regions with the smallest fitting error (less than the threshold) are merged, which continues until no further merge is possible. Selecting break-points for curves and break-curves for surfaces is a major problem in shape segmentation. Region merging is prevented across lines of high curvature that are computed first using a separate approach. The problem is that the surface fitting error is measured at each point, which does not correspond to the RMS (root mean square) error measured over the whole region.

Besl and Jain [1986] proposed a data-driven method to do segmentation using the surface curvature sign primitives composed of eight types based on the signs of the Gaussian curvature and the mean curvature A coarse segmentation is given by the KH sign image, which is refined by using iterative variable order surface fitting in the parallel region growing before the region test and merging. Other surface features, *e g* , flatness, region area *etc*. are obtained at the same time. This method obtains good results with a certain set of data [Besl and Jain 1986],[Besl 1986]. However, as will be discussed in Section 5.5, attention should also be paid to its performance in the general case such as with a "Monkey saddle" surface (see Figure 2.6.4), since the inclusion of cubic bi-polynomials into the set of basis fitting functions is questionable. Another problem is their lack of discussion on scale effects which may lead to the zero thresholds empirically imposed on a single scale no longer being appropriate at many other scales.

Yang [1988] discussed the difficulties in using surface curvatures for the purpose of segmentation and classification in the presence of noise. He proposed a strategy that first draws out the planar patches from the range data image using a planarity measurement, then uses the mean curvature to segment the remained curved surface and does further segmentation, if necessary, using the Gaussian curvature. This strategy succeeds with synthetic data. However, the acknowledged lack of theoretical analysis for deriving zero thresholds for surface curvatures is worrying and should be addressed by testing the strategy with real data.

Snyder and Bilbro [1985] discussed the difficulties of segmentation of range data images. The paper proposed first determining the points of high surface curvature; then performing connected-component analysis on the remaining points to determine regions where the surface curvature changes smoothly; and finally grouping high-curvature (boundary) points with associated regions The contrast between pixels is now introduced into the region-based method. But like all the papers mentioned above, this method works at a single scale thus the fundamental problem of scale effects is ignored.

Among these main published papers, both [Snyder and Bilbro 1985] and [Ponce and Brady 1987] can be thought of as key steps towards the curvature-based method where the contradiction between the edge-based and the region-based methods vanishes. And this is also the viewpoint advocated here that both edge-based and region-based methods can be united under one flag of curvature-based methods

5.2.3. Reconciling the incompatibility between the edge-based and region-based methods

The contradiction between the edge-based method and the region-based method looks irreconcilable. However, there is a counter example.

The edge-based method defines the surface patches by edges, so it requires that all detected edges form a set of closed curves However, in practice, not all the edges detected by an edge-based method can form closed curves Unclosed interior boundaries may appear at the crease of a surface as shown in Figure 5.2.1. They also frequently occur even on a non-folded surface, because of noise experienced during measurement and quantisation. Edge linking techniques, such as heuristic search, dynamic programming, relaxation or curve fitting need to be called on to overcome this difficulty.

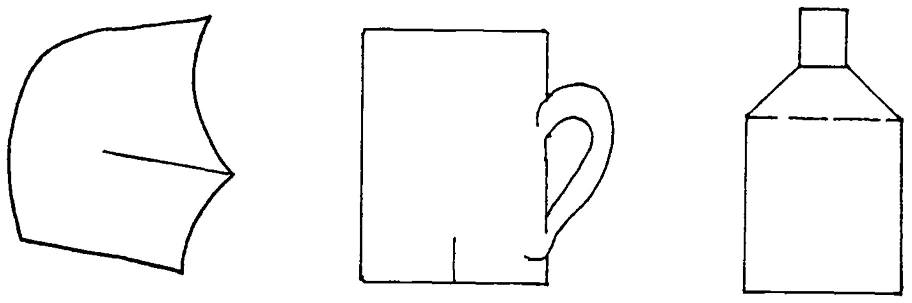


Figure 5 2 1. Unclosed edges detected on the surfaces.

If the edges are defined as the intersections of patches, region-based segmentation might seem to avoid this difficulty Unfortunately, this is not true. Corresponding to an edge in edge-based segmentation, a narrow strip-like region on the surface, which is expected to separate two big regions, may be split into several sections. Large regions are thus joined by these gaps.

This shows that both methods have things in common. Actually, the region-join phenomenon is closely related to the edge-open phenomenon although they may not occur in exactly the same position. Both phenomena can be explained uniformly in terms of surface curvatures. It is evidence in favour of the viewpoint advocated here.

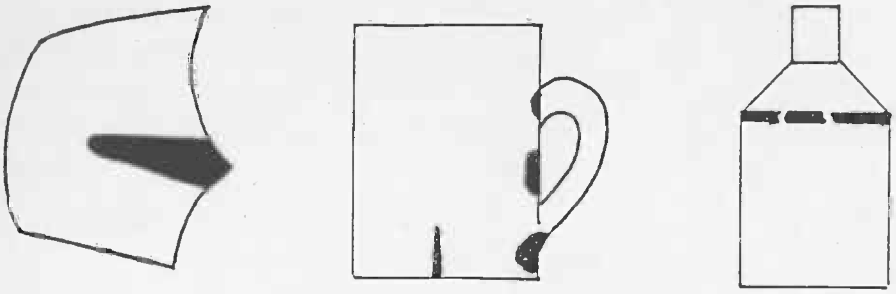


Figure 5.2.2. Different regions join together through the gaps of a split strip region.

5.2.4. Curvature-based segmentation

From the viewpoint of surface curvatures, the contrast between image pixels is due to the difference in signs or magnitudes of surface curvatures at these pixels. When either difference becomes large enough, pixels yield edges at these pixels due to the discontinuity of curvature. Otherwise, they form a surface region due to the similarity in the curvature. So, the edge-based and the region-based methods are complementary to each other like both sides of a coin. The key point is to give a *proper* criterion to judge the strength of difference in curvature signs and magnitudes so as to describe the surface using region separability or region homogeneity. And this is, in nature, a problem of how to approximate surface shape by properly extracting surface curvatures.

A. Region homogeneity and the surface primitives

Since curvature-based methods emphasise the surface properties, they may seem to behave like region-based methods, but it should be remembered that edge detection can also be regarded as a curvature-based problem now.

Region homogeneity requires that every region produced in the partitioning should be perceptibly uniform in a certain sense. So far a variety of surface primitives have been defined to give meaning to this idea.

Marr [1976] proposed the Primal Sketch using zero-crossings, blobs and edges *etc.* as primitives. Asada and Brady [1984] proposed the Curvature Primal Sketch using corners, smooth joins, ends, cranks, bumps and dents as primitives. Ponce and Brady [1987] proposed the Surface Primal Sketch using steps, roofs, smooth joins and shoulders as primitives. Primitives proposed by these edge-based methods will form the borders of homogeneous regions.

On the other hand, Haralick *et al* [1983] proposed the Topographic Primal Sketch using peaks, pits, flats, ridges, ravines (valleys), saddles, convex hillsides, concave hillsides, saddle hillsides, slopes as primitives. And Besl and Jain [1986] proposed Surface Curvature Sign primitives, such as, peaks, pits, ridges, flats, valleys, saddle ridges, minimals and saddle valleys, corresponding to eight possible combinations of the Gaussian and mean curvature signs. Primitives proposed by these region-based methods can directly form homogeneous regions by merging a set of the same primitives.

Between these two types of proposal are Intrinsic patch primitives proposed by [Brady *et al* 1985]. They are actually a mixture of critical surface curves and patches including planar lines of curvatures, planar asymptotes, lines of principal curvatures of constant values or geodesic curvatures of constant values, umbilic patches and planar patches.

Anyway, homogeneous regions are formed by directly or indirectly using these primitives, although the meaning of *homogeneity* may vary from one method to the another. The definition of region homogeneity applied in this thesis is given below, which is based on the Gaussian and mean curvatures. Any break in this homogeneity suggests the presence of a border.

Region homogeneity: *A surface patch is a uniform region when it is connected and invariant in the (K,H) sign category, where K is the Gaussian curvature and H the mean curvature of each pixel in the region*

Using curvatures to define the homogeneity is fairly reasonable since it relates the shape perception to the intrinsic properties of surfaces, which are invariant to rotation and translation. It is also quite easy to handle since the Gaussian curvature and the mean curvature can be computed directly from the given data. Hence, the method of using curvatures as features means it can be applied to a broad range of surfaces and at multiple scale levels. It also leads to a data-driven algorithm, which, however, has not yet lost all the features of a model-driven algorithm owing to using surface classification

B. A data-driven algorithm retaining some features of the model-based algorithm

Model-based[1] processing is widely used in high-level vision, where cylinders, cones and spheres, *etc.* are traditionally adopted as the primitives So, using a model-driven segmentation in the earlier vision stages may provide consistency and simplicity to the later work. This reason is particularly relevant to computer aided design or computer aided manufacture, where many manufactured surfaces are simple, regular forms.

However, there are many free-form surfaces, such as sculptured surfaces, which are difficult to represent with the above models. Using a data-driven approach allows more flexibility and may be better. The curvature-based method proposed in this thesis is such a data-driven algorithm. But it should be remembered that a data-driven algorithm based on the *(K,H)* sign category actually uses eight surface types as models to replace those conventional surface models, such as plane, cylinder, cone and sphere, *etc.* So, this data-driven algorithm still retains some features of a model-driven method. It will be significant and helpful for getting a symbolic description after the segmentation and for recognising objects in high level visual processing.

C. Surface curvatures at different scale levels

The results of segmentation change across spatial scale levels. There are two well-known ways to explore the behaviour of segmented surfaces. One way is to derive fine-to-large scale descriptions by performing additional smoothing at larger scales. Marr and Hildreth [1980], Canny [1983], Witkin [1983], Ponce and Brady [1987] have worked along this direction.

The other way is to use the methods of regularisation theory [Poggio *et al* 1984][Torre and Poggio 1984][Blake and Zisserman 1986]. For example, using the regularisation theory, searching for the optimal smoothing operator can be formulated as the following problem:

---

[1] The term "model-based" here is restricted to algorithms based on identifying *parts* that make up specific objects rather than a model of the entire object

$$\sum_{i=1}^{n}(I(x_i)-S(x_i))^2 + \lambda\int |S''(x)|^2 dx \rightarrow min.$$

where $I(x_i)$ denotes the raw data curve and $S(x_i)$ a smoothed version of the curve. The parameter λ controls the trade-off between the smooth measure and the approximation measure. $S(x)$ closely approximates $I(x)$ if λ is small and $S(x)$ becomes increasingly more smooth when λ increases. So λ is the scale at which edges are detected. Torre and Poggio [1984] further proved that the solution of this problem is equivalent to the image convolution with a cubic spline which is very similar to the Gaussian. More general conclusions are drawn in [Bertero *et al* 1987]

As a stringent surface property which changes gradually in the scale space, and also as a quantity of smooth measure related to 2nd order derivatives, curvature is very suitable to be chosen as the indication of how the segmentation varies across different smoothing scale levels or different λ levels. Besides, it was proven in Chapter 3 that spline smoothing is a special case of repeated averaging or explicit diffusion smoothing. So descriptions of segmented surfaces based on diffusion smoothing would be consistent with those based on regularisation theory. This is further evidence in favour of segmenting surfaces with a curvature-based method

5.2.5. Summary

In this section, edge-based and region-based surface segmentation methods and the contradictions between them have been reviewed. Understanding these methods from a *uniform* viewpoint of surface curvatures has been emphasised. Hence, a curvature-based segmentation method is preferred, which approximates the surface shape by properly extracting surface curvatures at multiple scales. Such a method will use a data-driven algorithm and use the KH sign image as the chief tool in accordance with the region homogeneity. But it will not exclude other means as long as they are curvature-related.

5.3. Roles of surface curvatures in segmentation

The Gaussian curvature K and the mean curvature H, or the principal curvatures C_1 and C_2 are stringent properties of surfaces. They play a fundamental role in surface segmentation. Either the (K,H) or (C_1, C_2) pair may be used to extract shape features from surfaces. Based on these features, surfaces can be partitioned and described.

In this section three problems that are related closely to surface curvatures will be discussed. They are:

1) how to get a better quality (K,H) sign image?
2) why it is possible to partition a sculptured surface with the KH sign image from diffusion smoothing alone? and
3) why surface fitting is indispensable for Besl's segmentation method.

Before starting the discussion, it should be pointed out that using depth data alone to produce the KH sign image of a surface is mathematically an under-determined problem.

5.3.1. An under-determined problem

To segment a surface in accordance with the region homogeneity needs a KH sign image of good quality. As the calculation of surface curvatures involves the 2nd order derivatives, the surface segmentation can be posed as a surface approximation problem, where an approximate version should be close to the surface not only in the depth but also in the 1st and 2nd order derivatives, *i e* , the approximation is *up to* the degree of curvatures. Conventional methods, such as Hermitian interpolation, can produce such an approximate version which has the same depth values and derivatives to the surface at a series of given nodes. However, they are no use in the present situation because of two reasons First, the noise in the data should be filtered rather than retained in the approximation, so the approximation should be a fitting not an interpolation. Second, there is only data for depth and nothing for derivatives at all, so it lacks constraints of derivatives to approximate surfaces up to the degree of curvature signs using the depth data alone. Hence, solutions of this problem will exist, may be stable but usually not unique. In other words, it is mathematically an under-determined problem

As will be addressed, this difficult situation can be improved by using some surface shape preserving treatments And feasible solutions of the problem can be obtained by imposing restrictions on the solution space One way to do so is to directly search for a type of surface in the solution space with a restriction that whose curvature signs can be automatically approximated when depth data is approximated using diffusion smoothing. Another way is to impose restrictions to the set of basis functions when depth data is approximated using a piecewise recursive surface fitting.

5.3.2. How to get a better quality curvature sign images?

The quality of the curvature sign image can be improved in many ways, such as using surface shape preservation in computation of derivatives, in surface smoothing and in surface fitting, or using proper zero thresholding of curvature, *etc*. In this research five treatments are involved.

A. Shape preservation using the spline smoothing method

As the computation of numerical derivatives is sensitive to measurement noise, quantisation error and computational error, both noise suppression and stable computing are required.

Owing to the properties of the bi-cubic B-spline functions [Ahlberg *et al* 1967][Reinsch 1967][Li and Huang 1978], the spline surface approximation is adopted to calculate the derivatives. This processing will be applied to any surface version regardless of whether it has been filtered by diffusion smoothing or not, and has the following characteristics: (1) The surface approximation is of C^2 continuity and $O(h^2)$ accuracy. (2) The local convexity/concavity of surface is preserved. (3) The computation can be done locally and in parallel within 3×3 windows.

Once the spline approximation has been obtained, all derivatives are calculated from the spline approximation rather than by direct numerical differentiation of the input surface data. Better results are obtained in the presence of noise. In particular, the computation of the signs of curvatures is improved.

Indeed, as shown in Chapter 3, spline smoothing will introduce a "convex corner cling" effect at the boundary. This is a weakness when spline functions are adopted to repeatedly implement scale space smoothing. However, this effect will have little influence on the result now because the derivatives are to be computed by a *single* spline smoothing. It may even disappear completely since "small leakage" diffusion smoothing will have successively supplied the subsequent spline smoothing with a proper shape tendency along the boundary. Hence, using the spline approximation to calculate the derivatives of the sculptured surface is reasonable in theory and feasible in practice.

B. Shape preservation using the "small leakage" diffusion model

Ponce and Brady [1987] observed that during scale space smoothing, an object's surface, *e.g.*, an oil bottle, will gradually "melt" into the background.

In scale space smoothing, the depth of a surface changes and the global shape becomes flatter, which leads to the convexity and concavity of the surface changing. Such a distortion must introduce a narrow strip along the exterior boundary of the surface in its KH sign image. So how to preserve the shape tendency along the boundary is an essential problem in scale space processing.

A "small leakage" model in multi-scale diffusion smoothing in [Cai 1988a,b] can be used as a means to maintain the convexity or concavity at most of the boundary pixels as described in Chapter 3. Surface smoothing with this model is, in fact, an approximation to depth data with curvature sign preservation at the surface boundary.

This model also ensures an unconditionally stable computing where computational errors over the surface will be reduced automatically in the process. With this model the approximation of sculptured surfaces can be up to the degree of curvature signs using depth data alone, as will be further addressed in subsection 5.3 3.

C. Shape preservation using piecewise surface fitting up to curvature signs

Curvatures K and H are local properties of surfaces Using them to describe a surface will proceed under the fundamental assumption of differential geometry that the surface is a graph surface or a smooth surface. A noise corrupted surface no longer satisfies this assumption. As a result the "odd texture" of the oil bottle occurs due to the 2nd order derivatives' sensitivity to noise, and it cannot be easily removed by a *local* smoothing operation as shown by the "convexity remaining" theorem in Chapter 4.

In this situation global operations such as surface fitting are useful to model the surface shape. The shape preservation thus becomes the surface fitting problem, and it is an under-determined problem as has been pointed out earlier. Here, the imposition of restrictions to the solution space can be achieved by choosing a set of polynomial basis functions. These functions, associated with a recursive piecewise fitting procedure, will ensure the surface approximation being not only to the depth but also to the derivatives signs (or the KH sign). It thus ensures obtaining a "clean" KH sign image where each region holds a homogeneous KH sign, as will be addressed in the next Section 5.4.

D. Formulated zero thresholding to Gaussian and mean curvatures

Even after surface curvatures have been calculated, zero thresholding of the Gaussian curvature and the mean curvature is still a difficult problem that is not avoided in generating a KH sign image. For a single scale segmentation method the zero thresholds can be determined by trial and error, and be empirically imposed during processing. However, if the thresholds have to be obtained in such a way, the method will not be a non-supervised method, and the threshold values have to vary from one object to another and depend on individual judgements as to which value is acceptable.

Until now, it has not been established how to thoroughly solve or avoid this problem. But by an analysis of the stability of the (K,H) sign category in Chapter 4, the relationship between the zero thresholds of the Gaussian curvature and the mean curvature in an estimation has been formulated. By this formula the zero threshold of the Gaussian curvature can be automatically determined when the zero threshold of the mean curvature is given. It applies to different scales and different surfaces, and contributes to producing a good quality KH sign image.

E. Noise reduction using a pair of quasi-invertible operations

Small and unsignificant noise spots scattering on the KH sign image can be removed by the erosion-expansion operation based on the idea of using a pair of quasi-invertible operations to form an irreversible processing cycle. This is not a new idea, but it is a simple and efficient one.

Applying a pair of quasi-invertible operations, *e g* , the Fourier transformation F and the inverse Fourier transformation G, to a functional x may produce the result $G \circ Fx = x$, *i e.*, $G \circ F = I$, when the functional x is absolutely integrable and satisfies the Dirichlet condition over the integral interval [Goodman 1968][Lighthill 1960]. In this case, F and G form a *reversible cycle*, after which the functional x remains the same, but complex calculations could be simplified. Another example is to calculate numerical differentiation via numerical integration to avoid numerical instability [Li and Huang 1978].

However, a pair of quasi-invertible operations will not form a reversible cycle if the functional x fails to satisfy some certain conditions. For example, given a 10×10 pixel region, a 3×3 pixel region and a 2×2 pixel region, applying an 4-direction erosion F followed by a 4-direction expansion G to these regions produces a 10×10 pixel region, a 3×3 pixel region and a blank region respectively. For the first and the second regions, which remain the same as the original ones, the erosion F and the expansion G form a reversible cycle as the regions have large enough size. In the contrast, the third region disappears after the operation F as it does not have a large enough size, thus nothing is left for the expansion G. Hence, F and G form an *irreversible cycle.* $G \circ F \neq I$ *This shows that whether or not a pair of quasi-invertible operations F and G can form a reversible cycle is closely related to the attributes of the functional x they apply to*

In computer vision real objects are usually given in discrete data corrupted by noise. With a proper ratio of signal to noise, the idea of this irreversible cycle $G \circ F \neq I$ is widely used to suppress the noise in images, to smooth the boundaries among different regions and even be possible to reconstruct a surface by scale space smoothing followed by deblurring using a regularisation technique. The experimental results in Figure 5.3.1 shows that the irreversible erosion-expansion cycle can efficiently remove those isolated small spots of noise from the KH sign image and raise the quality of the KH sign image, where the left half is for a light bulb and the right is for an oil bottle. Among the three images of a light bulb, the leftmost one is the reflection (cosine-shading) image of the raw data, the middle one is the KH sign image and the remaining one is the KH sign image after erosion-expansion operations. First, every region (of different shape type) is shrunk one pixel along the boundary, then all surviving regions are expanded outwards till the gaps among them are filled out Apparently isolated small spots disappear and larger neighbouring regions take their places. A large region (of the spherical shape) appears as the dominant region on the bulb part and saddle shape regions dominate the neck part. These improved the image quality quite a lot.

In short, the above five treatments **A** to **E** can be used in scale space processing to preserve surface shapes, improve calculation of derivatives and reduce computational errors. These lead to producing a better KH sign image, with which a sculptured surface can even be segmented directly.

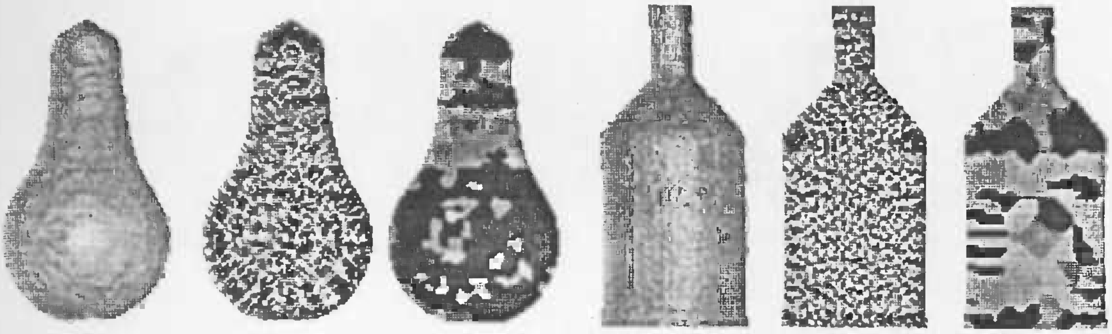


Figure 5.3.1. Isolated small spots of noise can be removed from the KH sign image by an irreversible operation cycle.

5.3.3. Why a sculptured surface can be directly segmented with the KH sign image from diffusion smoothing alone?

In Chapter 4, the stability of (K,H) sign categories and the scale space behaviour of small features were discussed. It has been argued that the KH sign image can be relied on in detecting a C^2 discontinuity at the join points of convex and concave segments, and these join points can be located very well. It is probably the most significant advantage of the KH sign image as a tool in detecting surface shape features. According to Canny's criteria for a good edge detector (cf. [Canny 1983]): 1) a good detection ability, 2) a good localisation ability, and 3) the uniqueness of the edge detected, a good KH sign image can be a powerful tool in detecting a C^2 orientation discontinuity on surfaces. Indeed, this performance can be maintained over a large range in scale space as pointed out in subsection 4,5.1.

Hence the KH sign image is particularly suitable for segmenting those curved surfaces which are rich in the alternation of large convex and concave patches, as happens, for example, with sculptured surfaces. Cai [1988b] argued that a sculptured surface, which is typically a free-form surface and therefore complex in description, can be approximated directly by diffusion smoothing and segmented by the KH sign image. Experimental results of some objects are illustrated in Figure 5.3.2 through 5.3.7. They are a torus, a drill, an egg on a cup, a Renault car part, a human face and a light bulb. Each surface is segmented *only* using the KH sign image resulting from scale space smoothing without any operation relevant to region growing. These surfaces are partitioned into patches as expected. Meaningful surface features are stable over a large range of scales thus are convenient for model matching. For example, the segmentation result of the light bulb is obviously better than Besl's where the surface of the light bulb appears as a whole patch! (An explanation for his result will be given later on.) It is possible that other segmentation methods can obtain results that are as good as this, but it is likely that they will pay a higher cost in computation, whereas applying the KH sign image for segmentation is quite simple and straightforward. In fact, surface fitting and region growing have been applied as the post-processing to these segmentation results, but it is hard to find any significant improvement, as shown in Figures 5.3.6 and 5.3.7.

Figure 5.3.2. Segmentation of a torus (synthetic data) using only the KH sign images resulting from scale space smoothing (scale t = 6).

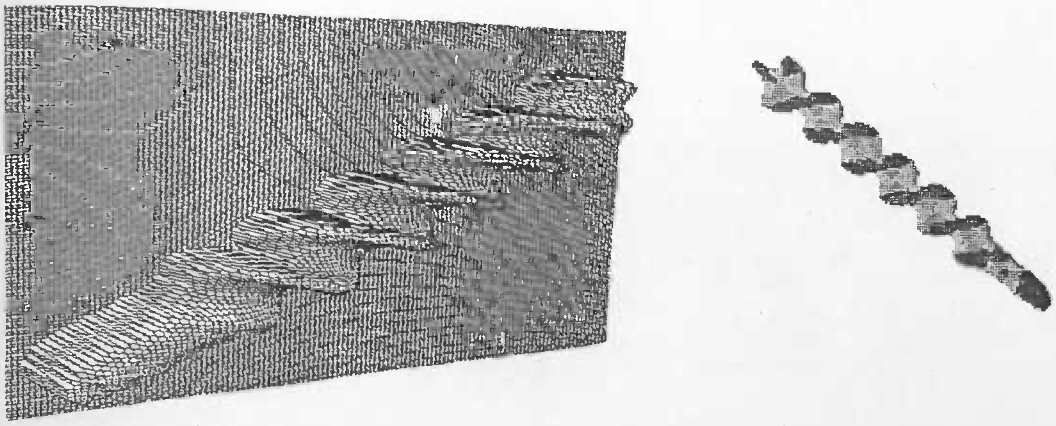


Figure 5.3.3. Segmentation of a drill using only the KH sign images resulting from scale space smoothing (scale t = 16).

Figure 5.3.4. Segmentation of an egg on a cup using only the KH sign images resulting from scale space smoothing (scale t = 25).

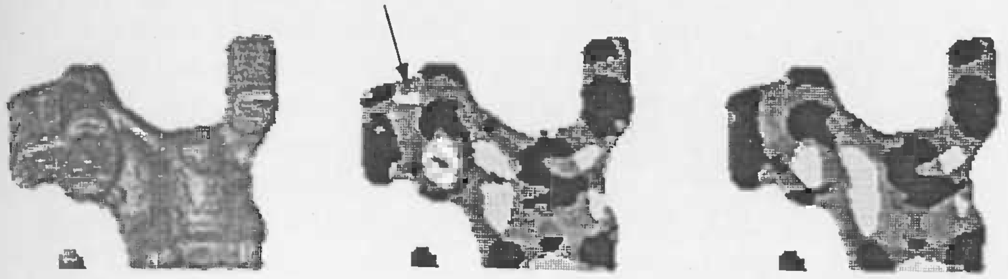


Figure 5.3.5. Segmentation of a Renault part using only the KH sign image resulting from scale space smoothing (at the scale t = 9, 16). Note those small patches around the small pit patch that is pointed to by an arrow, they will be seen as the same saddle surface type when using (C_1, C_2) category; but they can be distinguished as the saddle ridges and saddle valleys when using the (K, H) sign category. This will be significant for object recognition.

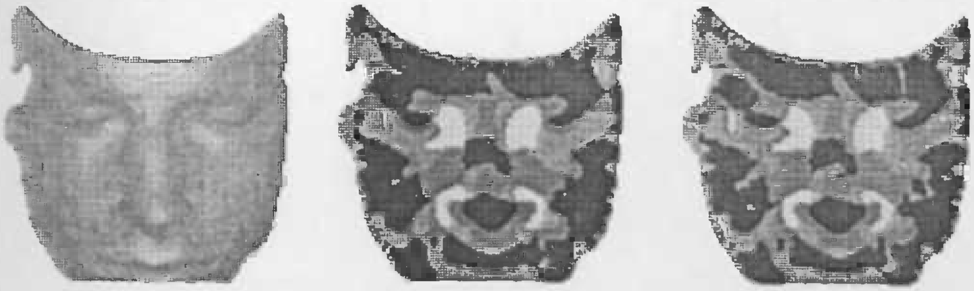


Figure 5.3.6. Segmentation of a human face using only the KH sign image (middle) vs. an additional region fitting/growing (right) (scale t = 25).

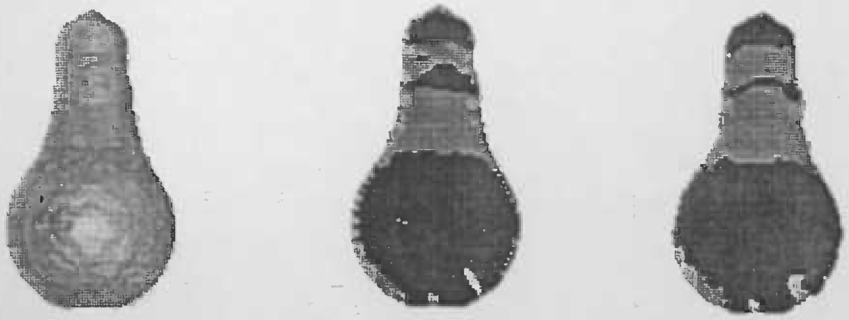


Figure 5.3.7. Segmentation of a light bulb using only the KH sign image (middle) vs. an additional region fitting/growing (right) (scale t = 25).

It is interesting and paradoxical that a human face, a complex generic surface can be easily segmented. However, it should be remembered that skin has some elasticity and its shape is constrained by two factors, one is the shape of the bones it covers, the other is the elasticity of the skin itself. To avoid stress concentration, sharp edges rarely appear on the bone surfaces or at their connections. Hence, the elasticity of the skin makes the skin surface take the shapes similar to those bone patches which have peak or pit shapes, and take a variety of saddle shapes that follow the smooth joints joining these peaks and pits. For this reason, a sharp edge rarely appears on the skin surface. As a result the skin surface is a highly smooth surface and is able to satisfy the smoothness assumption of differential geometry. Typically, only peak, pit and saddles appear on the skin surface. The shapes of these surface types can be preserved in diffusion smoothing and their smooth joints can be maintained as well The good performance of the KH sign image in segmenting the human face is therefore quite reasonable. Similarly, the light bulb is composed of a conic top, a short cylindrical stem, a saddle neck and a spherical body with all of them being smoothly joined to each other After several smoothing operations only the peak, pit and saddle shapes remain and satisfy the region homogeneity. Hence the surface can be easily segmented using only the KH sign image resulting from diffusion smoothing.

In other words, sculptured surfaces composed of peak, pit and saddle patches are surface types whose curvature signs can be automatically approximated when the depth is approximated by the diffusion smoothing in scale space. Thus they are surface types being searched for as the desired solutions of the under-determined problem raised in subsection 5.3.1.

When a segmentation method is unable to preserve the surface shape, the distortion will propagate from the border to the central part as the scale increases. It is difficult to segment the surface with such an unqualified KH sign image alone. Problems also arise when the smoothness assumption of differential geometry is broken. One problem is an occasional failure to be able to detect a roof edge on the surface due to the scale effects as well as the corner-rounding effects (cf. Appendix 2). Information about high curvature magnitude points from other sources is helpful in determining the roof edge position so as to set up curvature-based descriptions in individual regions separated by the roof edge Another problem is the difficulty met in the detection of cylindrical patches due to the "odd texture" effect. Surface fitting is thus used as a tool to even out such noise perturbation and retain some global smoothness for the planar and cylindrical patches as [Besl and Jain 1986] and [Besl 1986] did.

5.3.4. Why surface fitting is an indispensable global operation for Besl's method?

Besl's segmentation method succeeds in certain cases. However it has some problematic features: 1) It works on a single scale, so it is impossible to track the movement of features in scale space to retain some meaningful information. 2) It uses Gaussian smoothing to reduce noise. Thus it fails to preserve surface shape along the boundary. As a consequence, surface boundaries are distorted and sharp edges inside the surface are blurred. 3) Many depth images are used as test data, but they are mainly polyhedral or cylindrical patches, whose KH sign images are sensitive to noise as shown in the last Chapter. As a result, the "odd texture" effect must appear on the cylindrical patches. All these suggest that the KH sign image produced by his method can only be used as a coarse initial segmentation of the surface even when the surface is a sculptured surface like a human face.

Probably, due to lack of information to correct surface distortion, Besl adopted a strategy that first erodes a patch in the KH sign image to a small core part to remove the distortion as well as noise effect so as to get a small "reliable" seed, then applies region growing to the seed to get a uniform patch. To guide region growing in this situation, the only "reliable" thus indispensable tool is surface fitting, where significant emphasis is given to depth approximation rather than curvature signs approximation.

5.4. Fitting surfaces up to the degree of curvature signs

The KH sign image in Besl's method only gives a coarse segmentation of the surface. He did not intend to partition a surface directly with the KH sign image alone. Instead of improving the quality of the KH sign image, he used surface fitting to improve the coarse segmentation. But a correct result cannot be guaranteed by the fitting method he used. For example, his segmentation of a light bulb is a poor result, where the bulb appears as a whole region. Using the same fitting method to segment the "monkey saddle" behaves similarly (cf. Figure 2.6.4).

Yang [1988] recognised the difficulty in using the KH sign image for the purpose of surface classification and segmentation in the presence of noise. He proposed first separating planar patches from the noise-corrupted image, and then using surface curvatures to partition the remaining surface.

However, the difficulty here is not only due to noise. Although local operations, such as Gaussian smoothing, can be used to reduce noise, the smoothing effect propagates locally too. From the "convexity remaining" theorem in Chapter 4, the influence of noise in the curvature signs will persist through many iterations of smoothing. Changing to other smoothing methods makes little difference to this problem so long as the sum of all smoothing weighting coefficients keeps positive (cf. Remark 4.4, 4.5 and 4.8). The final results for the "oil bottle" KH sign image turn out to be similar to each other as shown in Figure 5.4.1. This has been verified by a variety of experiments, even using conservative smoothing[1].

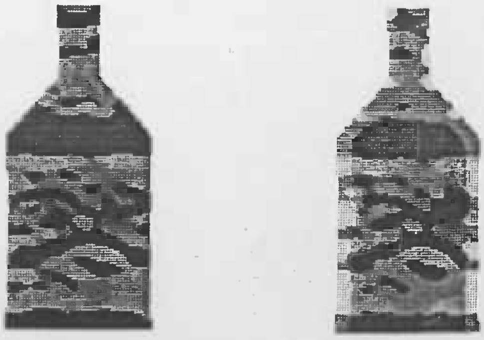


Figure 5.4.1. "Odd textures" in the KH sign images resulting from different smoothings (at the scale $t = 9$) are similar to each other in the central parts.

---

[1] When conservative smoothing is applied to a pixel, if the depth value of this pixel is beyond the passable range determined by both extreme depth values of the neighbouring pixels, it will be replaced by the weighted average of the extreme values near the central pixel value; otherwise, its depth value is maintained.

The reason is that although a pure smoothing can reduce noise in data, it works only by an approximation to surface depth. It can be a simple zero-degree approximation[1]. Such an approximation is not always enough to produce a good KH sign image for a surface (region) where surface shape (curvatures) must be involved in. To handle surface shapes, global information and global operations are useful. Hence, consideration should not only be put on the noise reduction but also be drawn to the fitting method itself.

5.4.1. Surface fitting for region growing vs. KH sign image generation

Surface fitting for region growing and for KH sign image generation are different. Surface fitting for region growing pursues a surface version close to the original surface of depth data within a given error band. It first fits a surface function to the depth data in a local region (called the seed), then expands the seed, under the guidance of this function, till the deviation from the fitting surface cannot be tolerated. The shape of the patch over the expanded seed is implied in the algebraic form of the fitting function. Unfortunately, there is no guarantee of the shape consistency between the expanded patch and the original patch over the seed, since the fitting could be only a zero-degree approximation to the original version. For example, within the error band a seed can be fit by either a planar patch or a spherical patch of large radius, both approximations are available. They are close to each other in depth. However they are different in convexity and lead to different KH sign images before and after the expansion.

On the other hand, surface fitting for KH sign image generation pursues an approximation surface version whose KH sign image should be close to that of the original surface in the presence of noise. That is, both the approximation version and the original version should be close *first* in shape (*i.e.*, convexity) then in depth. This is a higher degree approximation and, as mentioned earlier, it is also an under-determined problem due to lack of constraints to derivatives. So, there might be many unacceptable solutions mixed with the feasible solutions. And efforts should be made to select out those feasible solutions.

Fortunately, the surface shape approximation here does not need a degree as high as to the value of 2nd order derivatives. A weaker approximation, *i.e.*, only an approximation in the signs of the 2nd derivatives (curvatures) is enough. Figure 5.4.2 illustrates an extreme example that a circular curve $f_1(x)$ can be a good approximation to a waving curve $g(x)$ in the depth but it would be a bad approximation in the shape, thus being a bad result for the segmentation of $g(x)$ (at a certain spatial scale); whereas the waving dash curve $f_2(x)$ is a worse approximation of $g(x)$ in the depth, however, it can be a good approximation in the shape since it shares the same curvature signs at each sections, thus being a nice result for the segmentation of $g(x)$.

1: An n-degree approximation means that the approximation surface is close to the given surface up to the degree of n-order derivatives including magnitude and sign at each node, that is, they are not only close in the depth but also in the 1st, 2nd, $\cdots$, till the nth derivatives. For example, given the approximation error ε, if $f(x)$ is a 2-degree (piecewise) approximation of $g(x)$ over interval I, there will be: $|f(x)-g(x)|<\varepsilon$, $|f'(x)-g'(x)|<\varepsilon$, $sgn(f'(x))=sgn(g'(x))$, $|f''(x)-g''(x)|<\varepsilon$ and $sgn(f''(x))=sgn(g''(x))$ $\forall x \in I$.

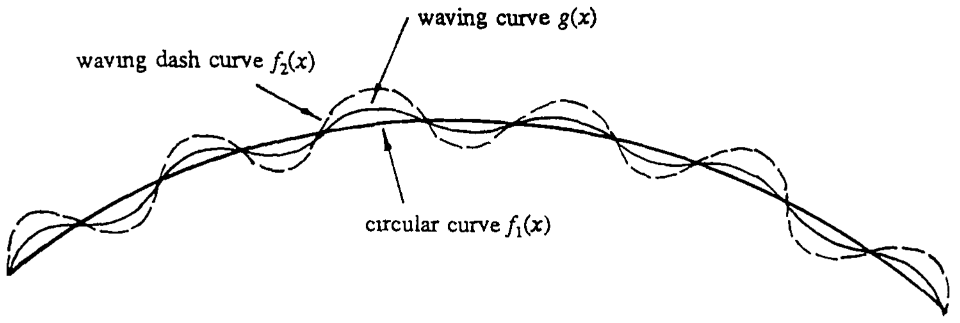


Figure 5.4.2 An extreme example that a better segmentation results from a good approximation in the shape rather than in the depth

Hence, in the case of surface approximation for surface segmentation, larger error of the approximation to the depth will be tolerated as long as the approximations to the signs of K and H are acceptable[1]. Therefore, more attention should be paid to the *curvature signs* (i.e., shape) than to the curvature magnitudes and even the depth. This interesting and important observation gives a clue as to how to obtain a desired solution of the above under-determined fitting problem.

5.4.2. Selection of basis fitting functions

Given a piecewise smooth surface $z(x,y)$ in the form of discrete depth data: $\{z_{i,j}, i,j = 1,2, \ldots, n\}$, the task is to segment it into some patches of uniform K and H signs.

It is plausible to use the depth data alone to fit a surface up to the degree of curvature signs, which produces a good quality KH sign image for segmenting the surface directly. The simplest way to achieve this seems a recursive piecewise surface fitting with some restrictions to the basis functions. And as usual the surface fitting is preferred to be decomposed into a pair of 1-D curve fittings.

Four definitions are given below. Then main results are given in the form of six lemmas, a theorem and two corollaries. Their proofs are in Appendix 4.

Definition 5.1: *Given two continuous curves $f(x)$ and $g(x)$, the distance between their points $f(u)$ and $g(v)$ is measured by*

$$\|f(u) - g(v)\| \equiv \sqrt{|u-v|^2 + |f(u)-g(v)|^2} \qquad (5.4.1)$$

Definition 5.2: *Given a piecewise smooth curve $g(x)$ with its 1st order derivative $g_x(x)$, the arc length of $g(x)$ over the interval $I = [x_0, x_1]$ is measured by the integration with respect to x over I:*

$$S_I(g) \equiv \int_{x_0}^{x_1} \sqrt{1+(g_x)^2}\, dx \qquad (5.4.2)$$

Definition 5.3: *Given two piecewise smooth curves $f(x)$ and $g(x)$, the maximum deviation of $f(x)$ from $g(x)$ over an interval $I = [x_0, x_1]$ is defined as the maximal minimum distance between $f(x)$ and $g(x)$ over the interval I:*

---

[1] That is, given error ε for the approximation of a surface $f(x,y)$ to the surface $g(x,y)$ over the same domain D up to the degree of K and H signs, there should be: $|f-g|<\varepsilon$, $|f_x-g_x|<\varepsilon$, $|f_y-g_y|<\varepsilon$, $sgn(K(f))=sgn(K(g))$ and $sgn(H(f))=sgn(H(g))$ $\forall(x,y) \in D$.

$$D_I(f,g) \equiv \max_{u\in[x_0,x_1]} \min_{v\in[x_0,x_1]} \| f(u)-g(v) \| \qquad (5.4.3)$$

Definition 5.4: *Given two piecewise smooth curves, $f(x)$ and $g(x)$ over an interval $I = [x_0, x_1]$, the significance of the approximation $f(x)$ to $g(x)$ over I is measured by the ratio of the maximum deviation $D_I(f,g)$ to the arc length $S_I(g)$:*

$$R_I(f,g) = \frac{D_I(f,g)}{S_I(g)} \qquad (5.4.4)$$

The first lemma shows the behaviour of a concave curve in a uniform approximation to a straight line segment. It puts emphasis on the approximation of the slope and the curvature signs rather than on a trivial approximation to the depth. This leads to the strategy of approximating a curve's *curvature sign* by a two-stage processing: first use the polygonal lines to approximate a curve in depth, then use parabolic curves to correct the polygonal approximation in order to approximate the degree of curvature sign. This will be addressed in Lemma 5.2 and Lemma 5.3.

Lemma 5.1: *(Existence:) Given a straight line segment $g(x)=d+kx$, there must exist a smooth and concave (or convex) curve $f(x)$ passing through both ends of the same interval $I = [x_0, x_1] \subset [0, \infty)$ $f(x_0)\equiv g(x_0)$ and $f(x_1)\equiv g(x_1)$ such that when the significance measure $R_I(f,g)\}$ is sufficiently small, not only the concave curve $f(x)$ will be a zero-degree uniform approximation of the straight line $g(x)$ over the interval I but also*

1) $f(x)$ is an approximation of $g(x)$ in both slope and curvature magnitudes;
2) $f(x)$ is an approximation of $g(x)$ in the slope sign if $k\neq 0$, but never in the curvature sign;
3) $f(x)$ is invariant in its slope sign over I if $k\neq 0$ And
4) $f(x)$ can be obtained within a finite number of successive approximations to $g(x)$

Corollary 5.1: *A parabolic curve $a+bx+cx^2$ (a, b, c parameters) can be taken as the concave curve $f(x)$ in Lemma 5.1 When the significance measure over I is sufficient small,*

1) parameter c will be close to zero, parameter b will be close to k, and
2) the parabolic curve's slope sign becomes invariant over I if $k \neq 0$.

Lemma 5.2: *Within a finite number of recursions of polygonal approximation, a smooth concave curve can be piecewise approximated by a set of folded straight lines such that the concave curve and the folded line are close in each section up to the degree of curvature magnitude but not the curvature sign*

Lemma 5.3: *Within a finite number of recursions, a smooth concave curve can be piecewise approximated by a set of smooth parabolic curves such that both concave and parabolic curves are close in every section up to the degree of the curvature sign*

There are two problems about the calculation of surface derivatives: How the 1st and 2nd order derivatives affect the signs of Gaussian and mean curvatures? And what is the condition under which the calculation of surface derivatives can be decomposed into x- and y-directions? The results are concluded in Lemma 5.4 through 5.6 below.

Lemma 5.4: *On a smooth surface, the signs of the Gaussian curvature K explicitly depend only on 2nd order derivatives of the surface whereas the mean curvature H also involves the 1st order derivatives*

Lemma 5.5: *When the parameter of the first fundamental form* $F = 0$ *(cf. (2.2.12)), the surface must be a developable one where* $sgn(K) \equiv 0$ *and* $sgn(H)$ *will be related only to the sign of the 2nd order derivative.*

Lemma 5.6: *On a smooth surface when the parameter of the second fundamental form* $M = 0$ *(cf (2 2 13)), the calculation of derivatives can be decomposed into x- and y- directions*

Based on the above lemmas, a practicable way is given in the following theorem, which shows how to select bi-polynomials as the basis fitting to make the signs of *K* and *H* invariant over each region of the approximation surface.

Theorem 5.1 *(Feasibility) To ensure that within a finite number of steps a recursive piecewise fitting can approximate a surface up to the degree of the signs of K and H,*
1) the maximal order of polynomial basis function should be 2; and
2) the feasible form of the fitting bi-polynomial is

$$z(x,y) = a_0+a_1x+a_2y+a_3x^2+a_4y^2 \tag{5.4.5}$$

Corollary 5.2: *Using the bi-polynomial in (A 4 16) to approximate a smooth surface up to the degree of the signs of K and H, the derivatives can be calculated in x- and y- directions independently, i e , the 2-D surface fitting can be decomposed into 1-D parabolic curve fittings*

The above theorem and its corollary gives the restriction to the basis fitting functions so that a surface can be piecewise approximated by bi-polynomials up to the degree of *K* and *H* signs of the surface in order to determine the KH sign types for all patches. Based on the form (5.4.5) a 2-D surface fitting can also be split into a pair of 1-D curve fittings in the x- and y-directions respectively, as is the calculation of surface derivatives.

5.4.3. Two stages of the recursive piecewise curve fitting

To make a recursive uniform approximation to an arbitrary curve using low order functions restricted as in (5.4.5), it is convenient to choose the recursive endpoint fitting method in [Duda and Hart 1973][Lowe 1987][Rosin and West 1988] as the first stage of the approximation to produce a set of polygonal lines, where the significance measure of the polygonal line to the curve is used to judge the goodness of the uniform convergence and to determine the best fit. In this stage, an arbitrary curve is approximated by a polygonal curve (and a surface is approximated by a set of planar facets). Meanwhile, noise is reduced by extracting as long straight lines as possible. Unfortunately, a long smooth curve, *e.g.*, a long conic arc, will be approximated by polygonal lines too. Of course, polygonal lines can be a good approximation to a curve in the depth, even in the slope (as shown in Lemma 5.1). However, it cannot approximate the curvature sign of a curve unless the approximated curve itself is a polygonal curve. For this reason, it is necessary to retain some information which will be useful to improve the result in order to approximate the curve not only in depth and slope but also in curvature sign.

The second stage is therefore a parabolic curve fit to correct, where appropriate, the polygonal approximation yielded by the first stage. A parabolic test curve is produced over each segment. The significance measure of the parabolic segment is used to determine whether a replacement of the line segment by a parabolic segment is appropriate. In this way, the fitting to a conic arc can be corrected by a set of connected parabolic segments, where the approximation is now up to the degree of the curvature sign. Hence, the final result of the second stage will be a set of connected line segments and parabolic segments giving partial derivatives along the x-direction. Partial derivatives in the y-direction can also be obtained after a similar two-stage processing. Based on these derivatives a KH sign image is produced. In this way a given surface $z(x,y)$ is approximated by a set of patches, resulting from fittings in both x- and y-directions, in the forms of planar and quadric patches, *e.g.*, plane, cylinder, cone, saddle, peak or pit, *etc.*

5.4.4. Implementation of recursive piecewise curve fitting

Suppose a curve $C_j = z_{1_j}z_{2_j}$ $z_{n-1_j}z_{n_j}$ is given as the j-th cross section of the surface S along the x-direction as shown in Figure 5.4.3. Where (for convenience the subscript j is omitted,) z_1 and z_n are the end points of the curve, and $z_2, z_3, \ldots, z_{n-1}$ are intermediate points.

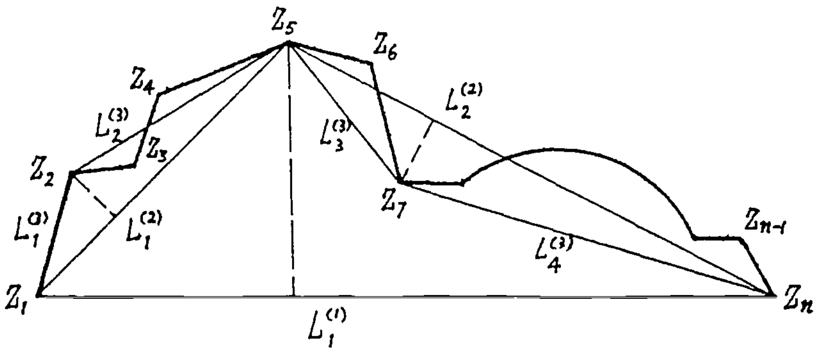


Figure 5.4.3 Recursive end point fitting to the given curve C.

The curve C is first fitted by a straight line segment $L_1^{(1)}$ passing through end points z_1 and z_n. Then C is split at an intermediate point, say, z_5, which has the maximum deviation from $L_1^{(1)}$, and is fitted by two line segments $L_1^{(2)} = z_1z_5$ and $L_2^{(2)} = z_5z_n$. The curve is to be further split at new maximum deviation points, say, z_2 and z_7 and fitted by more line segments, say, $L_1^{(3)}$, $L_2^{(3)}$, $L_3^{(3)}$ and $L_4^{(3)}$. This binary process is recursively repeated until the deviation is zero or the length of each straight line segment is less than 3, the length of the smallest line segment containing non-zero deviation. The structure of this recursive process is a multiple level binary tree where fittings at one level are a refinement of the higher level. After this the unwinding of the recursion tree starts, the significance measure of the curve segment to the straight line segment is used to judge the goodness of the fitting (which is somewhat similar to the significance measure given by [Rosin and West 1988]). Only when the significance measure value is smaller than that of the high-level's, the fitting at the current level will be retained and passed up to the higher level for

---

1: The significance measure (5.4.4) applies to general case of the approximation between curved curves. In the case of a curved segment approximates a line segment, it degenerates to the significance measure given by [Rosin and West 1988].

comparisons in the next turn. So, after the unwinding the final polygonal approximation will be the most significant fitting to the given curve C.

Suppose the polygonal approximation to an arbitrary curve C is $\{L_{k,k+1}\ k = 1,2,...,m-1\}$ $(m \leq n)$, whose vertices $\{v_k\}$ divide C into segments $\{C_{k,k+1}, k = 1,2,...,m-1\}$. Each curved $C_{k,k+1}$ segment will be fit by a parabolic curve $P_{k,k+1}$ determined by both vertices v_k, v_{k+1} and z_{k_0}, the maximum deviation point of $C_{k,k+1}$ from the line segment $v_k v_{k+1}$, as shown in Figure 5.4 4.

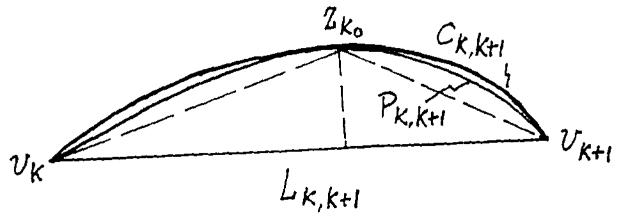


Figure Figure 5 4 4. Using parabolic fitting to correct the result of polygonal fitting.

Then compare the significance measure of $P_{k,k+1}$ to $C_{k,k+1}$ with that of the polygonal line $v_k z_{k_0} v_k$ to $C_{k,k+1}$. If the former measure value is smaller, the line segment $L_{k,k+1}$ is replaced by the parabolic segment $P_{k,k+1}$. So the fitting to a conic arc can be corrected from polygonal segments to a set of connected parabolic segments, where the approximation is now up to the curvature sign.

As shown above, the recursion process of endpoint and parabolic curve fitting will partition an arbitrary curve at the high curvature points, so it virtually gives a segmentation result. And as the significance remains invariant to change of the curve size, this result solely depends on the shape of curve. This is an advantage for curve segmentation.

Certainly, as pointed out in [Duda and Hart 1973] "a single 'wild' point can drastically change the final result". Pre-smoothing is therefore required to reduce noise in discrete data. When smoothed at different scale levels, the shape of the curve (or the surface) will gradually change. The subsequent segmentation that results from the recursive fitting processing will also gradually change. This means that the above requirement can be satisfied automatically and the recursive fitting processing can be embedded in the scale space without inconvenience.

5.4.5. Summary

The discussion on the surface approximation up to the degree of the curvature signs using the depth data alone gives a theoretical analysis to this problem. Although the experimental implementation is still in progress, the theoretical discussion has shown that it is plausible to use a recursive piecewise fitting with a specific set of basis functions to produce a KH sign image from the depth data alone. Such a KH sign image can be directly used to partition surfaces. This gives another feasible solution for the underdetermined surface approximation problem by imposing restrictions on the basis functions of the solution space.

As the surface fitting is useful for dealing with the "odd textures" sticking effect occurring on the KH sign images of cylindrical surfaces, it will increase the power of the KH sign image in surface segmentation. Although this piecewise recursive fitting is particularly suitable for polyhedral and cylindrical surfaces, where straight line and conic curves are elementary curves, it also applies to the segmentation of general surfaces by splitting a 2-D surface fitting into a pair of 1-D curve fittings, where general curves still can be piecewise approximated by straight line and conic curves with the same curvature signs. Thus surfaces can be piecewise approximated by planar and quadric facets with the same KH signs.

As the 1-D piecewise approximation with parabolic modification may produce a better result than the polygonal approximation, it is also an improvement to the method proposed by Pavlidis [1972], where a 2-D problem was also decomposed into 1-D problems so that methods effective in the waveform segmentation, *e g* , the polygonal approximation could be applied.

Besides, the discussion on the selection of basis fitting functions will be helpful for properly treating the surface fitting for region growing as shown in later (cf. Section 5.6).

5.5. Detecting cylindrical patches using principal directions

The piecewise recursive surface fitting in Section 5 4 can produce a feasible solution for surface segmentation. But surface fitting is a *global* operation requiring an intensive computation. It is therefore worth considering other ways to detect a cylindrical surface using non-fitting and curvature-related methods, that is, to detect cylindrical surfaces based on *local* estimation of surface curvatures in the presence of noise.

5.5.1. Extended Gaussian Image

The first possible method is the Gaussian map which transforms each surface normal **n** onto the unit sphere surface S as a radial vector **r**. The Gaussian map is a rotation-independent transformation for a convex object. Horn [1983], Brou [1984], Little [1985] and Ikeuchi [1984] have developed the Extended Gaussian Image (EGI) representation for recognition and attitude determination.

For 2½-D data only a hemisphere surface is involved. Using the 1-1 correspondence between the half sphere surface S and its plate domain D, the top point of each unit radial vector **r** can be projected onto a point p on D then the distribution of p over D can be used to test the shapes of some surfaces as illustrated in Figure 5.5.1.

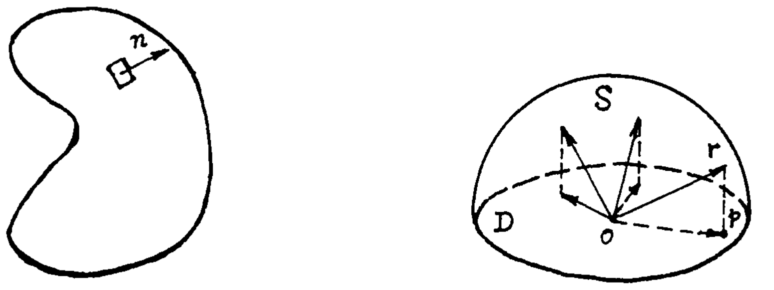


Figure 5 5.1. Correspondence among the unit normal vector **n**, unit radial vector **r** and point p on the plate domain D.

In theory if the symmetry axis is parallel to the plate, the distribution on the plate D will be a straight line for a cylinder, an elliptic arc for a cone surface, a single point for a plane, and a wholly filled plate for an ellipsoid (which is a non-developable surface) as illustrated in Figure 5.5.2.

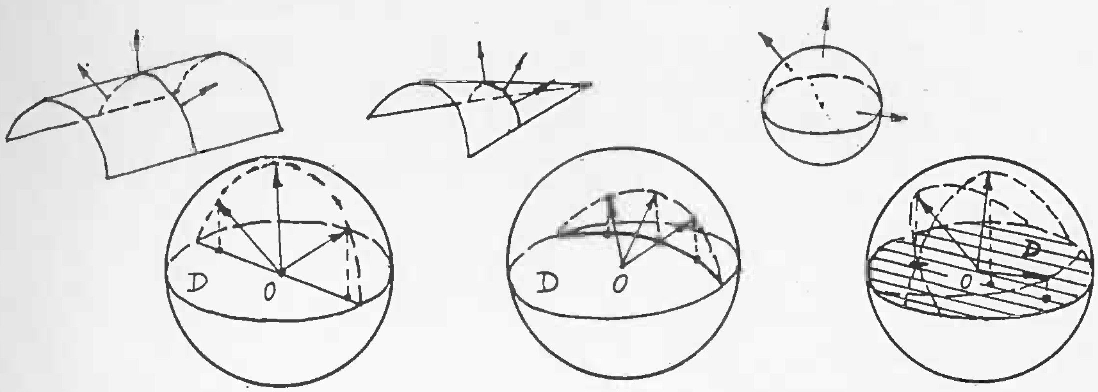


Figure 5.5.2. Cylinder, cone and sphere's Gaussian maps projected on the plate domain respectively.

However, the practical results are not as good as the ideal case. Because the direction of a normal vector is only a local property of a surface, it is not robust to noise. As Ponce and Brady [1987] pointed out that "the EGI is an information-presentation only for complete maps of convex objects, a rare situation in practice". A noise-corrupted cylinder, cone or plane may have a scattered distribution on the plate and a noise-corrupted sphere may have a somewhat sparse distribution as shown in Figure 5.5.3. Usually, the situation is worse for non-convex surfaces. Even in the extended form of the EGI [Ikeuchi 1984], where each discrete view of a non-convex object has its own orientation histogram, the situation is still disappointing.

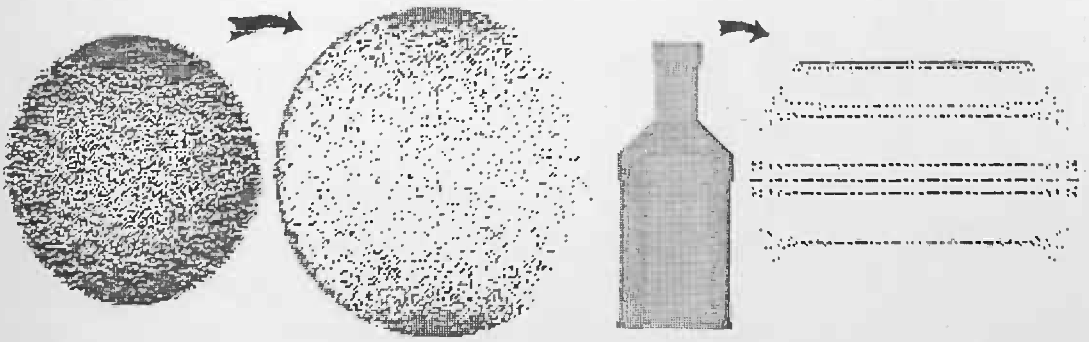


Figure 5.5.3. Sphere and cylinder/cone's Gaussian maps projected onto the plate domain in the presence of noise respectively.

5.5.2. Lines of curvature

The second possible way to detect a developable surface using local estimation of surface curvatures is to use the lines of curvature as proposed in [Brady 1983]. At each point on the surface, there are two principal curvatures C_1 and C_2 associated with their principal directions ϕ_1 and ϕ_2 respectively, where $|C_1|= \min\{|C|\}$ and $|C_2| = \max\{|C|\}$, and the direction ϕ_1 is orthogonal to ϕ_2 on the tangent plane at the surface point.

Once the principal directions at each pixel have been obtained, the curvature primal sketch method [Brady *et al* 1983] links them into the lines of curvature. Since the principal direction is a local vector but the line of curvature is a global description of the surface, it is very useful for representing a surface.

For example, Figure 5.5.4 illustrates that a complete ellipsoid possesses exactly three closed planar lines of curvature. All three lines of curvature are elliptical curves and they uniquely determine the ellipsoid itself in turn. This suggests that a global description may be obtained via a local estimate of the surface properties.

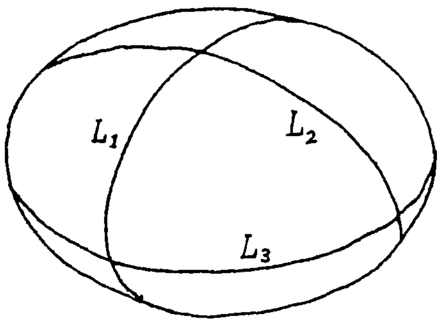


Figure 5.5.4. Three planar lines of curvature L_1, L_2 and L_3 on the ellipsoid.

This method succeeds in producing novel descriptions of a surface in many cases. Nevertheless, some important problems, as pointed out in [Brady *et al* 1983],[Ponce and Brady 1987], are still open to research, because:

1) Lines of curvatures on a smooth surface are not necessarily planar.
2) Lines of curvature are difficult to use to describe an incomplete surface, such as, a partial ellipsoidal patch or a sculptured surface patch.
3) Linking principal directions into lines of curvature is a computationally expensive and tedious process as too much ambiguity has to be cleared up in the processing.
4) The noise in the data is difficult to suppress and aggravates the other problems everywhere.

5.5.3. Principal direction (PD) image

To avoid the above difficulties, it is proposed here that the principal directions will be directly applied instead of having to link them into lines first.

Consider the behaviour of principal directions on some typical surfaces as illustrated in Figure 5.5.5. For a cylinder surface, its principal curvatures are $C_1 \equiv 0$, $C_2 \neq 0$ and the principal direction ϕ_1 is parallel to the symmetry axis s. For a cone surface, its principal curvatures $C_1 \equiv 0$ and $C_2 \neq 0$, and principal directions ϕ_1s at all points are oriented towards and intersect at the top of the cone. For a plane, its curvature is zero everywhere and there is no room for principal direction. And for a sphere (a non-developable surface), curvatures are the same everywhere and every direction can be a principal direction.

Correspondingly, by examining $\mathbf{d}_1^*$ and $\mathbf{s}^*$, (the projections of ϕ_1 and s on the surface domain D,) as illustrated in Figure 5.5.5, it can be seen that the range of the directional angle θ (from $\mathbf{d}_1^*$ to $\mathbf{s}^*$) of a cylinder is quite different from that of a cone as shown in Figure 5.5.6.

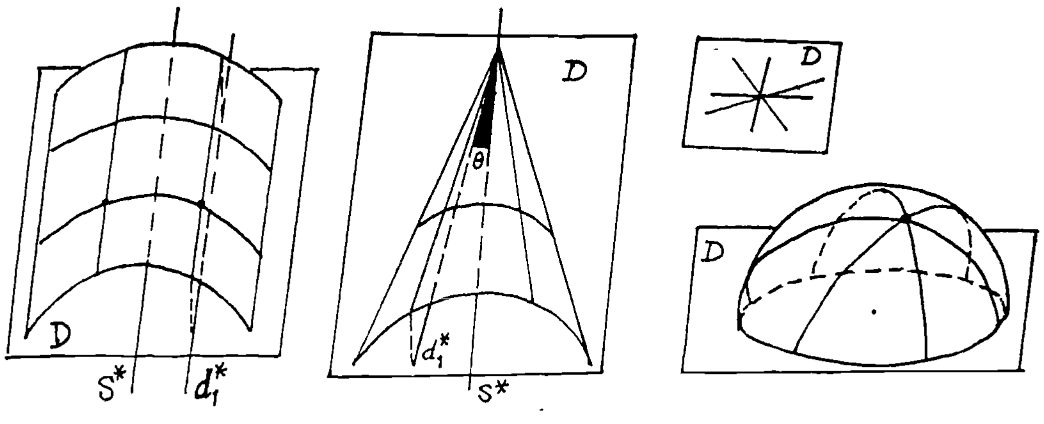


Figure 5.5.5 Principal curvatures/directions of cylinder, cone, plane and sphere.

The angle-distribution of a cylinder concentrates at the zero point of the range whereas the angle distribution for a cone covers a sub-interval from 0 to $\omega/2$ (ω is the conic angle on the axis-section). So a cone can be distinguished from a cylinder using the principal directions. As to umbilic surfaces, such as a plane and or a sphere, they can be identified either by curvature computation or by the principal direction test.

Figure 5.5.6 Distributions of PD angle θ of the cylinder and cone.

According to the above observation on the behaviour of the principal directions, a simple surface category is set up to help in the detection of cylindrical surfaces.

PD category: p–PD type (parallel principal directions, *e g.*, of a cylinder),

u–PD type (umbilic principal directions, *e g* , of a plane or a sphere) and

r–PD type (remaining principal directions, for remaining surfaces).

Associated with a colour denotation for principal directions, the PD category leads a surface to a PD image, where each pixel is dyed in one of three colours respectively:

1) If a pixel is of the p–PD type, where its direction d_1^* coincides with the symmetry axis s^*, the pixel is YELLOW
2) If the pixel is of the u–PD type, the pixel is WHITE.
3) Any other pixel is RED.

So, a sphere is WHITE in colour; a cone surface is RED in colour with a thin YELLOW line along its axis; and a cylinder surface is YELLOW in colour. Each is easily identified from its PD image.

Remark 5.1: *In the definition and denotation of the PD category, nothing is mentioned about the coordinate system. Because all the orientational comparisons are made with respect to the symmetry axis of each (cylindrical or conic) surface itself, the denotation is thus orientation-independent.*

Given a Cartesian coordinate system, the symmetry axis direction of a cylinder at an oblique position can be estimated by averaging all the projection's directions $\{d_1^*\}$ on the cylindrical surface's domain **D**.

In noisy circumstances, the principal directions would experience some deviations from their expected positions. Fortunately, allowing the directions a small perturbation would not destroy the appearance of symmetry [Barlow and Reeves 1979],[Zucker 1983]. When a principal direction falls into the given angular tolerance area, it is seen as coinciding with the symmetry axis.

Note that the tangent principal directions $d_p=\frac{dv}{du}$ $(p=1,2)$ are both roots of the quadric ordinary differential equation:

$$(GM-FN)\left[\frac{dv}{du}\right]^2 + (GL-EN)\,\frac{dv}{du} + (FL-EM) = 0 \tag{5.5.1}$$

The following criteria are given for classifying pixels:

1) If $|d_p|$ falls into $[0,\,\varepsilon]$ or $[\varepsilon^{-1},\,\infty)$, the pixel belongs to p-PD type, and is YELLOW.

2) If $|d_p|$ falls into $(\varepsilon,\,\varepsilon^{-1})$, the pixel is RED.

3) If there exists no $|d_p|$ or $\frac{L}{E}=\frac{M}{F}=\frac{N}{G}$, the pixel is WHITE.

where the angular perturbation threshold ε is a small amount, say, ε = 0.1.

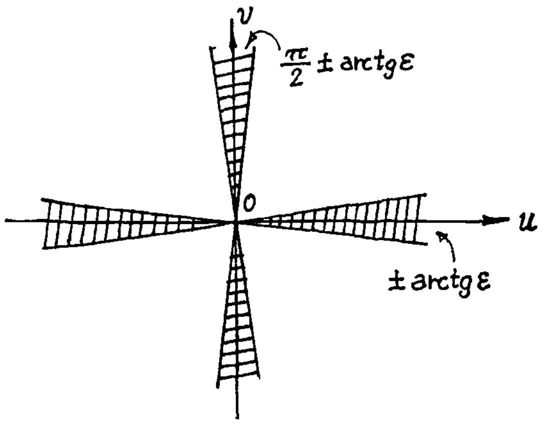


Figure 5.5 7 Division of the u-v plane for the principal direction test.

5.5.4. Surface "roof" detection

When a cylindrical patch adjoins with other type of patches as is the case with the oil bottle, although cylindrical patches can be detected by PD image, the KH sign image is still weak at eliciting the roof edge between the cone shoulder and the cylindrical body, where the surface is continuous but the surface normal is theoretically discontinuous. The main difficulty is that the roof edge will gradually become a broader patch from a narrow strip in the KH sign image when the scale increases. On one hand, this monotonic change provides a strong clue for recognition of roofs, thus the roof eliciting might be left to a recognition stage to solve. On the other hand, as a smoothed roof is still a place of high curvatures, the surface primal sketch method [Ponce and Brady 1987] can be used to detect it.

Ponce and Brady [1987] carefully analysed the scale space behaviour of surface primitives, such as step, roof, shoulder, smooth join and thin bar, *etc.*, and gave some scale-related formulae for the behaviour. For example, the position of the extreme curvature of a "step" is proportional to σ^2:

$$x_\sigma = \frac{\delta}{h}\sigma^2 \qquad (5.5.2)$$

where σ is the scale, $\delta = k_2 - k_1$ is the difference of slope between both sides, h is the "step" height. They also pointed out that when $h=0$, the "step" becomes a "roof", satisfying

$$C(x,\sigma) = \sigma^{-1}C(\frac{x}{\sigma},1) \qquad (5.5.3)$$

This means curvature $C(x,\sigma)$ is a homogeneous function of degree one. In particular, this implies that the extrema value of C is proportional to σ^{-1}, and its position x is proportional to σ, thus the extremum curvature threshold varies proportionally to σ^{-1}.

By referring to this formula, the roof can be detected at different scale σ levels along the maximal principal direction $\mathbf{d}_2$, where the non-maximum suppression in [Canny 1983] can be used to reject the local extrema As for some undesired lines parallel to the axis of a cylinder, which cannot be eliminated by thresholding [Ponce and Brady 1987], they can be rejected by the PD image of the cylinder, where for all pixels $\mathbf{d}_1^*$ coincides with the symmetry axis $\mathbf{s}^*$, thus all pixels are p–PD type and are YELLOW in the KH sign image Hence, using the curvature extrema to detect the "roof" features would be more coherent with a surface curvature-based method with local estimation.

So far it has been shown that the KH sign image is powerful at detecting sculptured surfaces though weak at detecting cylindrical patches without fitting, but the PD image can detect cylindrical patches in the presence of noise, and curvature extremum can be used to detect the "roof" position in scale space. Hence, it is prospective that the combination of information from the KH sign image, the PD image and the "roof" detection can be used to segment general surfaces including generic surfaces and developable surfaces. Surface segmentation in this way is based only on the curvature properties and it works as a local operation at multiple scale levels. Therefore, it forms a curvature-based surface segmentation method in scale space as desired at the beginning of this chapter.

5.6. Comments on the extension of Besl's method from single scale to multi-scale

The discussions in Sections 5.3 through 5 5 have shown the feasibility of the curvature-based surface segmentation in scale space. These discussions are also useful to extend Besl's method (cf. Section 5.2) from single scale to multiple scales.

5.6.1. Essential differences between single and multiple segmentation methods

Indeed, Besl's method (cf. Section 5.2) can produce good results for his test data, particularly, for polyhedral and cylindrical surfaces. However, applying his method at different scale levels will encounter significant difficulties. A multi-scale segmentation method is not a simple copy of a single scale method at each scale. The difference lies in the assumption that the single scale method works at a scale level where noise has almost been eliminated and the surface distortion remains small. This implies that the scale is nearly optimal to identify significant features for the segmentation task.

Unfortunately, this is not always true for multi-scale processing due to the scale effects pointed out in Section 4.1. First, features and their importance change at different scales. Hence, the means of detection of features should be variable over scales to fit this circumstance. Because a single scale segmentation method pays no attention to this phenomenon, it may not retain meaningful information by tracking the movement of significant features across scales whereas the multi-scale segmentation method may For example, zero thresholds of surface curvatures imposed empirically at a specific scale are unlikely to be appropriate at other scales. Second, small features will be filtered along with noise when the scale increases, and some surface distortion will be unavoidable at larger scales. Hence, for example, the same erosion-fitting-growing procedure, which yields a boundary around the roof at a small scale, must fail at large scales where the roof has been flattened a lot. These effects have to be accepted and referred to as the basic facts in scale space processing.

Therefore, an interesting and similar situation met in the previous sections is now faced again in connection with Besl's region fitting/growing method in scale space. In this case also the curvature signs (therefore the surface shape) are more significant than the curvature magnitudes (therefore the depth values). As the curvature sign based segmentation has dealt with the scale effects, it may contribute to the extension of Besl's method from a single scale processing to multi-scale processing. Many changes can be expected, including strategies for noise removal by erosion, noise estimation, region fitting and region growing as well as tactics for the computational efficiency's sake in region searching and filling and so on.

5.6.2. Changing "erosion to core" to "erosion once" to remove small noise spots

A coarse initial segmentation can be produced by the KH sign image, and the segmentation may be improved by the region erosion/growing operation, which needs a set of region seeds. An important point here is how to create these seeds. Besl and Jain [1986] proposed to erode each patch into a small core part to form the seed with respect to a given threshold. Seed are then expanded by using a bi-polynomial least squares fitting. By empirically imposing thresholds on the seed size, fitting error and zero curvatures, this single scale segmentation method obtained good results with a certain set of data. But problems may arise when this "erosion to core" strategy is adopted in scale space.

A. A KH sign image in the form of a mixture of small spots

When the KH sign image appears as a mixture of small spots due to noise, *e.g.*, at scale level σ = o (raw data) or smoothing with fine scale levels, either too many small spots remain or too few large enough spots remain as seeds after thresholding as shown in Figure 5.6.1. Thus the KH sign image is still noisy. The ideas of "erosion to the core" and "biggest size region first" do not work properly there.

The KH sign image may look better as the scale increases to some scale, then the "erosion to the core" and "biggest size region first" strategies begin to work. So, if there were some *a priori* knowledge of the object and the starting scale, the scale space processing might start from this scale level, however, at the price of losing some information about the surface at the finer scales. So, when the data is measured from a completely unknown object, such a trade-off is still difficult to make.

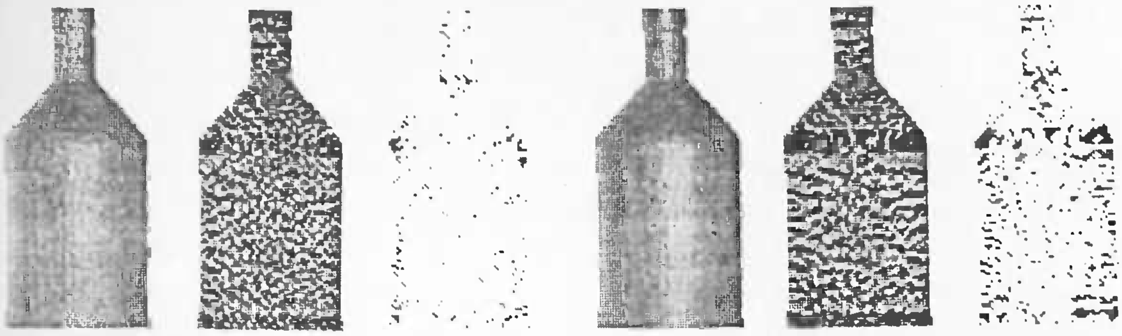


Figure 5.6.1. KH sign image of the oil bottle before and after erosion at a very fine scale level ($\sigma = 0, 1$).

B. Seed is located on a roof patch

For single scale processing small regions across a roof can be cleaned out by erosion, meanwhile other regions on two sides of the roof are either cleaned out as small noise spots or remain as seeds of larger surface regions. When subsequent region growing starts, regions on both sides will be prevented from merging by testing their different orientations. But for scale space processing, the roof part of a surface will gradually expand from a narrow strip to a mild area in nearly convex shape when the smoothing scale increases. The seed of this area usually locates on the centre of the area overlapping the roof. So the erosion-to-core strategy is unlikely to work correctly again.

Recall the argument in subsection 5.3.2.E and note that at fine scale levels many noise spots are very small (such as 2×2 pixel in size), they will be removed at larger scale levels; whereas many signal spots may be only slightly larger than the noise spots, but they will remain at larger scale levels. Therefore to avoid the above difficulty the "erosion-to-core" strategy is replaced with the "erosion-once" strategy (but not the "erosion-twice" strategy, which could also remove some some signal spots in 4×4 pixel size). Consequently, only some small noise spots will be removed by one erosion, the remaining noise spots would be left for diffusion smoothing to remove at larger scale levels and the surface roof distortion would be temporarily accepted as a natural result in scale space processing. As the roof patch will always expand when the smoothing scale increases, it can be recognised by tracking its movement across scale levels, or be found by the "roof" detection due to the high curvatures near the roof.

5.6.3. Changing method to estimate noise in raw data

The threshold for region fitting and growing is closely related to the estimation of data noise. By comparing the raw data and its smoothed version the root mean square deviation or the maximum deviation can be estimated. However, sharp edges of the surface sometimes introduce serious error to the estimation using local planar fitting as Besl did, where he proposed to estimate the root mean square deviation only in the internal area of regions to avoid sharp edges. As the spline smoothing has been used in this thesis to calculated the surface derivatives, it is very convenient to estimate the root mean square error from the spline smoothed version without having to avoid these sharp edges. The influence of deviation at the sharp edges is reduced owing to the spline function's C^2 continuity and $O(h^2)$ accuracy as well as the preservation of local convexity/concavity as mentioned before. It improves the estimation of noise and save computation as well.

5.6.4. Changing basis functions of region fitting/growing

Many model-driven methods, for example, [Hebert and Ponce 1982] have classified surfaces into either plane, cylinder, cone or ellipsoid (sphere) types, thus quadric surfaces are used as basis functions.

Besl's data-driven methods based surface classification on eight types: peak, pit, saddle ridge, minimal, saddle valley, ridge, flat and valley. He treated surface growing as a conventional depth fitting problem, so his bi-polynomial fitting function is up to the 4th order, the highest order accepted by practical computation.

Even though some surface distortion is accepted in the scale space processing, two facts are still essential to surface fitting for growing. First, the region homogeneity should be satisfied in both the surface partitioning and region growing. Second, the basis fitting functions play a fundamental role in the region description and consequently affect the region growing. Thus, to extend Besl's method into the scale space processing the choice of the basis functions should coincide with the requirement of regional homogeneity.

Being faithful to Besl's surface fitting for region growing as a variable order bi-polynomial fitting, the extended version maintains the bi-polynomial fitting function up to the 4th order and still treats the growing as a conventional depth fitting problem. However, note that there is no guarantee of shape consistency in the region extrapolation with Besl's fitting method, efforts in this research are made to grow a region within the seed's shape type. A selection of fitting functions is thus required to reduce the risk while region growing. This selection will refer to the conclusions made for the surface fitting used for KH sign image generation although the restriction can be looser now.

A. Selection of the basis functions

Besl and Jain [1986] were the first to introduce the variable order polynomial fitting to the vision area. In total, they adopted fifteen basis functions and combined them into four nested groups according to the surface orders:

$$1,\, x,\, y,\ /\ xy,\, x^2,\, y^2,\ /\ x^2y,\, xy^2,\, x^3,\, y^3,\ /\ x^3y,\, x^2y^2,\, xy^3,\, x^4,\, y^4 \qquad (5.6.1)$$

This is a familiar technique in the area of numerical computation, where the fitting usually terminates itself using a procedure which selects the optimal fitting order. Besl and Jain, instead of using the optimal selection procedure, used a RMS (root mean square) error selection in fitting when the order ascends and treated the termination in this way: when the surface fitting of a region fails at the highest order, that region loses its independence and will be divided by other regions during their growing. So, Besl and Jain's variable order polynomial fitting is an *open* process.

While the least squares surface fitting may work well in the seed regions, it is still possible to get undesired results from region growing (extrapolation). The "Monkey saddle" surface shown in Figure 2.6 4, which is a bi-cubic function $z = x^3 - 3xy^2$, can be taken as an example.

According to the (K,H) sign classification, the surface should be partitioned into six patches, and the KH sign image is really composed of six patches around an isolated flat type point (0,0). After region erosion, six seeds are produced as expected too. They can be fitted by bi-cubic functions respectively. However, starting from any one of the six seeds, region growing must enter all the other five patches and eventually forms a nearly unique surface. This result would clash with the aim of segmentation. So including bi-cubic functions in the fitting basis functions is risky, it is better to exclude them.

Referring to the results in Section 5.4, excluding bi-cubic polynomial functions from the set of basis fitting functions is allowable. First, it is feasible to use bi-quadric surfaces to piecewise fit the bi-cubic patches to get the desired segmentation results in this case. Second, the emphasis on region fitting has already been changed in scale space processing. Speaking in detail, a single-scale segmentation pursues a depth fitting as close as possible to the object itself at a specific scale no matter what the importance of each feature of the object, whereas a multi-scale segmentation pays more attention to the extraction and tracking of those perceptually significant shape features in the scale space where distortion is unavoidable and permissible. So emphasis is now put on the approximation to surface shape rather than the depth itself, which introduces more flexibility and error-tolerance into the fitting and the growing so long as the recognition will not suffer from them.

Regarding the stability analysis of the KH sign category in shape feature extraction, the situation is similar. As most of the quadric surfaces can maintain their shape tendency (convexity/concavity) in surface extrapolation, *all* bi-quadric polynomials fitting functions can be included in the set of basis functions, as are those 4th order even functions; whereas all the bi-cubic polynomials should be excluded from the set to reduce the risk of getting the wrong shape during surface extrapolation (region growing) and to protect the region's curvature homogeneity from damage. It is interesting that this exclusion coincides with the catastrophe theory [Poston and Stewart 1978], where the bi-cubic polynomials are also classified as a kind of functions of structural instability.

The basis function set adopted by the extension of Besl's method in the multiple scale case is therefore composed of:

$$1,\ x,\ y,\ /\ xy,\ x^2,\ y^2,\ /\ xz,\ yz,\ z^2,\ /\ x^2y^2,\ x^4\ ,y^4. \tag{5.6.2}$$

which introduces xz, yz and z^2 as basis fitting functions for all bi-quadric patches.

Compare with Besl's fitted surface in 3-D space:

$$\begin{aligned} 0 = f(x,y,z) &= w_0 + w_1x + w_2y + w_3xy + w_4x^2 + w_5y^2 + w_6x^2y + w_7xy^2 + w_8x^3 + w_9y^3 + \\ &+ w_{10}x^3y + w_{11}x^2y^2 + w_{12}xy^3 + w_{13}x^4 + w_{14}y^4 - z \end{aligned} \tag{5.6.3}$$

the fitted surface proposed here is of an obviously different form:

$$\begin{aligned} 0 = f(x,y,z) &= w_0 + w_1x + w_2y + w_3xy + w_4x^2 + w_5y^2 + w_6xz + w_7yz + w_8z^2 + w_9x^2y^2 + \\ &+ w_{10}x^4 + w_{11}y^4 - z \end{aligned} \tag{5.6.4}$$

All planar and quadric surfaces can be described with these basis functions. But different patches of odd order surfaces, *e g* , the "monkey saddle", are now unable to grow to encompass the whole surface. Instead, the segmented surface will be approximated by six quadric surface patches.

B. A variable "order", optimal surface fitting

This approach is still a variable order polynomial fitting. Its viewpoint has simply switched from the "order" of the surface to the "size" of the set of basis fitting functions, which gives more flexibility to the fitting without any theoretical difficulty.

Given an error band, the fitting starts with the simplest planar form using three basis functions, 1, x, y and counts the fitting error. Afterwards, more basis functions may join the fitting one by one, in a set-size ascending order along with selection for the best fit.

1) If the current fitting error is below the error band, the fitting succeeds. The next basis function joins the fitting, and its result will be compared with the current result to test whether the current fitting has reached the optimal one. If so, the fitting stops. Otherwise, let the "test fitting" be the current fitting, and introduce a new basis function for another test fitting. Continue this process till an optimal fitting is obtained.
2) For each failed fitting, there is no test for optimisation. A new basis function joins the fitting. Repeat this process till the fitting succeeds.
3) If all fittings fail, select the last fitting as the optimal one, then stop.
4) An optimal fitting should satisfy two conditions: i) It is the first time the current fitting's error is less than the testing fitting error; ii) There is no fitting following the test fitting that can reduce a significant percentage, say, 40% from the current fitting error. Otherwise, it is a local minimum, and from now on the first fitting which can reduce 40% error from the current error will be selected as the optimal one, and the fitting will end (cf. [Clenshaw 1960][Makinson 1967]).

5.6.5. Changes in region growing

A. Dominant colour (shape) first

Before region growing begins, the question arises as to which region should be the first to start. It is hard to say which answer is the best. Besl started from the region of largest size. Referring to the importance of surface shape in segmentation as well as in description (as well be addressed in Chapter 6), the following precedence rules are proposed instead:

Rule 5.1. (Growing precedences):

1) (Dominant shape first:) Among different shape types (colours), regions of the "dominant shape type (colour)" in the KH sign image will first grow.

2) (Dominant size first:) Among different size regions of the same shape type, the largest region will first grow.

The reason for the "dominant shape first" choice is that in the presence of noise, the dominant colour (surface type) seems also the dominant factor in the perception of the global shape of an object. In particular, the KH sign image at a fine scale level is a mixture of many small colour spots due to the noise. This noise effect is something like a veil whose texture separates the face and makes it vague. It is possible that a human can detect the face behind the veil not by the strategy of searching the largest patch which may be a part of the veil, but by the strategy of considering all the separated, small patches of the same colour (or shape) as a whole. The size of any single patch is thus less decisive in this situation.

The reason for the "dominant size first" choice is that after smoothing, many small regions vanish or merge rapidly into big ones. Big regions usually persist over a large range of scales. So the region with a dominant size is relatively important and stable, and should be first to be grown.

B. Updating coefficients in a ring-wise vs. a point-wise fashion

During region growing the basis fitting function set remains invariant. A region expands itself by calculating its fitting function and undergoing growing tests. Any point that joins a growing region will change the fitting coefficients of this region, thus affecting the final result of the growing. The curvature magnitudes of a non-planar region may also gradually change during growing even though its basis function set is invariant. So the fitting coefficients must be updated to fit the changing shape.

There are two updating modes. One mode is to update the coefficients as soon as a candidate point has been accepted. So both the region expansion and the coefficient updating are going on in a single point-wise fashion until the whole growing finishes. The other mode is to expand the region ring-wise, and to repeat it again and again till no more expansion is possible.

The latter mode is chosen for its reliable performance in a relatively small region as well as for its higher computational efficiency in a large region. (cf. the updating of Householder transformation in Appendix 6.)

C. Growing tests

Once the least squares fitting result has been obtained, the region is represented by the basis functions and their fitting coefficients. The region growing goes on with the surface extrapolation using this fitting function The acceptance or rejection of a candidate point will be given after some growing tests are made.

The first one is the connectedness test. A perceptually homogeneous patch is a connected region. As this requirement has been adopted in region erosion and fitting, it should also be retained in region growing to avoid any confusion. The 4-adjacency test checks whether or not a candidate point is connected to the growing region.

It is possible that after occlusion with other surfaces, a surface (*e g* , a sphere surface) remains as two or more separated sub-surfaces. In this case a fitting function built up on one of these sub-surfaces may also be applied to the others This may be an important clue for subsequential geometrical reasoning or surface description, but is not yet a reason for the growing region to accept a point in a disconnected region in the segmentation stage, even though both regions really belong to the same (sphere) surface.

Besl's method allows a disconnected point to join the current growing region as long as it satisfies the fitting function. The viewpoint of this research is different. Every big enough region has its own chance to build up its fitting functions and expand itself. Two regions can be united if they touch in the growing and passed the merge test as proposed in [Besl and Jain 1986]. Early acceptance of a disconnected point into the current growing region seems against intuition. The algorithm and the data structure will become more complicated and the local scannings become global ones, which are much more expensive.

In contrast, the connectedness requirement makes it possible to identify candidate points within a local rectangle box, which gets the initial size from the region searching stage and gradually expands along with the region growing so that it always covers the growing region. As the initial rectangle box is usually smaller than the whole image, such a local operation reduces a lot of the computation.

Other tests such as best fitting test, C^0, C^1 and C^2 tests are discussed in Appendix 5.

5.6.6. Other changes

Another change is the algorithm for locating regions (cf. [Ledley 1964],[Rosenfeld and Kak 1976],[Gonzalez and Wintz 1987]). The wall-following algorithm instead of the region-expansion algorithms is used in region searching, filling and boundary pixel indexing. It reduces both the computational complexity and the memory requirement from $O(l^2)$ to $O(l)$, where l is the perimeter of a region. Also, the updating Householder transformation is used in the least squares fitting. This change is very important to raise the computational efficiency of Besl's method in scale space. However, it is not closely related to the main topic of this section, so it is put into the Appendix 6.

5.7. Summary

In this chapter the theoretical results in previous chapters are applied to the surface segmentation. The curvature-based segmentation method is advocated.

Segmentation based on surface curvatures in terms of the KH sign image is posed as an under-determined mathematical problem of approximating surfaces up to the degree of curvature signs using depth data alone. This shows the direction along which efforts are made to obtain desired solutions of the segmentation problem. Sculptured surfaces have been shown as the surface type whose curvature signs can be automatically approximated when the depth data is approximated by the "small leakage" diffusion smoothing, thus, segmentation can be achieved using the KH sign image resulting from the diffusion smoothing alone. Developable surfaces have been shown to be feasible to segment via a recursive piecewise fitting with restriction to the basis fitting functions in order to ensure the surfaces are approximated up to the curvature signs. Meanwhile, cylindrical surfaces have been shown to be able to detect using the curvature directions directly rather than first linking them into the lines of curvatures. Hence the discussion on the 3-D surface segmentation is virtually around how to reliably extract the curvature features of surface in terms of a good quality KH sign image. To reach this aim, when using smoothing to approximate the surface, shape preservation along the boundary is necessary. When using surface fitting to approximate the surface to extract curvature signs, the function xy and any function whose order is 3rd or higher should be excluded from the basis fitting function set. And when using surface fitting for region fitting/growing, the bi-cubic polynomials should be excluded from the basis function set to reduce the risk of getting undesired shapes in surface extrapolation (region growing). Finally, when the scale varies, the zero thresholds of the Gaussian curvature and the mean curvature should be changed according to the consistent zero thresholding inequality given in the subsection 4.3.4.

Based on these results, general surfaces may be segmented in scale space by using the KH sign image, PD image and roof detection together. This is the main version of the curvature-based segmentation method using local, non-fitting operations. Another feasible version adopts a piecewise, recursive surface fitting associated with a parabolic curve modification, thus uses curvature-related global operations. These results are also helpful in extending Besl's region-based method from single scale level to multiple scale levels.

Chapter 6

Surface shape representation for recognition

6.1. Introduction.

Early visual processing produces a lot of useful information. But simply gathering it in a heap is not yet computer vision To understand surfaces properly, this information must be well organised within a scheme of representation. The most fundamental property of a scheme of representation is that it makes some types of information explicit, and this property can be used to bring the essential information to the foreground allowing smaller and more easily manipulated descriptions to suffice.

Some simple surfaces, such as a cylinder, a polyhedron or a sphere, can be described with graceful mathematical forms However, it is difficult to use mathematical equations to describe some complex surfaces, especially, sculptured or noise-corrupted surfaces. A piecewise and symbolic form is therefore convenient to represent them, *i e*, a surface is decomposed into several meaningful patches using primitives or patterns, then the separated patches are symbolically described and assembled up into a compact version, where more emphasis is given to a subset of patches and the inter-relations amongst them due to their salient roles in subsequent object recognition.

There are several ways of representing surface shapes. Marr [1976] proposed a volumetric representation which describes shapes and their spatial organisation in an object-centred coordinate frame, using a modular hierarchical representation that includes volumetric primitives (*i e*, primitives that represent the volume of space that a shape occupies) as well as surface primitives, and attach them to a few sticks or axes within a spatial configuration. Thus these primitives can be used to represent both a detailed description of part of an object, and a gross description of the whole object.

Haralick and Watson [1983] proposed the *Topographic Primal Sketch (TPS)*, which labels each pixel of the surface with edge (step) label, or one of ten possible topographic labels: peak, pit, flat, ridge, ravine (valley), saddle, convex hillside, concave hillside, saddle hillside and slope TPS quantities are invariant to monotonic gray scale transformations. However, all TPS labels are, in general, not invariant to small changes in viewpoint relative to a surface.

By relying on surface curves to constrain surface shapes, Asada and Brady [1984] proposed the *Curvature Primal Sketch*, which is used for planar curve shape description of surface curves. Ponce and Brady [1987] proposed the *Surface Primal Sketch* presentation, which detects, localises and symbolically describes significant surface discontinuities in scale space. Brady *et al* [1985] proposed an *Intrinsic Patches* representation, which is based on concepts of differential geometry, such as lines of curvatures, asymptotes, bounding contours, surface intersections and planar and spherical (umbilical) surface patches, and provides a hierarchy of increasingly stringent surface descriptions. Together with the *Surface Primal Sketch*, the *Intrinsic Patch* approach aims at such a representation, which, for example, can describe a light bulb as the figure below:

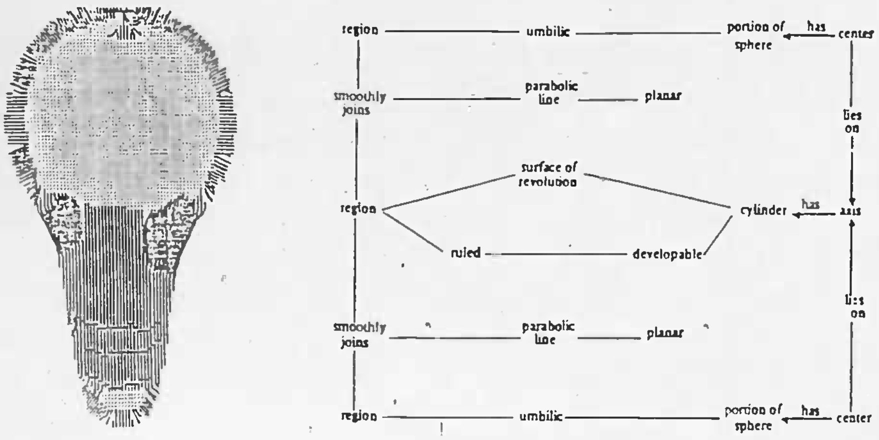


Figure 6.1.1. Representation of a light bulb from Brady.

Even though the scope is limited (as mentioned in Section 5.2) and the emphasis is mainly put on curve information, the *Intrinsic Patch* representation shows well how a representation based on region information can be done.

Medioni and Nevatia [1984] presented space description ideas in the context of differential geometry but the approach limits itself to the zero-crossings of the Gaussian curvature and the maximum principal curvature, and the maxima of the maximum principal curvature. Besl and Jain [1986] proposed a piecewise-smooth surface model for image data that possesses surface coherence properties. The KH sign image provided an initial coarse image segmentation, which was refined by an iterative region growing method based on a variable-order surface fitting. But these methods were not designed to work in scale space.

In this thesis, a surface shape representation for recognition, namely, the *Flexibly Constrained Graph Tower* (FCGT) is proposed, which works in scale space. Scale is important for this scheme of representation because the perception of surface structure is dynamic [Koenderink and Doorn 86], *i.e.*, surface shape will change over different scales, which leads the curvature-based segmentation to change accordingly. To reflect this fact, the representation takes a tower-like form in its architecture. It consists of many layers corresponding to different scale levels. On each layer there is a flexibly constrained graph, which is an adjacency graph describing the topological properties of the segmented surface, constrained partially or wholly by an attribute list describing the geometrical properties of individual surface patches. Between each neighbouring layer linkages are set up to assemble them from a pile of isolated layers into a solid tower as whole.

Marr [1982] addressed three aspects of a representation's design: 1) the representation's coordinate system; 2) its primitives, the primary units of shape information used in the representation; and 3) the organisation that the representation imposes on the information in its description. In terms of these aspects, the FCGT representation will have 1) an object-centred coordinate system; 2) eight primitives for surface patches of different KH signs; and 3) an organisation decoupling the topological properties and the geometrical properties into the forms of an adjacency graph and an attribute list within a multi-layer architecture.

In the FCGT representation, the first two aspects are direct consequences of the curvature-based segmentation using differential geometry. The third aspect is motivated by the fact that topological properties of surfaces are free from the choice of coordinate system. Thus they are less constrained and more general than geometric properties.

Hence they are suitable to form the core part of the representation.

This decoupling of topological and geometrical properties can improve the effectiveness of a shape representation. First, adjacency graphs at different scale levels can be grouped into a family with respect to a single scale parameter. Meaningful changes in surface segmentation prompt changes in the adjacency graph, so the adjacency graphs themselves could represent the *structural* changes of a segmented surface over scale levels. Thus, the decoupling favours the representation's accessibility. Because a relatively simple adjacency graph will simplify the processing of inexact surface matching [Shapiro and Haralick 1981][Eshera and Fu 1986], decoupling also favours the representation's stability.

Second, the adjacency graph of a segmented surface concerns itself with the *existence* rather than with the *content* of the difference between surface patches. Thus, it can cater for some flexibility in the surface This favours the representation's scope as well as its accessibility.

Third, some uncertainty will appear in matching using the adjacency graph alone, but it will be reduced by using constraints from a partial or a fully associated attribute list, and the more constraints the less the uncertainty. This favours the representation's sensitivity as well as its uniqueness and accessibility.

In subsequent sections the organisation and formation of the scale space representation, the flexibly constrained graph tower will be discussed in more detail. Section 6.2 discusses the forms and functions of the adjacency graph Section 6.3 discusses the forms and functions of the attribute list Section 6.4 further supports the separation of topological and geometrical properties of a segmented surface by referring to some observations. Section 6.5 discusses flexibility in the surface representation and specifies the formal expression of the flexibly constrained graph Section 6.6 discusses the extension of Witkin's method into the 3-D case and produces the natural scale description of a surface in the 3-D case. An experimental example will be shown.

6.2. The adjacency graph of a segmented surface.

After segmentation a surface is partitioned into homogeneous regions, where the intersection of regions forms their boundaries. A surface representation is organised to describe the properties of these regions and boundaries. For an edge-based method, boundaries are the main components of the description, where the regions are implicitly represented, a wire-frame model [Porrill *et al* 1988] of an object can illustrate the method. In contrast, for region-based or curvature-based methods, regions are the main components of the description and are explicitly represented. So the representation has to reflect this fact. The dual form of a wire-frame would illustrates this method.

6.2.1. From segmented surface to adjacency graph.

In dual relationships a region corresponds to a point, and a line corresponds to a line. A node will be marked inside each region and each pair of nodes will be linked with one arc when their regions are adjacent. This set of nodes and arcs gives a graph of the segmented surface, where the number of nodes is the same as that of the regions, and the number of arcs is the same as that of the boundaries[1]. This graph is called the

---

[1] If two or more arcs appears between a pair of nodes, they are merged into one. This is somewhat different to the traditional case in Graph theory.

adjacency graph (AG) of the segmented surface since it represents the adjacency and connectivity among all patches on the segmented surface. Figure 6.2.1 illustrates some segmented surface and their adjacency graphs.

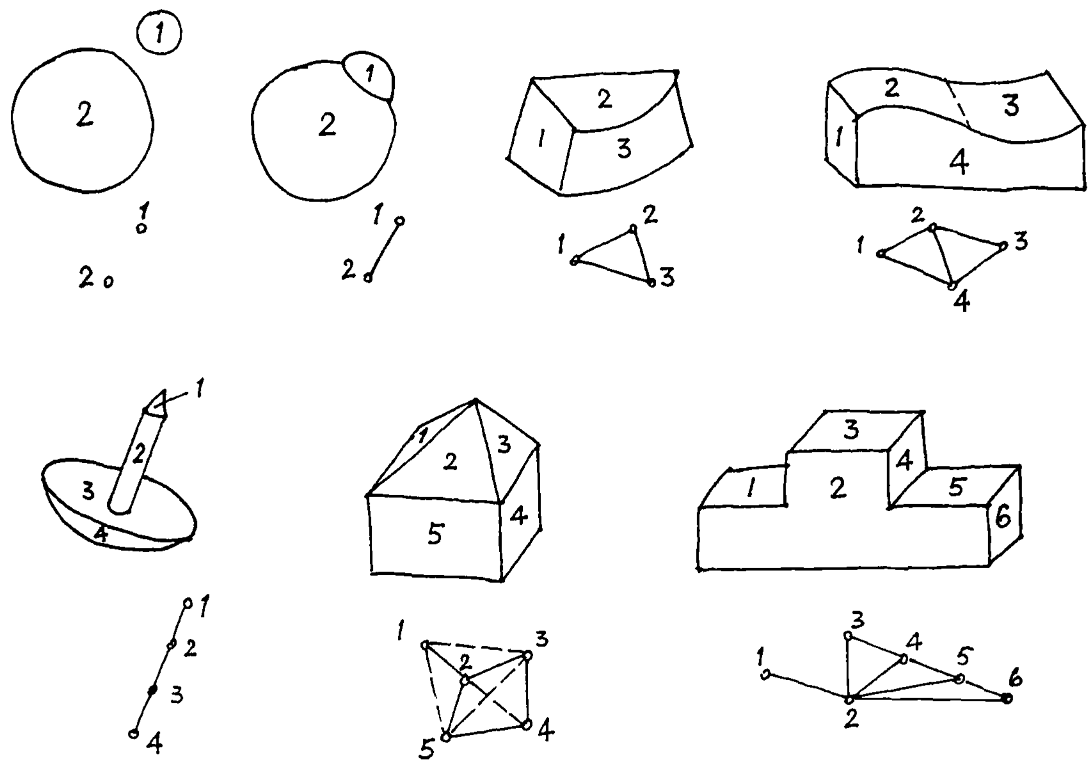


Figure 6.2.1. Segmented surfaces and their adjacency graphs.

Obviously, some surface information is lost when extracting the adjacency graph from a segmented surface. For example, a region's convexity/concavity, orientation and area size cannot be recognised from the node or arc of the adjacency graph, nor can the length of the boundaries, because they are related to the concept of distance in geometry. An adjacency graph presents only the topological properties, connectivity and adjacency, among all patches.

Remark 6.1: *For the above reasons a node's position in the graph is only notional, it can be marked at any position inside the patch. So, for convenience and without loss of generality, the node is marked at the centre of the patch But this does not introduces any idea relevant to arc length in the graph. A longer or a shorter arc denotes the same adjacency between the patches.*

6.2.2. The topological meaning of the adjacency graph.

Topology is the study of the properties of structures which are invariant under bicontinuous transformations. Topology permits a surface to be thought of as a rubber sheet. It rejects any measure relevant to the notion of distance, and allows a surface to be stretched, compressed and bent, *e.g.*, to change from a convex shape to a concave shape and so on. All these operations will not change the adjacency of the surface, since all adjacent points still remain adjacent. Tears, or the joining of separate bits together are prohibited as non-topological operations, since the neighbourhoods are disturbed. So the adjacency graph of a segmented surface is unique in the topological sense. For examples, the AG itself is invariant under rotation, scaling and translation operations, as they will not change the topological property of the AG. Hence the AG makes contribution to the uniqueness of the representation.

The separation of a surface is invariant under a topological transformation. Consider the following example of drawing a closed curve C on the surface of a sphere. Put graphically, it can be seen in Figure 6.2.2 that the closed curve separates the surface of the sphere into two distinct parts, S_1 and S_2. To get from a point p_1 in S_1 to a point p_2 in S_2, it is necessary to cross C. When the sphere is mapped to an ellipsoid by a topological transformation, the closed curve C is also mapped to a continuous closed curve C^* on the ellipsoid. Thus the ellipsoidal surface is still separated into two distinct parts, S_1^* and S_2^* by C^*. Hence, it is possible to maintain the separation among all patches on a segmented surface under certain topological transformations.

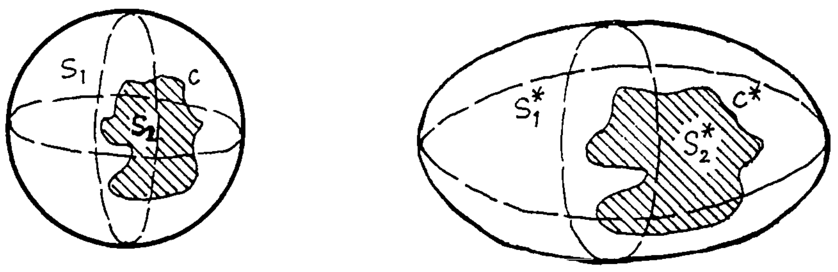


Figure 6.2.2. The separation of surface maintains invariant under a topological transformation.

Note that the intersection results in the boundaries of different patches. To present the segmented surfaces, the sorts of general topological transformations have to be limited to satisfy the homogeneity requirement raised in Chapter 5. That is, a surface should be thought as some piecewisely connected rubber sheets. More specifically, surface deformations of bending, stretching and compression, even being large ones, are still allowed if they retain a region's homogeneity in curvature sign; but the twisting, even being a small one, is excluded because it damages the uniformity of the surface curvature sign in a region. Under these assumptions, no matter that the real shape of a region represented by a node on the adjacency graph cannot be directly seen in a graph, it is known that a node must correspond to a patch of a certain shape class, and a pair of linked nodes must correspond to two adjacent patches of different attributes. This further implies that the adjacency of a segmented surface changes only when the segmentation of the surface changes significantly The adjacency graph is thus an indication of surface segmentation in scale space. An example is illustrated in Figure 6.2.3:

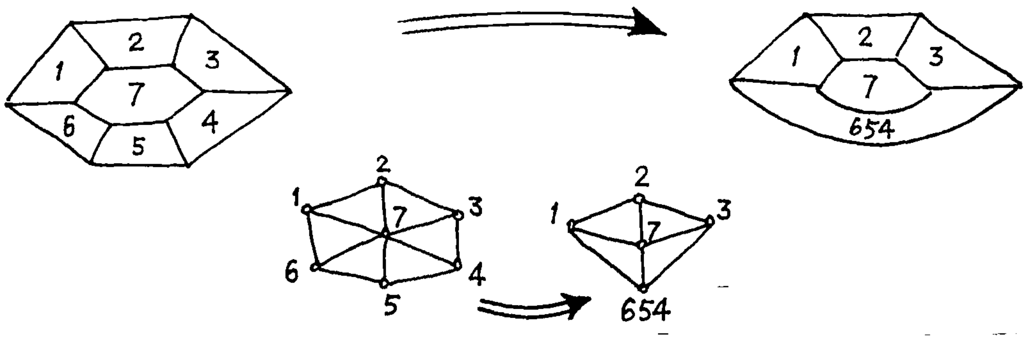


Figure 6.2.3. The adjacency graph changes when the structure of a segmented surface changes.

On the left is shown a segmented surface of seven patches and its adjacency graph. When the surface changes (*e g*, because of smoothing) to a new version of five patches as shown on the right, patches 6, 5 and 4 merge into a new patch 654, their boundaries shared with patch 7 also merge into a new one. These changes can be indicated by the changes of the corresponding nodes and arcs in the adjacency graph. By merging nodes 6, 5 and 4 into a new node 654 and arcs 6-7, 5-7 and 4-7 into a new arc 654-7, relabelling arcs 6-1 as 654-1 and 4-3 as 654-3, a new adjacency graph is produced, which is exactly the AG of the new surface version.

6.2.3. AG changes in scale space

Now consider the changes of an image, its segmentation and description in scale space. In the 1-D case, a curve may be partitioned into several segments at some critical points, then be symbolically represented by these points and segments As any arbitrary curve will be smoothed to a *straight line* when the smoothing scale increases continuously, all these segments will be gradually merged into one in the process. Similarly, in the 2-D case, any closed curve can be partitioned at some critical points and can be symbolically represented by these points and segments. All these segments will gradually merge into a *circle* during the smoothing process.

In the 3-D case according to the scale space behaviour theory addressed in Chapter 4, *(structure rule:)* nest-structured patches organise their scale space behaviours within a nest structure on the KH sign image; and *(Final fading shape theorem:)* when a non-minimal surface changes to flat through successive Gaussian smoothing, for any zero thresholds of K and H satisfying the consistent zero thresholding inequality (4.3.10), the final fading non-flat patch must be cylindrical: a ridge shape for a globally convex surface and a valley shape for a globally concave surface Hence, when using the adjacency graph as an indication, a similar phenomenon will appear: all nodes on the adjacency graph will eventually merge into a single node when different patches of a *free-form* surface merge into one patch with a nearly homogeneous curvature sign.

Being consistent with the scale space behaviour of 1-D and 2-D representations, the adjacency graph shows itself to be a tool qualified to describe the topological structure of a segmented surface in the 3-D case[1].

6.2.4. AG of flexible surfaces

As the AG presents only the topological properties of surfaces and it permits surface deformations of bending, stretching and compression, it is free from geometrical constraints. This leaves space for representing some kinds of flexible surfaces.

There are various flexible objects such as a flag, a rolled carpet, a balloon, a snake or a cup of water. These objects are quite different from each other except that their surfaces are all flexible in shape, to a greater or lesser extent The idea of "flexible" here is too soft and broad to handle. A specific sort of flexible object will be concentrated on.

---

[1] More precisely speaking, it should be the "2½-D case" rather than in the "3-D case", because they are essentially different, even though the adjacency graphs in both cases will eventually and similarly become a single node when the smoothing scale increases First, to present those self-occluded patches, the AG of the surface in 3-D space will have more nodes than the AG of the surface in 2½-D space, which presents only those visual patches. Second, the 3-D surface will eventually contract to a spherical shape when the smoothing scale increases, whereas the 2½-D surface tends to a flat shape.

They are not shape-free objects but folding rigid objects as in [Goldgof *et al* 1988], even some elastic objects which can change an individual patch's shape without damage to the patch's homogeneity and the boundaries among patches. In other words, when the surface changes, the topological property of the segmentation remains invariant. In Figure 6.2.4, a robot arm illustrates a folding rigid object and a vibrant drum head illustrates an elastic object.

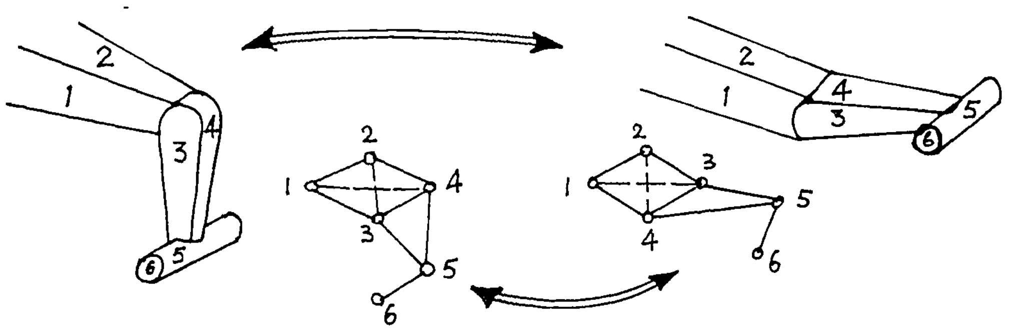


Figure 6.2.4 The shapes of a moving robot arm changes while the adjacency graphs remains invariant.

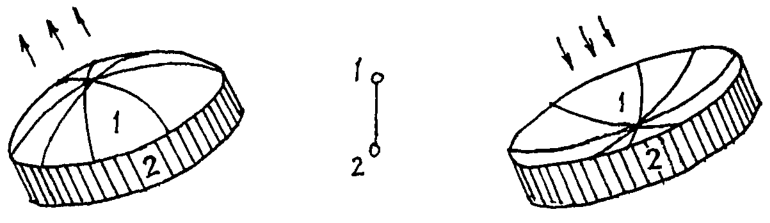


Figure 6.2.5. The shapes of a vibrant drum head changes while the adjacency graphs remains invariant.

As the small changes of flexible objects can be temporarily ignored and attention only paid to their adjacency graphs, this representation is convenient for the symbolic description and model matching of flexible objects.

6.2.5. Using the AG in matching

For the purpose of recognition, descriptions are produced for a unknown object and some models. Recognition is achieved by matching the object with the models. This is virtually a matching between their descriptions.

As usual, matching a description which has many attribute items suggests an expensive computation. Some *a priori* knowledge is therefore introduced to improve the searching efficiency. This gives some features a higher precedence in searching to speed up the process of rejecting unqualified candidates. In this sense, using an AG first strategy in model matching assigns the highest precedence to searching the topological attributes. As the AG represents properties that all surfaces possess, this assignment does not empirically impose a precedence order relating to specific types of surfaces, nor any threshold. This will be important and helpful for processing unknown objects.

As the AG itself is topologically invariant, using the AG to match the model and the object is a conventional problem in Graph theory [Deo 1974]. If a strong difference is found between two AGs, the candidate object will be rejected. On the other hand, if the matching succeeds, this is strong evidence in favour of the candidate being recognised as the desired object.

6.2.6. The uncertainty introduced by the AG

Successful AG matching suggests the candidate surface and the model surface have the same topological structures. However, it is too early to say that both surfaces match. As the relationship between the AG and the surface is not a 1-1 correspondence, it might happen that an AG is shared by objects with distinct geometrical appearances as illustrated in Figure 6.2.5. As topological properties are very general properties of surfaces, where metric attributes or geometrical properties are completely ignored, uncertainty may appear when the topological properties alone are used in the description or the matching. On the other hand, it is a price paid for gaining capability to represent some flexible surfaces using AGs.

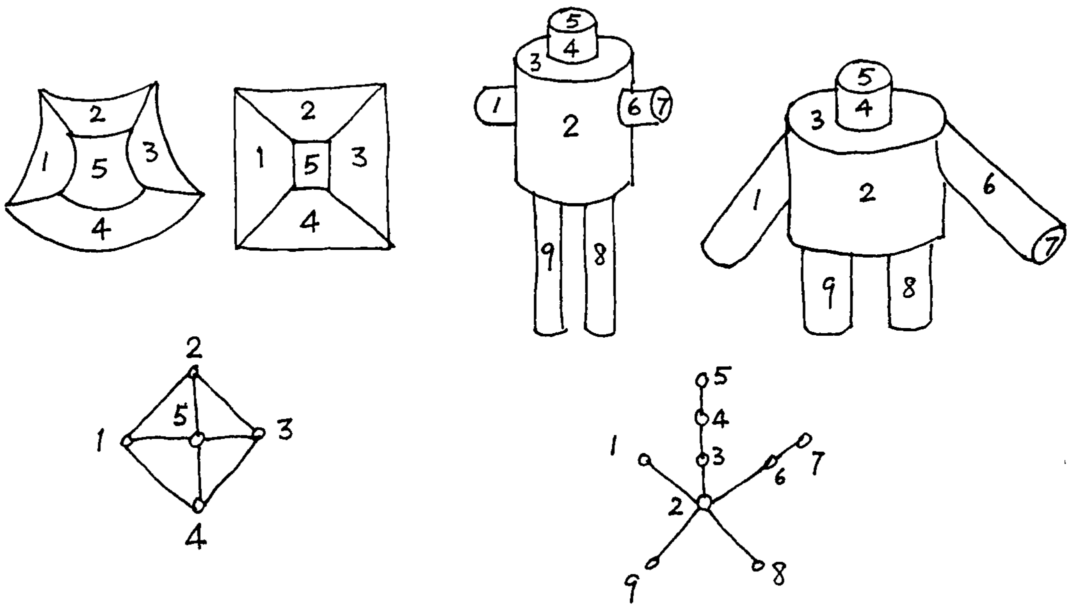


Figure 6 2.6. An adjacency graph shared by different surfaces.

6.3. The attribute list of the segmented surface

After extracting topological properties from the shape information, geometrical properties and some other visual properties of surface such as the reflection, colour and so on remain. As this thesis is concerned with range data, only the geometrical properties are of interest. All these geometrical properties are organised into an attribute list (AL), as the complement to the adjacency graph of the same segmented surface. They include some stable information that captures the more general and less varying properties of a shape as well as some information that is sensitive to finer distinctions between shapes.

6.3.1. Using the AL to reduce uncertainty

The AG and the AL complement each other in describing a segmented surface. Their functions also complement each other in recognition, *i e*, the AL can be applied as a constraint to reduce some uncertainty introduced by the AG. As the representation is to be used for recognition, the shape description must be unique. Hence, after a successful matching between AGs of the object and the model further compensatory matching of the ALs follows, where any inappropriate candidate will soon be rejected if strong differences are found between the ALs.

In practice, different candidate surfaces seldom share both the AL and AG of the model surface. And candidate surfaces that are less different than others can be distinguished when more items of AL are used in the comparison. Therefore, uncertainty can be reduced by imposing the AL's items on the AG's nodes, and the more items that are used, the less uncertainty is likely to remain. For example, an adult's AG will be the same as a child's AG, but the sizes of their body parts (head, trunk, arms and foot) will be different. By testing the sizes the desired object can be found. On the other hand, there is a limit to the number of AL items that are useful, since a very detailed AL will be derived less consistently by earlier processes such as segmentation.

6.3.2. The structure of the geometrical attribute list.

The structure of the attribute list is a table, whose lines corresponds to nodes (surface patches) in the adjacency graph. Each line has attribute items to describe both large and subtle differences among objects, including:

1. curvature (K,H) signs. 2. patch area 3 shape centre. 4. perimeter p. 5. number of neighbours *nb*. 6. orientation 7. Least Square Fitting parameters LSF. 8. area of patch support. 9. ... *etc*. An attribute list is shown by the table below.

node	(K,H)	area	centre	p	nb	orientation	LSF	...
1	$K_1\ H_1$	A_1	$(x\ y\ z)_1$	p_1	nb_1	$(\alpha\ \beta\ \gamma)_1$	$(a\ b\ c\ \ .)_1$	...
2	$K_2\ H_2$	A_2	$(x\ y\ z)_2$	p_2	nb_2	$(\alpha\ \beta\ \gamma)_2$	$(a\ b\ c\ \ )_2$	...
.	.	.					.	. . .
n	$K_n\ H_n$	A_n	$(x\ y\ z)_n$	p_n	nb_n	$(\alpha\ \beta\ \gamma)_n$	$(a\ b\ c\ .\ )_n$	...

Figure 6 3 1 A geometrical attribute list of segmented surface

Although the number of attribute items can be increased arbitrarily and these geometric attributes can be further grouped into intrinsic and extrinsic attributes, or position independent and position dependent attributes and so on, the structure of the AL is limited largely by the attribute items that can be reliably derived by processes prior to the representation.

There is no linkage between ALs at different scale levels. Each item in the AL is only constrained by the node concerned in the AG, and indexing any items in the AL at the neighbouring level has to pass through the neighbouring AG, as the AGs are the skeleton part of the *FCGT* whereas the AL is the flexible flesh part as will be shown in the later section.

6.4. Why is the separation between topological and geometrical properties needed?

Since surface matching with the AG alone introduces some uncertainty in recognition, and geometrical properties can be used as constraints to reduce the uncertainty, a problem is raised: *Why is the separation between topological and geometrical properties needed?*

Earlier it was argued that the separation of topological and geometrical properties extracts special information of shapes, which is irrelevant to the coordinate system, as the core part or the first important component in the shape representation for recognition. By this strategy, matching can be significantly simplified and leave a frame compatible with some flexible surfaces as well. Apart from these considerations there are three more reasons based on the following observations for adopting this approach for representing surfaces.

6.4.1. Patch's shape information prior to size information

In human perception the shape of a surface can be grasped independently of its size and is usually more significant than its size For example, when an object is described as an ellipsoidal without referring to the lengths of its long and short radii, people can get a rough impression of it. The same applies to an object with a saddle or a cylindrical surface.

On the other hand, given a mixture of plates, balls and sticks, if their sizes are not extremely different, it is easy to sort them into three heaps without paying attention to their sizes. A large and a small ball are still balls! Furthermore, with the same resolution a 14" screen TV can show the same story as a 20" screen TV, where the only difference is the spatial size not the shape

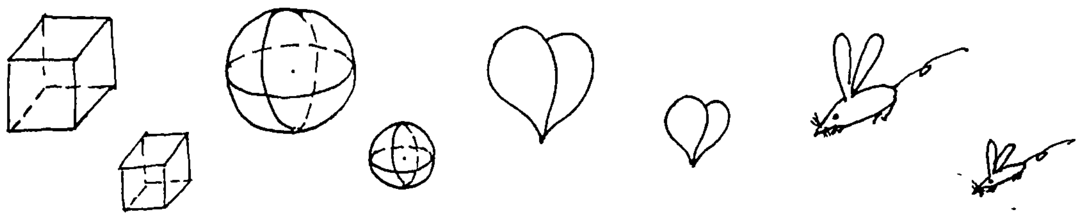


Figure 6 4 1 Objects of the same shape but different size.

6.4.2. Surface's structure effects prior to patch shape effects

A similar situation exists with the structure of a surface. Changes in surface structure are usually much more impressive than local changes in patch shapes. The Seven-pieces-puzzle (Chinese tangram) is an ancient game, where a variety of figures can be created by different arrangements of the same seven elementary shape blocks. 12 figures selected by [Clark and Wilkes-Gibbs 1986] from a book with 4000 such figures collected by [Elffers 1976] are illustrated below.

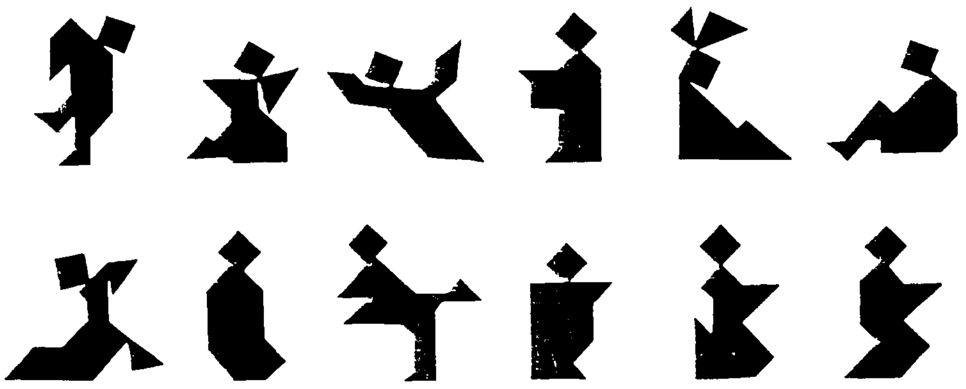


Figure 6 4 2 Different object shapes created by the seven-pieces-puzzle

Another example is models of human faces as shown in Figure 6.4.3. Regardless of using squares, rectangles, circular plates, elliptical plates or their mixtures, all these models suggest a *human* face. It is the structural relation between features and not the features themselves that cue recognition of a face [Hurlbert and Poggio 1986].

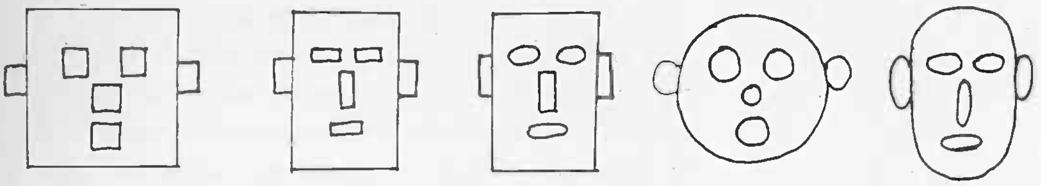


Figure 6.4.3. Human face models constructed by squares, rectangles, circular plates, elliptical plates or their mixture respectively.

The above two figures illustrate that it is possible and helpful to represent different surface patches and their relationships in a qualitative and symbolic form for recognition, where the shapes of patches can be left for later consideration.

6.4.3. Recognition-aimed representation

A third reason for adopting this approach to representing surfaces is that a recognition-oriented scheme of shape representation does not have to include all features of an object, nor to reproduce a shape's surface like a surface reconstruction. Significance varies for different features, and significance of a specific feature changes according to the visual processing task as well as the circumstance it stays in. Indiscriminately using all features in matching is not a sensible strategy. To recognise or distinguish an object, it is worth finding those *salient* features, which reflect the similarity between two shapes and yet allow subtle differences to be expressed as well, and which are stable across the scale levels. Salient features should always be paid attention to first in object recognition. As an example, stick-portrayed animal shapes from [Marr and Hildreth 1978] are illustrated in Figure 6.4.4, where each animal shape is first partitioned into different stick patches corresponding to a large part (such as trunk, leg, *etc.*) or junctions of the body, then the associated AG is given[1].

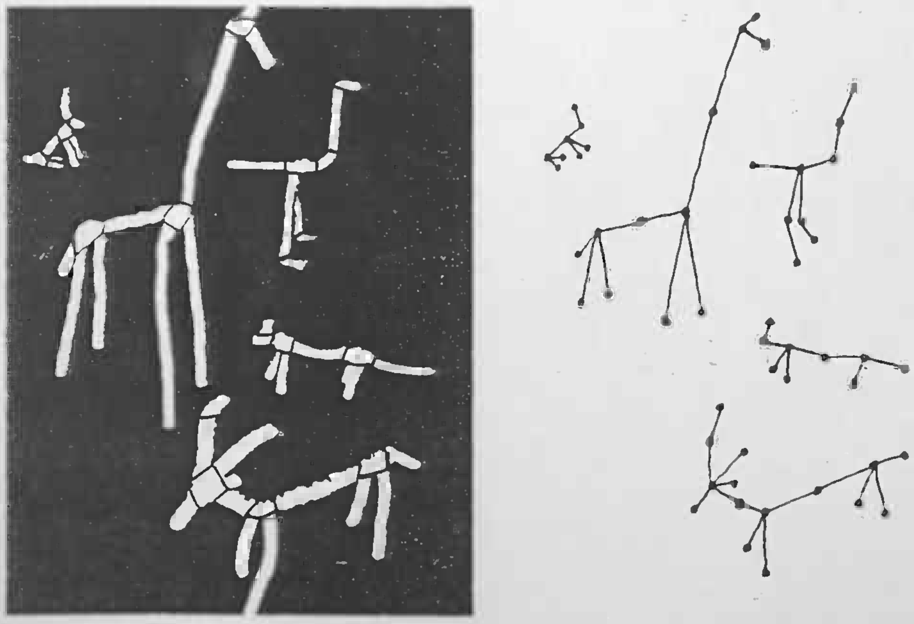


Figure 6.4.4. Stick-portrayed figures and AGs of animal shapes.

---

[1] It would be easy to claim that the given AGs are deceptive because the AG have encoded length and position information like the stick-portrayed figures due to their similar appearances. However, this claim is

From the scale space behaviour of surface features discussed in Chapter 4, it can be seen that if a mild concave patch exists on a globally convex surface (a common case) it is appropriate to be chosen as a salient feature for recognition. Briefly speaking, as the global shape is convex, the surface will eventually become convex during scale space filtering, so convex patches will gradually predominate and concave patches, if they exist, become scarcer and scarcer. As a concave shape is the most stable one among the eight possible surface shapes, a mild concave patch is qualified to be a very infrequent but most stable type among the non-convex shape patches, *i.e.*, to be a salient patch in recognition.

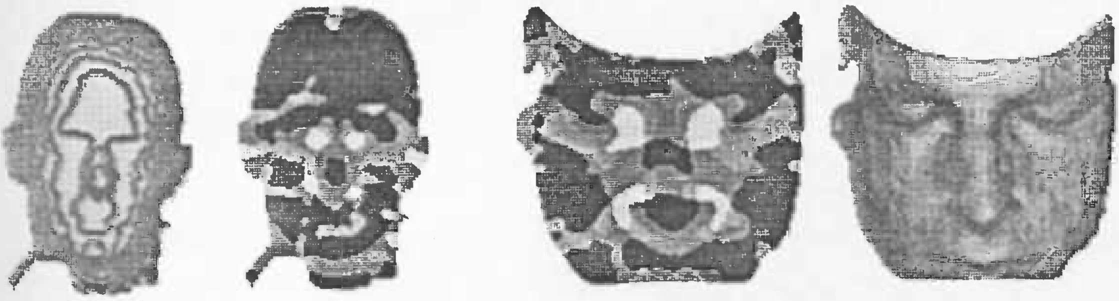


Figure 6.4.5. Recognition of two human faces with the same salient features.

To illustrate the above idea, the KH sign images of two human faces are shown in Figure 6.4.5. Human faces are globally convex. Some mild concave regions on faces can be chosen as the salient features. These regions are located at the inside corners of the eyes and on the lower edges of cheeks in GREEN (pit) on the KH image. (Meanwhile, convex patches are in BLUE and saddle patches are in RED or VIOLET.) During smoothing convex (BLUE) patches become dominant on both faces, though the concave patches resist the erosion from the convex patches and cling to both KH sign images and AGs for a long time. It can be seen that for two human faces the configuration of convex, concave and saddle patches are similar, so their AGs are similar too. However, for individual faces, the locations of the concave patches (GREEN regions) are different. Hence, by using the GREEN patches as salient features, the human faces can be distinguished from other surfaces, and the wanted face could be further determined by testing the ALs.

6.5. Flexibility of surface representation and its formal expression

6.5.1. Flexibility of surface representation

The representation of a flexible surface and the flexibility of a surface representation are two related but distinct concepts. A flexible surface needs a representation to represent the flexibility of surface adequately whereas the flexibility of a surface

---

apparently right but actually wrong. Recall that remark 1 in Section 6.2.1 has pointed out that a node's position in the graph is only notional. So the lengths of arcs in an adjacency graph have only notional meaning also. A shorter or longer arc length can change nothing about adjacency or connectedness represented by an AG, where any notion relevant to distance has been excluded. Hence, a giraffe and an ass, as well as a human being and a gorilla, may share an AG no matter the stick lengths in their stick-portrayed figures are quite different, thus it leaves some ambiguity. But a giraffe will never be confused with a human even though the arc lengths are arbitrarily changed. Interestingly, the above claim, on the contrary, shows that the topological property of a surface is different from and more general than the geometrical property. Thus, it is worth first extracting the topological properties to get the AG and leaving the geometrical constraints, such as length and position, in the AL for reducing the above ambiguity in later stages.

representation concerns an attribute of the representation itself which applies to each surface no matter whether it is a rigid or a flexible one.

So far, a segmented surface is represented in the form of an adjacency graph associated with an attribute list. Initially, the motivation is to provide a frame for representing some flexible surfaces. And, in return, by organising shape information in this way, a shape representation gains itself some flexibility in description, even in matching. Applying this representation form at each scale level produces a description of a segmented surface, namely, a *flexibly constrained graph (FCG)*.

6.5.2. Formal expression of flexible constrained graph

The *flexibly constrained graph (FCG)* can be formally expressed by an adjacency graph (AG) and an attribute list (AL) with an *addition* operator +:

$$FCG = AG + AL \tag{6.5.1}$$

where the addition operator is defined as an algebraic operation of the sets {AG} and {AL}, resulting in a unique element of the set {FCG}. As the adjacency graph (AG) can be seen as a fixed skeleton part of FCG whereas the attribute list (AL) can be seen as the flexible flesh part, with *whole* items from the AL, the strongest constraints are imposed on the AG, with *partial* items from the AL, a weaker constraint set is imposed on the AG; and with a *null* AL = ∅, no constraint is imposed on the AG, that is, the FCG is reduced to its core part, the AG. So, the representation reaches the extremes of its flexibility range in both the *full* and *null* AL cases. To emphasise the flexibility in representation, the above expression is re-written as.

$$FCG = AG + \mathbf{P} \circ AL \tag{6.5.2}$$

Where, **P** is the *partial* operator, bounded by the *identity* operator **I** and the *null* operator **Θ** as the follows:

$$\mathbf{\Theta} \le \mathbf{P} \le \mathbf{I} \tag{6.5.3}$$

$$\emptyset \equiv \mathbf{\Theta} \circ AL \subseteq \mathbf{P} \circ AL \subseteq \mathbf{I} \circ AL \equiv AL \tag{6.5.4}$$

So, both extreme cases of the FCG are:

$$\mathbf{P} = \mathbf{I} \quad => \quad FCG = AG + \mathbf{I} \circ AL \tag{6.5.5}$$

$$\mathbf{P} = \mathbf{\Theta} \quad => \quad FCG = AG + \mathbf{\Theta} \circ AL \tag{6.5.6}$$

Note that a full AL is the strongest constraint on the AG which leaves a small feasible solution field when matching, thus leads to a risk of failure as well as expensive computation; whereas a null AL could be a rather weak constraint on the AG, which introduces much uncertainty, thus failing to match the unique object. A trade-off is therefore to be expected between both extremes. In implementation the optimal operator **P** can be determined in such a dynamical way that items in the AL are put into a descending order of precedence so that matching can start from those of highest precedence items, say, the curvature signs; and when sufficient evidence is found the matching is stopped, where **P** is nearly null. So the partial operator **P** virtually gives the precedence order. Some *a priori* knowledge can be used in determining the order of precedence.

6.5.3. A formal expression of surface matching

Within this method of surface representation, matching between data and model becomes a matter of matching two of FCGs under a similarity measure. The matching task can be expressed formally as follows:

> Given a model **M**, data **D**, segmenting operator S, description operator FCG, difference operator $-$, metric $\| \ \|$ and numerical tolerance ε, it can be expected that:

$$\| FCG \circ S(\mathbf{M}) - FCG \circ S(\mathbf{D}) \| = \| FCG \circ S(\mathbf{M} - \mathbf{D}) \| < \varepsilon \quad \Rightarrow \quad \mathbf{M} = \mathbf{D} \tag{6.5.7}$$

So the match is in a weaker sense — an identity under the operator $FCG \circ S$.

From (6.5.2), there is

$$FCG \circ S = AG \circ S + \mathbf{P} \circ AL \circ S \tag{6.5.8}$$

Hence, the weakest case will be an identity under the operator $AG \circ S$, *i e.*, after the data surface and the model surface are segmented, they can be thought as identified if they share the same adjacency graph,

Certainly, without introducing the formalism of functional symbols as above, the whole idea in this subsection can still be stated using literal expressions. However, a further formalism, if possible, would be helpful to consider the matching problem from some other points of view

For example, when

$$\| FCG \circ S(\mathbf{M} - \mathbf{D}) \| \leq \| FCG \circ S \| \ \| \mathbf{M} - \mathbf{D} \| \tag{6.5.9}$$

because the norm $\| FCG \circ S \|$ is finite, a sufficient small $\| \mathbf{M} - \mathbf{D} \|$ can lead to

$$\| FCG \circ S(\mathbf{M} - \mathbf{D}) \| < \varepsilon \tag{6.5.10}$$

Thus, whatever the processing $FCG \circ S$ is, the model **M** must match with the data **D**. It further implies that to define a norm for measuring the difference between a model and a data the weightings of significant features should be fairly larger than that of insignificant features; otherwise, a mismatching may occur.

On the other hand, when

$$\| FCG \circ S \| \leq \| FCG \| \cdot \| S \| \tag{6.5.11}$$

and

$$\| FCG \| \leq \| AG \| + \| \mathbf{P} \| \ \| AL \| \tag{6.5 12}$$

a small $\| AG \|$ and $\| \mathbf{P} \|$ lead to a small $\| FCG \|$ which implies a match between model and data. But how to define the norms for the AG and the **P**? Or how to evaluate the structure of the AG and the precedence order given by $\| \mathbf{P} \|$ so that once a certain forms of AG and AL appear, the match will succeed? It is unclear at the moment how to define these norms, but it is certain that all these inequalities from (6.5.9) to (6.5.12) will be valid once the above norms have been defined.

6.6. Natural scale description of a surface in scale space

To set up a symbolic description for recognising an unknown surface, the extraction of surface shape features is a necessary stage. Surface segmentation is based on the shape feature extraction, and, in turn, gives different patches of uniform shapes.

Extraction of shape features at a single scale may give a useful result if 1) the ratio of signal to noise is fairly high; or 2) the statistical distribution of the noise is known or its characteristics can be obtained in the processing; or 3) the surface shape types are known in advance. Otherwise, there is too little information to distinguish the signal from noise at a single scale.

Nevertheless, human perception appears capable of choosing the "best" or the "natural" scale to recognise or distinguish unknown objects It is well known that the human visual system initially analyses the retinal image through a number of spatial filters which differ in the amount of smoothing that is performed in space and in time [Wilson and Bergen 1979]. A multiple scale processing was first proposed by [Rosenfeld and Thurston 1971] to extract more information about the unknown object than a single scale processing. However, due to scale effects, it is rather difficult to let the computer judge which scale is the "best" or the "natural" one to describe objects in visual processing Feature extraction and the scale choice thus become a "chicken and egg" problem.

Witkin [1983] showed how to deal with this problem in the 1-D case. His scale space filtering method efficiently broke the dead-lock by combining analysis at different scales. This method can be directly extended to the 2-D case [Mokhtarian and Mackworth 1986]. Unfortunately, a direct extension for the 2½-D case is not a trivial task.

In this section, the part of Witkin's ideas which are still valid in the 2½-D case, and the suitable form to carry out these ideas in the 2½-D case are discussed. Some problems in implementing them are then dealt with

6.6.1. Witkin's natural scale description

Witkin's method continuously smooths the 1-D waveform to a straight line, tracks the movement of the 2nd order zero-crossings (ZCs) that partition the smoothed curve from fine to coarse scales, then sets up a tree from the fingerprint-like trace of the ZCs, and finds the scale range where the structure of the tree is invariant. The segmentation of the signal curve, which corresponds to the tree, will persist over such a scale range. Hence it is stable. These scales are called the "natural" scales [Witkin and Tenenbaum 1983b] and the description of the signal's structure at these scales is called the "natural" scale description, which will be used in matching with model descriptions.

Briefly speaking, Witkin's method is composed of two procedures. The first procedure unfolds the scale space behaviour of all features across the scale levels. The second procedure extracts significant "events" from the trace of these features on the scale space image using the stability criterion. And common to both procedures, the key linkage is to select an appropriate carrier of segmentation information. In Witkin's case the carrier is the zero-crossing (ZC).

The fingerprint theorem [Yuille and Poggio 1983] points out that the ZCs of the 1-D curve cannot be created but may be deleted in moving from fine to coarse scale. This theorem provides a theoretical support for tracking the movement of the ZCs in scale space and describing it by a simple tree. So these ZCs are the right carriers of segmentation information, and the adjacency of the existing ZCs will never be disturbed even though some of them will disappear when the scale increases.

6.6.2. Extension of Witkin's method from 1-D to 2-D and 2½-D.

In the 2-D case, any arbitrarily closed curve that does not meet itself will be smoothed to a circular closed curve containing a smaller area than the initial one [Mokhtarian and Mackworth 1986]. The carrier of the segmentation information is still the 2nd order ZCs, and the adjacency of existing ZCs will never be disturbed, too So, the extension from the 1-D curve to the 2-D curve is fairly simple and direct Perhaps, the main difference between both cases is that the 1-D waveform is an open curve whereas the 2-D curve may be a closed curve.

However, in the 2½-D case, a surface point is controlled by neighbouring points not only on both sides but in all the directions. Difficulties arise due to the substantial differences between the 2½-D case and the 1-D or 2-D cases. Yuille and Poggio [1983] showed that in the 2½-D case while new ZCs cannot be created, they are free to split and merge as well. The 1-D results of the fingerprint theorem may not be appropriate in the 2½-D case. Ponce and Brady [1987] explored an edge-based method which formulated the scale space behaviour of surface features like a roof, step, smooth join and shoulder *etc.* in the application of the Surface Primal Sketch representation. The experimental results are nice. However, these features are models imposed on surfaces not the "intrinsic" attributes of surfaces themselves, thus the scope of representation is limited. And due to the behaviour of the ZCs in the 2½-D case, Ponce and Brady [1987] mentioned that the tracking forest (a group of trees) has never been explicitly built. Hence, for region-based surface segmentation and presentation it is unlikely to be a prospective way to use the ZC as an information carrier in the extension of Witkin's method in the 2½-D case. But, what can replace the ZCs as information carrier?

The AG node can be used as an information carrier in region-based processing, because even though AG nodes may be created at some intermediate scale, the number of AG nodes (as the abbreviates of regions) is much less than that of the ZCs, and a change of the AG is always consistent with a significant change of the segmentation not with a small change of an individual patch. So, within the network form, an AG is a simpler and more stable tool to implement the tracking as a ZC does in the 1-D case

6.6.3. Tracking the movement of shape features in scale space

According to the *structure rule* in Chapter 4, *nest-structured patches organise their scale space behaviours within a nest structure on the KH sign image* This means that corresponding to their geometrical distributions, different features behave in a nested way at each scale level, and echo with their counterpart features at the higher and lower scale levels respectively. Since only connected patches can influence each other directly, and this influence will not become explicit in the AG before an existing surface patch vanishes or a new patch is created. Hence, when changes in a FCG occur, they are only *local* events. And the subsequent changes will happen when the scale varies gradually. According to the *convexity remaining theorem*, a wild surface patch will remain on the KH sign image over several scale levels although its area may change. All pixels remaining on this patch must have the same curvature signs, and any pixel leaving from the patch will have the same curvature signs to the patch it joins. So the changes of patches can be tracked by comparing the KH sign images at neighbouring scale levels.

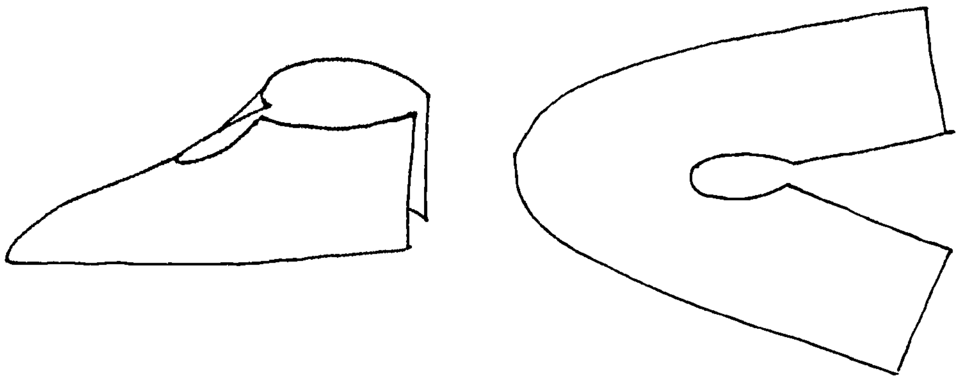


Figure 6 6.1. Curvatures of the shoe surface change gradually in the flattening.

Referring to the *Final fading shape* theorem, an analogy with the process of flattening a shoe (*i e* , reversing the process of shoe manufacture) in Figure 6.6.1 gives an intuitive model of how surface curvatures and types vary in scale space:

Rule 6.1 (Shoe flattening):

- *The maximum magnitude of surface curvatures* $|K|$ *and* $|H|$ *decrease monotonically.*
- *Most parts of a flat patch remain flat*
- *The Ridge/valley patch gradually changes to a flat patch*
- *The other patches gradually change to a flat patch via a ridge/valley patch.*

As the shape information of a specific surface patch (or a clique of patches) is represented by a node (or a clique of nodes) on the AG and the geometrical information is listed in the associated AL, tracking the movement of features in scale space means a comparison of FCGs at adjacent scale levels. And it is actually a comparison of the AGs, associated with the ALs as the constraints, at adjacent scale levels. On the AGs, the appearance and disappearance of arcs and nodes, as well as the movement of a patch's centre, indicate the changes of adjacency among the patches of the segmented surface. The items of the AL, such as, curvature signs, area size, perimeter and orientation, *etc.* indicate the geometrical changes of features. They will associate the nodes in tracking and in matching as well Among them the most important items are probably the curvature signs and the area.

According to the scale space behaviour of small surface features in Chapter 4, regions in the KH sign image may appear, expand, remain, contract and disappear as well, so tracking their movement in scale space should be implemented by a bi-projection operation between patches at neighbouring scale levels. "Parent → children" relationships will be assigned to the concerned nodes on both AGs after the projection from a patch at the higher scale level to some patches on the lower scale level. Conversely, "child → parents" relationships will be assigned to the concerned nodes after the projection from the patch at the lower level to some patches on the higher level Hence, as the bi-projections go up the levels, a "children ←→ parents" relationship *net* will be produced in vertical direction; therefore AGs at separate (horizontal) levels will be linked into a *tower* as whole.

The data structure of scale space is not a *forest* (a group of *trees*) but *nets* in both horizontal and vertical directions (*i.e.*, a *tower*). This is due to the substantial difference between the 2½-D case and the 1-D and 2-D cases [Yuille 1988]. However, such a tower may be too complex to use when all events are represented. Hence it should be simplified to expose the main events across different levels.

From the *interim shape creation* theorem, new shapes usually first appear at the intersections of patches. So when some inner pixels of a patch remain their shape types according to the *convexity remaining* theorem, some marginal pixels change their shape types and merge to neighbouring patches or to new created patches. Meanwhile those pixels with the same shape type as the background patch will not lose their shape attribute according the *response rule* Hence, the "inheritance right" is different from pixel to pixel (and from patch to patch). If both child and parent in a pair are of the same KH sign type, they have a *congenital* relationship admitted everywhere, the "inheritance right from blood". In contrast, if the child and parent in a pair are of the different KH sign types, they can only have a *acquired* relationship admitted by law, the "inheritance right from adoption". Some rules are therefore given to judge the significance of the "inheritance right" of a child-parent pair dependent on their position, shape type and overlapped area size.

Rule 6.2: The "congenital inheritance right" will always be significant for a child-parent pair stands in the AGs whatever their overlapped area size is.

Rule 6.3: The "acquired inheritance right" will be significant only when both areas of the child and the parent patch are greater than a least threshold A_{min}, and the percentage of their overlapped area in the parent or child's area is higher than a least threshold p_{min},

Denote the areas of the parent patch P_i and the child patch C_j as $A(P_i)$ and $A(C_j)$ respectively, their overlapped area as $A(P_i \cap C_j)$, Rules 6.3 can be formulated as:

$$A(P_i) \geq A_{min} \quad \text{and} \quad A(C_j) \geq A_{min} \tag{6.6.1}$$

$$\frac{A(P_i \cap C_j)}{A(P_i)} \geq p_{min} \quad \text{or} \quad \frac{A(P_i \cap C_j)}{A(C_j)} \geq p_{min} \tag{6.6.2}$$

To prevent too many arcs from appearing in the AG and prevent random effects from noise, the least thresholds in practice are heuristicly chosen, say, A_{min} = 15 pixels, and p_{min} = 25%.

Once having been justified, the nodes of the child-parent pair of significant "inheritance right" will be labelled by linking up them together from both AGs. And those nodes belonging to no of significant "inheritance right" pairs will be removed from the AGs they stand. So, as the bi-projections go up levels, the AGs will be simplified and linked up to form the skeleton of the tower

6.6.4. Structure of the AG tower

Once the AG tower is produced, subsequent processes of scale space processing are similar to the case where a coarse-to-fine tree is produced by Witkin's method. It should be noted that in practice a tower will not have infinite layers of interest.

This tower has some structural features, that might be the characteristics of the 2½-D scale space:

- The top layer will have fewer nodes than the bottom layer has.

- Generally speaking, the number of nodes at intermediate layers decreases from the bottom towards the top.
- However, at some layers this tendency can be reversed.

Hence the data structure of the tower is a *net*, rather than a *forest* — a group of *trees*. Applying a top down searching procedure to the tower, intermediate layers of the same AG will be found, where the linkage between the neighbouring AGs are a group of parallel pillars.

To make the idea clear, a simple example of a top cut quadrangular pyramid in scale space processing is illustrated in Figure 6 6.2, where the pyramid's AGs at different scale levels are linked into a tower

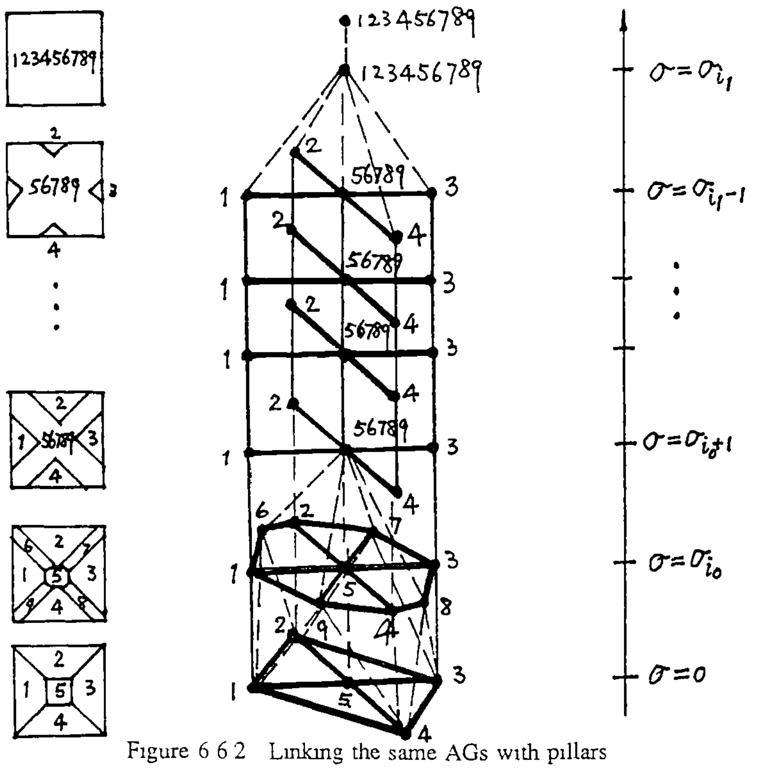


Figure 6 6 2 Linking the same AGs with pillars

At the scale level $\sigma = 0$, the pyramid's surface is segmented into five patches. Accordingly, the AG is composed of five nodes and eight arcs At the scale level $\sigma = \sigma_{i_0}$, four vertices and eight edges of the pyramid have been rounded, which creates four new narrow ridge patches 6, 7, 8 and 9 overlapping the position of the sideways edges. Accordingly, the AG consists of nine nodes and sixteen arcs, the new nodes 6, 7, 8 and 9, and the new arcs 6-1, 6-2, 7-2, 7-3, 8-3, 8-4, 9-4 and 9-1 denote the topological changes of the surface. Between both layers, the linkage 1-1, for example, denotes the congenital child-parent relationship; and the linkages 6-1 and 6-2 denote the acquired child-parent relationship. At the moment, the number of nodes increases, but it will begin to decrease at a higher layer. The shrinking of the top patch 5 and the broadening of the ridge patches 6, 7, 8 and 9 go on. When the scale level reaches $\sigma = \sigma_{i_0+1}$, they merge into a uniform peak patch 56789 while the sideways patches 1, 2, 3 and 4 remain flat in continuous reduction. Now the significant change of the surface structure happens: the AG is now composed of five nodes and four arcs. Compared with the bottom layer, the variation of the numbers of the arcs reflects the topological difference between the surfaces concerned.

Again, the linkages 1-1, 2-2, 3-3 and 4-4 denote the congenital child-parent relationships, and the linkages 56789-5, 56789-6, 56789-7, 56789-8 and 56789-9 denote the postnatal child-parent relationships. From now on till the scale level $\sigma = \sigma_{\iota_1 - 1}$, every patch changes its area gradually but the structure of the surface remains. So all AG layers in this scale range are of the same form. In addition, the "1-1" linkages grow into a pillar from layers σ_{ι_0+1} through σ_{ι_1-1}, so does the other "2-2", "3-3" and "4-4" linkages. When the scale reaches $\sigma = \sigma_{\iota_1}$, four residual flat patches "melt" into the central peak patch to form a uniform peak surface 123456789. And the concerning AG degenerates to one isolated node. From this node, the linkage 123456789-567889 denotes the congenital child-parent relationship while the other linkages, *e.g.*, 123456789-1 denotes the postnatal child-parent relationship. Before being flatten at a fairly large scale level, the surface shape may vary according to the *final fading shape* theorem. However, such change is beyond the scale range of interest, thus all layers from this point upwards are omitted in practice.

6.6.5. Extraction of stable segmentation and stable features.

A segmentation persisting over a large range of scales is a stable segmentation of the surface. Similarly, patches persisting over a large range of scales are stable patches. A stable segmentation provides information about the surface's structure for recognition, and stable patches provide information about the surface's salient features for recognition. Usually, the scale range from which a stable segmentation is extracted will overlap the scale ranges from which stable patches are extracted, but they are not necessarily the same one.

A stable segmentation can be found by using two measures produced by the top down searching of the tower. The first measure is a histogram of numbers of patches (nodes) versus scales. It shows over which range the numbers of patches change little. Note that the appearance of new patches and the disappearance of existing patches can occur at the same time. The above range is only a prospective range of the stable segmentation. And it can be improved by the second measure, which is a cluster net, retaining all children-parents relationships. Keeping all linkages and removing all arcs from the tower will generate such a net. Its function is similar to Witkin's coarse-to-fine tree but in a more complicated 2½-D case. It shows a specific scale range over which surface patches change but none of them appears or disappears. Thus the segmentation corresponding to the AG over this scale range is the stable segmentation. Certainly, a stable segmentation can also be extracted by directly comparing layers of the tower one by one. But using the above two measures can speed the process of searching for the stable segmentation.

As different surface patchs may have quite different stability in the scale space, it should also keep track of an individual patch's stability. Note that corresponding to stable patches, their AG nodes will cling to a series of AGs. These patches are either surface patches having the same shape type as the global shape of the whole surface, or surface patches having sufficient area to distinguish themselves from noise even though (some of them) eventually fade out like noise spots during the smoothing. In either case, the scale range of a patch p can be found by a top down searching of the tower for the scale levels σ_{born} where the patch p appears and σ_{die} where p disappears. This range is called the *span range* of the patch p on the AG tower, denoted as

$$Sp(p) = [\sigma_{born}, \sigma_{die}] \tag{6.6.3}$$

and the length of the span range is called the *life span* of p:

$$\| Sp(p) \| = \sigma_{die} - \sigma_{born} \tag{6.6.4}$$

Considering the noise effects, the least value of the life span of Gaussian scale for every patch in this research is set as

$$L_\sigma = 3 \tag{6.6.5}$$

The fourth rule related to individual patch's stability is then given below:

Rule 6.4: *To be considered as a stable feature, the life span of the candidate node should be no less than the given least value L_σ on the AG tower.*

$$\| Sp(p) \| \geq L_\sigma \tag{6.6.6}$$

Note that although the scale space itself prolongs from $\sigma = 0$ to $\sigma = \infty$, the scale range of interest D_σ from the scale $\sigma_{d_{\min}}$ to $\sigma_{d_{\max}}$, namely, the *dynamic range* of an AG tower, is actually limited in segmentation. Another rule in practice is given below:

Rule 6.5: *To get a surface segmentation of interest, only a limited dynamic range of scale $D_\sigma = [\sigma_{d_{\min}}, \sigma_{d_{\max}}]$ is needed, from which a stable segmentation may be obtained*

In this research the dynamic range of an AG tower is set as

$$D_\sigma = [2, 10] \tag{6.6.7}$$

From experience, the above bands (6.6.5) and (6.6.7) are suitable to obtain stable features and segmentations of surfaces from dense range data.

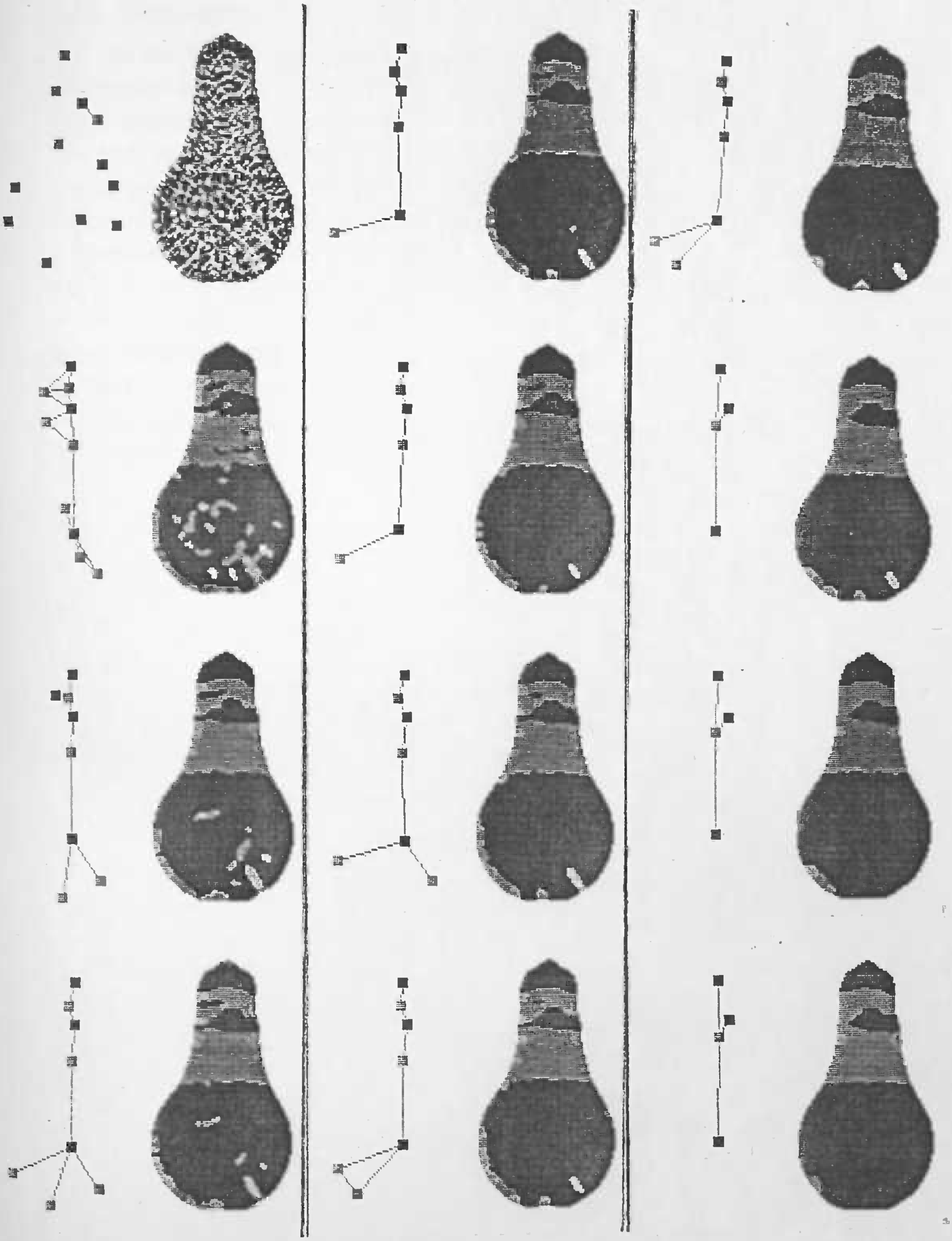
Obviously, the dynamic range of an AG tower should cover all AG node's span range:

$$[\sigma_{d_{\min}}, \sigma_{d_{\max}}] \supset [\sigma_{born}, \sigma_{die}] \tag{6.6.8}$$

By Rule 6.4, unqualified nodes will be removed from the AGs together with their direct arcs to other nodes in the same AGs and their direct linkages to parent or children nodes in neighbouring AGs. The corresponding removal is carried out within the AL lists. On this simplified FCG tower, the stable description of the surface can be obtained as layers within the dynamic range D_σ. Scales corresponding to these stable FCGs are thought as natural scales. These stable FCGs are usually those layers that link with parallel pillars as shown in Figure 6.6.2.

6.6.6. Experimental example.

While the research of surface representation of curvature-based method is still in progress, some results can be shown in Figure 6.6.3, where a light bulb is smoothed from the scale $\sigma = 0$ to $\sigma = 11$, and the AG of its stable description is extracted from layers of the FCGT over $\sigma = 2$ to $\sigma = 10$ regarding the curvature signs and the area size since they are usually the most important attributes of the AL. So the segmentation result of the light bulb is composed of three main patches: top, neck and body.



σ = 0 — 3

σ = 4 — 7

σ = 8 — 11

Figure 6.6.3. The KH sign image and AG of the stable FCGT layer of a light bulb (σ = 0, 1, 2, ..., 11).

6.7. Summary.

In this chapter, the *flexibly constrained graph tower* representation for surface shape recognition has been addressed. Emphasis has been given to a better organisation of shape information in the representation, which explicitly decouples the topological properties from the geometrical properties of a segmented surface and adopts nodes on the adjacency graph as the carriers of shape information. By this treatment the accessibility, scope, sensitivity and stability of the presentation are improved, and the capability of the representation to describe some sorts of flexible surfaces is gained. The concept of the flexibility of representation and its formal expression are introduced, where the traces of nodes on adjacency graphs form a skeleton, from which a stable and symbolic description can eventually be extracted. These show that this representation is a prospective extension of Witkin's scale space processing method from the 1-D or 2-D cases to the 2½-D case used for curvature-based or region-based segmentation. These results can supply a good basis for high level visual processing [Fisher 1989].

Chapter 7

Summary

7.1. Summary of results

In this research, the recovery of surface shape from a scale-related visual stimulus has been investigated. A computational theory of the curvature-based, scale space interpretation and representation of surface shape has been presented. The formulation of this theory is based on the investigation of diffusion smoothing and space scale behaviour of shape features, it is a $2\frac{1}{2}$-D extension of Witkin's 1-D scale space filtering.

First, diffusion smoothing theory has been pushed forward a substantial step. Different masks of Gaussian smoothing, repeated averaging and spline smoothing are uniformly derived from the general mask of diffusion explicit smoothing on a non-statistical basis, by which diffusion smoothing has proven itself as a more general filtering method. Thus, it is not surprising that diffusion smoothing can do more work than these widely-used methods mentioned above, such as the design of the optimal mask in the smallest neighbourhood. Meanwhile, a diffusion implicit smoothing scheme efficiently reduces the order of the computational complexity one by using a changeable time step owing to the unconditional stability of the computational scheme. Scale space surface processing thus becomes practicable. And more importantly, a "small leakage" diffusion smoothing model has been set up by making intuitive analogue to physical phenomenon, which treats the boundary condition properly and produces good quality KH sign images in scale space. This allows generic surfaces to be segmented using the KH sign images resulting from diffusion smoothing alone.

Second, the scale space behaviour (remaining, creation, fading and stability) of surface shape features is elucidated. The stability analysis of the surface KH categories (or large shape features) formulates the *consistent zero thresholding* inequality as the interrelationship between zero thresholds of the Gaussian curvature and mean curvature. It makes explicit that the zero thresholds ε_K and ε_H are not arbitrary but related via the mean curvature H, which brings the scale effects into zero thresholdings over different scale levels. By this formula, the zero threshold of the mean curvature can be automatically determined once the zero threshold of the Gaussian curvature has been chosen. This effectively prevents the "phantom case" of the KH classification occurring in scale space. On the other hand, the scale space behaviour of small shape features in 2½-D case has been formulated in three theorems and three rules. The *convexity remaining* theorem, the *interim shape creation* theorem and the *final fading shape* theorem made explicit the substantial differences of scale space behaviour between the 2½-D and the 1-D or 2-D cases. The conjecture in [Yuille 1988] has been verified for the case of surface smoothing with a symmetrical, positive function as the convolution kernel. The *conjugation rule* outlines the interaction of behaviour between the small and large features. The *response rule* facilitates an active detection to the shape of a large patch using a few small scattered shape-given patches as test signal, and may find application in robot navigation. The *structure rule* shows that the structure of nested features' behaviour is nested, thus a feature's behaviour can be localised and isolated temporally, which is the theoretical basis for tracking the movement of shape features in scale space. In fact, this has been implied in the *convexity remaining theorem*, the heart of this behaviour theory. The scale space behaviour analysis gives guidance for tracking the movement of features in scale space.

Third, both the diffusion smoothing and the behaviour theory are applied to surface segmentation in scale space. Surface segmentation based on curvatures in terms of the KH sign image is emphasised and posed as an under-determined mathematical problem of approximating surfaces up to the degree of curvature signs using the depth data alone. So, feasible solutions can be obtained by imposing restrictions to the solution space of the under-determined problem. One way to obtain a feasible solution is to directly search for a type of surface in the solution space with a restriction that curvature signs can be automatically approximated when surface depth is approximated using the "small leakage" diffusion smoothing Sculptured surfaces composed of peak, pit and saddle patches, *e g.*, human faces, have been shown to be of this surface type. As the principal directions have been shown to be able to detect cylindrical surfaces directly, the combination of the "roof" detection, the principal direction image and the KH sign image thus forms a curvature-based segmentation method using non-fitting and local surface operations. Another way to obtain a feasible solution of the above under-determined problem is to impose restrictions to the set of basis functions when surface depth is approximated using a piecewise recursive fitting associated with a parabolic shape modification. The feasibility of this treatment has been shown in the form of six lemmas, a theorem and two corollaries, which lead to another version of the curvature-based segmentation method using surface fitting and global operations. Both versions above may, along with other treatments, produce good quality KH sign images. Meanwhile, the essential differences between single scale and multi-scale segmentation methods have been discussed, which contributes to the extension of Besl's segmentation method to scale space. It makes clear that the bi-cubic polynomials should be excluded from Besl's basis fitting function set to reduce the risk of getting undesired shapes when region growing, even when surface fitting is treated as a conventional fitting problem.

Finally, the surface shape representation for recognition has been addressed. Most attention is paid to the organisation of a shape information representation. The advantages of separating topological properties of a segmented surface from its geometrical properties are elucidated through the discussion on the forms and functions of the adjacency graph and the attribute list, where nodes on the adjacency graph are adopted as the carrier of shape information. The concept of the flexibility of surface shape representation and its formal expression are introduced With this concept, the efficiency of an inexact matching can be raised by using the salient features first. Tracking the movement of surface features from fine to coarse scales yields a flexibly constrained graph tower representation, from which a "natural" scale description is eventually extracted under the judgement of the stability criterion. So this shape representation is a feasible extension of Witkin's scale space filtering from the 1-D or 2-D cases to the 2½-D case.

7.2. Contributions of this study

This study has set up a computational theory for scale space processing of surfaces from depth data to a shape representation for the purpose of recognition. The main contributions to computer vision are:

1. The diffusion smoothing theory has been pushed forward a substantial step in this study. The explicit scheme of diffusion smoothing in this thesis provides a theoretical basis to uniformly derive different masks of Gaussian smoothing, repeated averaging and spline smoothing, and to estimate their scale space behaviour. Meanwhile the implicit scheme of diffusion smoothing significantly improves the performance of diffusion smoothing. It is an unconditionally stable scheme which reduces the order of the computational complexity and yields a denser scale space without extra computation, thus makes the scale space processing practicable
2. A "small leakage" diffusion smoothing model is proposed in this thesis, which treats the external surface boundary properly, and thus prevents surface shapes along the boundary from distortion and produces good quality curvature sign images
3. Stability analysis of surface shape categories in the scale space processing is introduced. The *consistent zero thresholding* inequality is formulated as an estimation of the inter-relationship between zero thresholds of the Gaussian curvature and mean curvature This formula makes explicit that both zero thresholds are not irrelevant but related via the mean curvature, which brings the scale effects into zero thresholdings over different scale levels. Using this formula the "phantom case" of KH category can be avoided and the zero threshold of the mean curvature is determined automatically once the zero threshold of the Gaussian curvature has been chosen at each scale level
4. Three theorems (*convexity remaining*, *interim shape creation* and *final fading shape*) elucidate the scale space behaviour of small shape features in 2½-D case. Since the convexity remaining theorem applies to any surface smoothing with a symmetrical, positive function as the convolution kernel, the conjecture in [Yuille 1988] has been verified.
5. The *conjugate rule* outlines the interaction of behaviour between the small and large features. The *response rule* shows that scattered shape-known signals can be used to test the unknown shape of a large surface. And the *structure rule* makes explicit that nested shape features organise their behaviour in a nested structure, thus a shape feature's behaviour is a local action and can be isolated temporarily. This is the theoretical basis for tracking the movement of features in scale space.
6. Application of principal directions has been pointed out. Cylindrical surfaces can be detected directly by principal directions rather than first linking them into the lines of curvatures
7. Segmentation based on surface curvatures in terms of the KH sign image is posed as an under-determined mathematical problem of approximating surfaces up to the degree of curvature signs using depth data alone. This shows the direction along which efforts should be made to obtain feasible solutions. Sculptured surfaces have been shown as the surface type whose curvature signs can be automatically approximated when surface depth is approximated. Hence, it can be segmented well with the KH sign image resulting from the "small leakage" diffusion smoothing only. The combination of the KH sign image, the PD image and the roof detection gives the main version of the curvature-based segmentation method using local, non-fitting operations. Meanwhile, the feasibility of segmenting surfaces via a piecewise, recursive surface fitting has been discussed and expressed as a theorem, which leads to

another version of the curvature-based segmentation method using global, fitting operations.

8. A new type of shape information representation is described. Topological properties of a segmented surface are separated from its geometrical properties, nodes on the adjacency graph are adopted as the carriers of shape information, and the flexibility of shape representation is introduced. This leads to proposing the *flexibly constrained graph tower* as a scale-related form of shape representation for the purpose of recognition. From the tower a stable shape description can be extracted for model matching. Thus it is an extension of Witkin's 1-D curve scale space filtering in the 2½-D case.

7.3. Areas for further research

The computational theory presented in this thesis provides a basis for further research. Many problems are still unclear and worth further investigating, for example:

- How to suppress noise using local, context-sensitive smoothing with the diffusion equation so that the internal boundaries can also be preserved during the scale space processing? This means smoothing with a variable diffusion coefficient hence it is an extension of the "small leakage" diffusion smoothing.
- How to select the zero threshold of Gaussian curvature automatically with respect to scale level? That is, to characterise the scale space behaviour of the Gaussian curvature.
- When contours of the sliced object can be partitioned at corners and smooth joins using intrinsic properties of curves in scale space, how to merge these 2-D segmentation results as a whole to get a 3-D extension? This may be another version of curvature-based surface segmentation using local operations
- When surface segmentation is posed as an under-determined problem of approximating surface up to the degree of the curvature signs using depth data alone, how to separate the approximation to curvature signs from that to the depth, *i e*, first get a feasible partitioning according to the shape (convexity and concavity), then apply surface fitting in each region?
- How to produce a surface shape representation used for reconstruction? To achieve this task, regularisation techniques can be introduced to deblur the blurring resulting from the smoothing process. It would be a promising direction to work since the existing diffusion smoothing has produced a dense set of intermediate results as the basis for the deblurring operation.

Appendix 1.

Comparison of computational complexity between GS and DISCT

DISCT is a scale-incremental process. A lot of computation can be saved when a changeable time step as in (3.3.62) is adopted in each time sub-interval (3.3.63). A comparison with the GS process is illustrated in Figure A.1, where $b = \frac{1}{2}$ and $t = \sigma^2$.

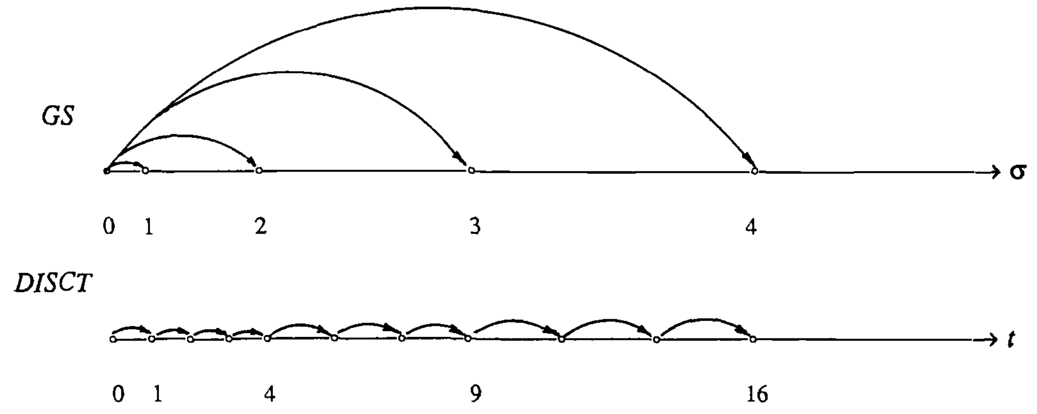


Figure A.1.1 A comparison of computational process between GS and DISCT in scale space

When the scale σ increases one, GS has to re-start the computation from the beginning thus the computing is independent of those results at previous scales whereas DISCT only needs to continue the sequential computation on the basis of the previous results. Hence, from (3.2.4) and (3.3.64), when $\sigma = 1, 2, \ldots, K$, the complexity estimations per pixel of GS and DISCT in scale space are:

$$C^*_{GS} = 8K(K+1) \tag{A.1.1}$$

$$C^*_{DISCT} = 10(3K-2) \tag{A.1.2}$$

When the Gaussian scale range $[0, K]$ varies ($K = 1, 2, \ldots$), a comparison of complexity estimations between GS and DISCT is given in Figure A.1.2:

K	1	2	3	4	5	6	7	..
C^*_{GS}	16	48	96	160	240	336	448	...
C^*_{DISCT}	10	40	70	100	130	160	190	...
$\frac{C^*_{GS}}{C^*_{DISCT}}$	1.60	1.20	1.36	1.60	1.69	2.10	2.36	...

Figure A.1 2. A comparison of computational complexity between GS and DISCT in scale space.

Note that a 3-D surface will be smoothed to flat within a narrower range of scales than that of a 1-D curve. The scale range of interest is actually limited for surface smoothing, *i e.*, the scale space will be a finite one in practice. Thus the ***average cost*** per smoothed version of image can be used to estimate the efficiency of smoothing methods in scale space. As shown in Figure A.1.1, when the Gaussian scale is K, DISCT can obtain

$3K-2$ smoothed versions of the data surface, of which $2(K-1)$ versions are the intermediate results obtained in the processing without any extra computation. Hence, the finite scale space generated by DISCT is *denser*[1] than that of GS.

From the complexity estimations (A.1.1) and (A.1.2), the average cost per pixel (of an $M \times M$ image) of DISCT is

$$\bar{C}^*_{DISCT} = \frac{C^*_{DISCT}}{3K-2} = \frac{10(3K-2)}{3K-2} = 10 \qquad \text{(A.1.2)}$$

which is a constant in scale space, and it is surprisingly one order of K lower than that of GS:

$$\bar{C}^*_{GS} = \frac{C^*_{GS}}{K} = \frac{8K(K+1)}{K} = 8(K+1) \qquad \text{(A.1.3)}$$

When the Gaussian scale range $[0, k]$ varies ($K = 1, 2, ...$), a comparison of average costs is given in Figure A.1.3:

K	1	2	3	4	5	6	7	...
$\bar{C}^*_{GS}$	16	24	32	40	48	56	64	...
$\bar{C}^*_{DISCT}$	10	10	10	10	10	10	10	...
$\frac{\bar{C}^*_{GS}}{\bar{C}^*_{DISCT}}$	1.6	2.4	3.2	4.0	4.8	5.6	6.4	...

Figure A.1.3. A comparison of average cost between GS and DISCT in scale space.

It is obvious that DISCT is a lower cost algorithm than GS in scale space processing.

---

[1] This will be useful when surface reconstruction is implemented via deblurring using the regularisation technique.

Appendix 2.

Corner-rounding effects resulting from finite differencing

The corner-rounding effect occurs when data smoothing and finite differencing are applied to calculate curvatures around a corner. If using a 3 pixel-width mask the effect propagates step by step outward from the corner (cf. subsection 3.2.2). This is a well-known fact, so the discussion here is only on the effect from the operation of finite differencing.

For simplicity without lose of generality, consider a corner as the intersection of a horizontal straight line l_q and an oblique line l_p as shown in Figure A.2.1:

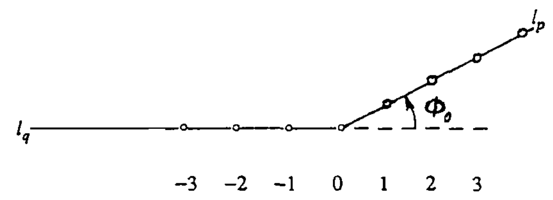


Figure A.2.1. A corner as the intersection of a horizontal line l_q and an oblique line l_p.

where on both sides of the corner x_0, slopes are different:

$$0 = q = f_{0^-} \equiv f(x_{0^-}) \neq f(x_{0^+}) \equiv \dot{f}_{0^+} = p \qquad \text{(A.2.1)}$$

$$\dot{f}(x_k) \equiv f_k = f_{0^-} \qquad k < 0 \qquad \text{(A.2.2)}$$

$$\dot{f}(x_k) \equiv f_k = f_{0^+} \qquad k > 0 \qquad \text{(A.2.3)}$$

As the C^1 discontinuity occurs at the corner, the orientation change, ϕ_k, of the curve instead of the curvature itself is calculated at each point x_k ($k = 0, \pm 1, \cdots$) to detect a corner. The tangent value of ϕ_k is

$$tg\phi_k = \frac{f_k^+ - f_k^-}{1 + f_k^+ f_k^-} \qquad \text{(A.2.4)}$$

Where, f_k^+ and f_k^- are numerical slopes resulting from the forward and backward finite differences with step jh respectively[1], where j can be 1, 2, 3, $\cdots$ and h is the pixel width.

$$f_k^- = \frac{f_k - f_{k-j}}{jh} \qquad \text{(A.2.5)}$$

$$f_k^+ = \frac{f_{k+j} - f_k}{jh} \qquad \text{(A.2.6)}$$

Hence, at the corner x_0 there are:

$$\dot{f}_0^- = q\,, \qquad \dot{f}_0^+ = p \qquad \text{(A.2.7)}$$

---

[1] The numerical slopes $\dot{f}_k^-$ and f_k^+ may be different from the curve slope $\dot{f}_k$ even though the curve is smooth at the point x_k, thus left and right slopes are equal at x_k: $f_{k^-} = f_{k^+}$.

$$tg\phi_0 = \frac{p-q}{1+pq} \tag{A.2.8}$$

Applying Taylor expansion to f_n at x_0, $n = 1, 2, \cdots$ gives:

$$f_n = f_0 + nhp\ , \qquad f_{-n} = f_0 - nhq \tag{A.2.9}$$

Thus, the forward and backward differences at points x_{-2}, x_{-1}, x_1 and x_2 are given below:

$$\dot{f}_{-1}^{-} = q\ , \qquad \dot{f}_{-1}^{+} = p - \frac{p-q}{j} \tag{A.2.10}$$

$$\dot{f}_{1}^{+} = p\ , \qquad \dot{f}_{1}^{-} = q + \frac{p-q}{j} \tag{A.2.11}$$

$$\dot{f}_{-2}^{-} = q\ , \qquad \dot{f}_{-2}^{+} = p - \frac{2(p-q)}{j} \tag{A.2.12}$$

$$\dot{f}_{2}^{+} = p\ , \qquad \dot{f}_{2}^{-} = q + \frac{2(p-q)}{j} \tag{A.2.13}$$

They leads to the orientation change ϕ at x_{-2}, x_{-1}, x_1 and x_2 as the following:

$$tg\phi_{-1} = \frac{\dot{f}_{-1}^{+} - \dot{f}_{-1}^{-}}{1 + \dot{f}_{-1}^{+}\dot{f}_{-1}^{-}} = \frac{(j-1)(p-q)}{(j-1)(1+pq) + 1+q^2} = \frac{p-q}{1+pq + \frac{1}{j-1}(1+q^2)} \qquad j > 1 \tag{A.2.14}$$

$$tg\phi_{1} = \frac{\dot{f}_{1}^{+} - \dot{f}_{1}^{-}}{1 + \dot{f}_{1}^{+}\dot{f}_{1}^{-}} = \frac{(j-1)(p-q)}{(j-1)(1+pq) + 1+p^2} = \frac{p-q}{1+pq + \frac{1}{j-1}(1+p^2)} \qquad j > 1 \tag{A.2.15}$$

$$tg\phi_{-2} = \frac{\dot{f}_{-2}^{+} - \dot{f}_{-2}^{-}}{1 + \dot{f}_{-2}^{+}\dot{f}_{-2}^{-}} = \frac{(j-2)(p-q)}{(j-2)(1+pq) + 2(1+q^2)} = \frac{p-q}{1+pq + \frac{2}{j-2}(1+q^2)} \qquad j > 2 \tag{A.2.16}$$

$$tg\phi_{2} = \frac{\dot{f}_{2}^{+} - \dot{f}_{2}^{-}}{1 + \dot{f}_{2}^{+}\dot{f}_{2}^{-}} = \frac{(j-2)(p-q)}{(j-2)(1+pq) + 2(1+p^2)} = \frac{p-q}{1+pq + \frac{2}{j-2}(1+p^2)} \qquad j > 2 \tag{A.2.17}$$

And generally, there will be

$$tg\phi_{-k} = \frac{(j-k)(p-q)}{(j-k)(1+pq) + k(1+q^2)} = \frac{p-q}{1+pq + \frac{k}{j-k}(1+q^2)} \qquad j > k \tag{A.2.18}$$

$$tg\phi_{k} = \frac{(j-k)(p-q)}{(j-k)(1+pq) + k(1+p^2)} = \frac{p-q}{1+pq + \frac{k}{j-k}(1+p^2)} \qquad j > k \tag{A.2.19}$$

Note that $q = 0$ in (A.2.1). Thus (A.2.8), (A.2.18) and (A.2.19) are simplified as the following:

$$tg\phi_0 = p \tag{A.2.20}$$

$$tg\phi_{-k} = \frac{(j-k)p}{j} = \frac{p}{1 + \dfrac{k}{j-k}} \qquad j > k \qquad \text{(A.2.21)}$$

$$tg\phi_{k} = \frac{(j-k)p}{j+kp^2} = \frac{p}{1 + \dfrac{k}{j-k}(1+p^2)} \qquad j > k \qquad \text{(A.2.22)}$$

So, from (A.2.20), (A.2.21) and (A.2.22), it can be seen that for the corner-rounding effects introduced by finite differencing:

a) When $j = 1$, $tg\phi_{\pm 1} = 0$. That is, little corner-rounding occurs at the neighbouring points around the corner x_0. Unfortunately, a single pixel-width step is rarely used in practice because ϕ may be very sensitive to noise in this circumstance.

b) Whatever j is, the scope of the corner-rounding effects is no greater than $[-j, j]$. Hence, a smaller j can be better for restraining the effects.

c) The edges near a corner will be rounded by finite differencing, leading to straight line edges becoming arcs, so false location of the corner may occur if only using a threshold to identify high curvature magnitudes

d) Corner-rounding effects decrease away from the corner, so the corner remains the point of the highest curvature magnitude. Hence, the corner should be detected by searching for the maximal magnitude in the high curvature neighbourhood of the corner.

Appendix 3.

Zero-drift effects in location of tangent joins of planar curves

When a straight line is tangent to an arc, a smooth join point is created. This is a very simple and common event of planar curves. But the position of a tangent join is not so easy to locate as it is supposed to be in visual processing.

Hoffman and Richards [1982] suggested segmenting plane curves into parts at (signed) curvature minima and using maxima and zeroes to describe the parts. However, an important join of planar curves, the tangent join, was totally neglected. Marimont [1985] described a representation for planar curves designed to serve as input to the subsequent computation While claiming that zeroes of curvature in the object curve always project to zeroes of curvature in the image, he also neglected the tangent joins of planar curves where the difference form of this claim is not true in the discrete case. Asada and Brady [1984] pointed out that "The detectability of a smooth join is essentially determined by the peak height h_{sj}, that is by the value of ϕ_{sj}. In particular, a smooth join is hard to detect .... It is not known how the threshold on ϕ_{sj} required to perceive a smooth join is related to that on ϕ required to perceive a corner." ϕ is the change of tangent orientation at a critical knot.

The difficulty of locating a tangent join is essentially due to the distinction between the continuous case and the discrete case In the continuous case, the location of the tangent join is really a trivial task since curvatures on both sides of the tangent join are obviously contrasting· curvature on one side is zero and on the other side is non-zero. Thus the discontinuity happens uniquely at the join point Around the join point, by searching either for a zero curvature point along the arc or for a non-zero curvature point along the straight line, the captured point must be the tangent join.

However, in the discrete case, shape distortion will be introduced by finite differencing and data smoothing around the tangent join They destroy the discontinuity by averaging and make the curvature at each point no longer an independent attribute but an attribute related to its neighbouring points According to the mean value theorem, curvatures at and around the tangent join become non-zero. Thus, the tangent join is no longer the discontinuity point separating the whole arc into zero and non-zero segments. Hence, the same strategy available to locate a tangent join in the continuous case may lead to a false or zero-drift result in the discrete case. In fact, tangent joins should now be found within points of non-zero curvatures. Therefore, it is necessary to set a zero curvature threshold to detect the tangent joins.

Noted that, although curvature magnitudes around a corner are also changed by the effects resulting from the same operations of data smoothing and finite differencing (see Appendix 2), the corner remains the maximum point there. Hence, it can be located by searching for the maximum magnitude in a small neighbourhood of high curvature magnitudes. In this sense the corner-rounding effects on location of critical points are substantially distinct from the zero-drift effects.

Zero-drift effect from data smoothing

The range of the zero-drift is closely related to the intensity of curvature discontinuity at the tangent join, the schemes of finite differences and the smoothing scale. For simplicity, consider a curve $f(x)$ consists of a circular segment $p(x)$ tangent to a horizontal line segment $q(x)$ at x_0 as shown in Figure A.3.1.

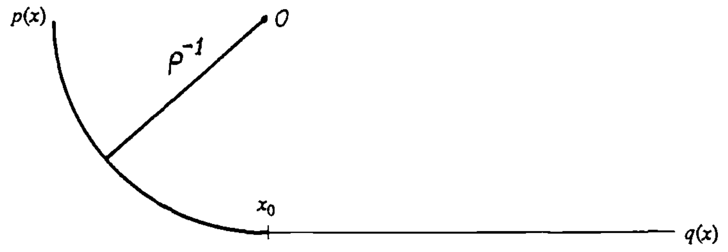


Figure A.3 1. A curve $f(x)$ consists of a circular segment $p(x)$ tangent to a horizontal line segment $q(x)$ at x_0.

where curvature $C(x_i) \equiv C_i = \rho > 0, \forall i < 0$ and $C(x_i) \equiv C_i = 0, \forall i \geq 0$ as shown in Figure A.3 2.

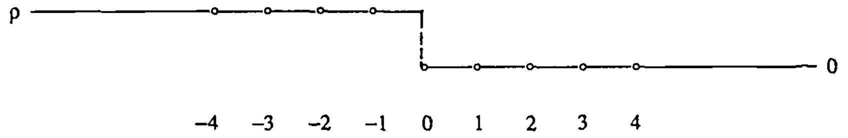


Figure A 3.2. The curvature $C(x)$ of the curve $f(x)$ above.

And post-smoothing is applied to the curvature $C(x)$ with a 3-pixel width mask (the least width mask) given in Figure A.3.3.

$$\frac{1}{2a+b} \times \boxed{a \mid b \mid a}$$

Figure A.3.3. The mask of the post-smoothing.

With one iteration of smoothing, curvature around the tangent point x_0 becomes:

$$C^{(1)}_{-2} = \rho \tag{A.3.1}$$

$$C^{(1)}_{-1} = \frac{a+b}{2a+b}\rho \tag{A.3.2}$$

$$C^{(1)}_{0} = \frac{a}{2a+b}\rho \tag{A.3.3}$$

$$C^{(1)}_{1} = 0 \tag{A.3.4}$$

With two iterations of smoothing, curvature around x_0 becomes:

$$C^{(2)}_{-3} = \rho \tag{A.3.5}$$

$$C^{(2)}_{-2} = \frac{3a^2+4ab+b^2}{(2a+b)^2}\rho \tag{A.3.6}$$

$$C^{(2)}_{-1} = \frac{3a^2+2ab+b^2}{(2a+b)^2}\rho \tag{A.3.7}$$

$$C^{(2)}_{0} = \frac{a^2+2ab}{(2a+b)^2}\rho \tag{A.3.8}$$

$$C_{1}^{(2)} = \frac{a^2}{(2a+b)^2}\rho \tag{A.3.9}$$

$$C_{2}^{(2)} = 0 \tag{A.3.10}$$

And with n iterations of smoothing, there will be:

$$C_{n-1}^{(n)} = \frac{a}{2a+b}C_{n-2}^{(n-1)} = \cdots = (\frac{a}{2a+b})^n C_{-1}^{(0)} = (\frac{a}{2a+b})^n \rho \tag{A.3.11}$$

$$C_{n}^{(n)} = 0 \tag{A.3.12}$$

Hence, using zero threshold of curvature to detect the tangent join x_{TJ} will give a false result:

$$x_{TJ} = x_n \tag{A.3.13}$$

where the zero-drift $d^{(n)}$ will be:

$$d^{(n)} = x_n - x_0 \tag{A.3.14}$$

as long as $C_{n-1}^{(n)}$ satisfies:

$$\varepsilon_{zero} \leq \left|C_{n-1}^{(n)}\right| = (\frac{a}{2a+b})^n|\rho| \tag{A.3.15}$$

On the other hand, the above inequality shows zero-drift will be damped when the smoothing scale increases. Given a, b, ρ and ε_{zero}, the possible zero-drift of the tangent join may, *at most*, reach:

$$\max n = \frac{log(\varepsilon_{zero}^{-1}) - \log|\rho^{-1}|}{log(2a+b) - log(a)} \tag{A.3.16}$$

It is likely that there will be little zero-drift if the following inequality is satisfied:

$$\varepsilon_{zero} \geq \frac{a}{2a+b}x|\rho| \tag{A.3.17}$$

This provides a strong clue for setting the least value of the zero curvature threshold ε_{zero}. For example, if using a 3-pixel width mask $a = 1$, $b = 6$, the zero curvature threshold may be: $\rho_{zero} \geq \frac{1}{8}|\rho|$.

Zero-drift effect from finite differencing

Now consider the effect resulting from finite differencing around the tangent join x_0. Suppose the 2nd order numerical derivative of $f(x)$ is calculated by the central difference with the difference step jh, where h is the pixel width, $j = 1, 2, 3, \cdots$.

$$\ddot{f}_i = \frac{f_{i-j} - 2f_i + f_{i+j}}{(jh)^2} \tag{A.3.18}$$

As $f(x)$ is of C^1 continuity and the segment $q(x)$ is a horizontal straight line and tangent to the circular segment $p(x)$ at x_0, there will be:

$$f_i \equiv f_0 \qquad \forall i > 0 \qquad \text{(A.3.19)}$$

$$\dot{f}_i \equiv 0 \qquad \forall i \geq 0 \qquad \text{(A.3.20)}$$

$$\ddot{f}_i = \ddot{f}_{0^+} \equiv \ddot{f}(x_0^+) = \ddot{q}(x_0^+) \equiv 0 \qquad \forall i > 0 \qquad \text{(A.3.21)}$$

$$\ddot{f}_i = \ddot{f}_{0^-} \equiv \ddot{f}(x_0^-) = \ddot{p}(x_0^-) \equiv \ddot{p}_{0^-} = \rho \qquad \forall i < 0 \qquad \text{(A.3.22)}$$

$$\dddot{f}_i \equiv 0 \qquad \forall i \qquad \text{(A.3.23)}$$

Applying Taylor expansion to f_{-j}, f_{1-j} and f_{2-j} at x_0 gives:

$$f_{-j} = f_0 + \tfrac{1}{2}(jh)^2\rho \qquad j > 0 \qquad \text{(A.3.24)}$$

$$f_{1-j} = f_0 + \tfrac{1}{2}((j-1)h)^2\rho \qquad j > 1 \qquad \text{(A.3.25)}$$

$$f_{2-j} = f_0 + \tfrac{1}{2}((j-2)h)^2\rho \qquad j > 2 \qquad \text{(A.3.26)}$$

Thus, 1st order numerical derivatives at x_0, x_1 and x_2 are:

$$\dot{f}_0 = \frac{f_j - f_{-j}}{2jh} = -\frac{jh}{4}\rho \qquad j > 0 \qquad \text{(A.3.27)}$$

$$f_1 = \frac{f_{1+j} - f_{1-j}}{2jh} = \frac{-(j-1)h}{4}\rho \qquad j > 1 \qquad \text{(A.3.28)}$$

$$\dot{f}_2 = \frac{f_{2+j} - f_{2-j}}{2jh} = \frac{-(j-2)h}{4}\rho \qquad j > 2 \qquad \text{(A.3.29)}$$

And 2nd order numerical derivatives at x_0, x_1 and x_2 are:

$$\ddot{f}_0 = \frac{f_{-j} - 2f_0 + f_j}{(jh)^2} = \tfrac{1}{2}\rho \qquad j > 0 \qquad \text{(A.3.30)}$$

$$\ddot{f}_1 = \frac{f_{1-j} - 2f_1 + f_{1+j}}{(jh)^2} = \tfrac{1}{2}\left(\frac{j-1}{j}\right)^2\rho \qquad j > 1 \qquad \text{(A.3.31)}$$

$$\ddot{f}_2 = \frac{f_{2-j} - 2f_1 + f_{2+j}}{(jh)^2} = \tfrac{1}{2}\left(\frac{j-2}{j}\right)^2\rho \qquad j > 2 \qquad \text{(A.3.32)}$$

So, when the size of difference step is h, tangent join x_{TJ} drifts at least from x_0 to x_1 as long as the curvature magnitude at x_0 satisfies:

$$\varepsilon_{zero} < \left|C_0\right| = \frac{\left|\ddot{f}_0\right|}{(1+\dot{f}_0^2)^{\frac{3}{2}}} = \frac{\frac{1}{2}\left|\rho\right|}{(1+\frac{1}{16}h^2\rho^2)^{\frac{3}{2}}} < \frac{1}{2}\left|\rho\right| \qquad \text{(A.3.33)}$$

When the size of difference step is $2h$, x_{TJ} will drift from x_0 to x_2 as long as the curvature magnitude at x_1 satisfies:

$$\varepsilon_{zero} < \left|C_1\right| = \frac{\left|\dot{f}_1\right|}{(1+f_1^2)^{\frac{3}{2}}} = \frac{\frac{1}{8}\left|\rho\right|}{(1+\frac{1}{16}h^2\rho^2)^{\frac{3}{2}}} < \frac{1}{8}\left|\rho\right| \qquad \text{(A.3.34)}$$

And when using a difference step larger than $2h$, say, $3h$, not only the magnitude of C_1 may rise, but also x_{TJ} will drift from x_0 to x_3 as long as the curvature magnitude at x_2 satisfies:

$$\varepsilon_{zero} < \left|C_2\right| = \frac{\left|f_2\right|}{(1+f_2^2)^{\frac{3}{2}}} = \frac{\frac{1}{18}\left|\rho\right|}{(1+\frac{1}{16}h^2\rho^2)^{\frac{3}{2}}} < \frac{1}{18}\left|\rho\right| \qquad \text{(A.3.35)}$$

And the larger the difference step jh, the broader the range of zero-drift.

Also from (A.3.34) the zero-drift is likely to be limited to one unit h or less if the following inequality is satisfied:

$$\varepsilon_{zero} > \frac{1}{8}\left|\rho\right| \qquad \text{(A.3.36)}$$

Similar to (A.3.17), this also provides a strong clue for setting the least value of the zero curvature threshold ε_{zero} when a difference step $2h$ is used in the computation of numerical derivatives with central differences.

In short, a proper zero curvature threshold is important to guard against zero-drift effects in location of tangent join of planar curve. Otherwise, the combined effects resulting from both of finite differencing and data smoothing may become far beyond than might be naively foreseen.

Appendix 4.

Approximating surfaces up to curvature signs using depth data alone

Referring to Definitions 5.1 to 5.4 given in Chapter 5, six Lemmas, one theorem and two corollaries are proven as the following.

Lemma 5.1: *(Existence) Given a straight line segment $g(x)=d+kx$, there must exist a smooth and concave (or convex) curves $f(x)$ passing through both ends of the same interval $I = [x_0, x_1] \subset [0, \infty)$ $f(x_0) \equiv g(x_0)$ and $f(x_1) \equiv g(x_1)$ such that when the significance measure $R_I(f,g)\}$ is sufficiently small, not only the concave curve $f(x)$ will be a zero-degree uniform approximation of the straight line $g(x)$ over the interval I but also*

1) $f(x)$ is an approximation of $g(x)$ in both slope and curvature magnitudes;

2) $f(x)$ is an approximation of $g(x)$ in the slope sign if $k \neq 0$, but never in the curvature sign,

3) $f(x)$ is invariant in slope sign over I if $k \neq 0$, and

4) $f(x)$ can be obtained within a finite number of successive approximations to $g(x)$.

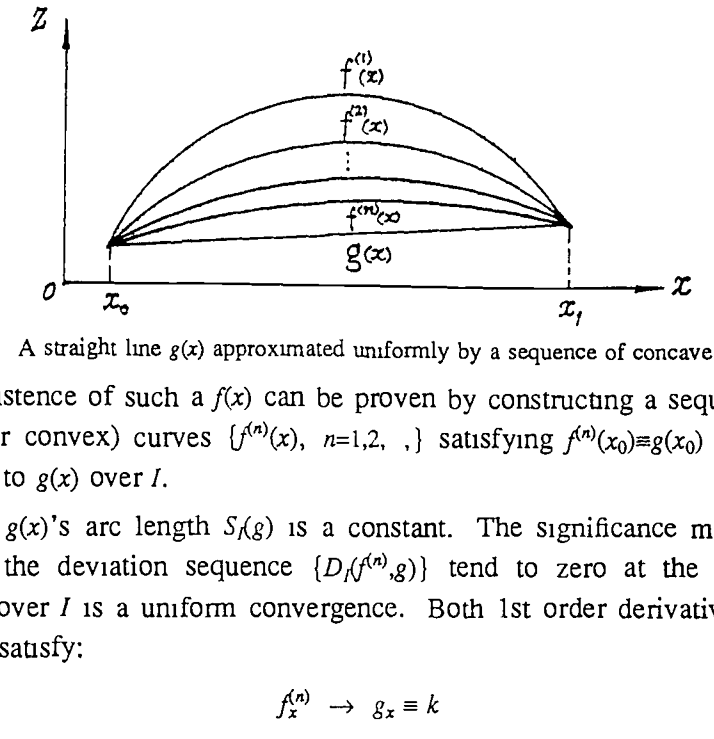


Figure A 4 1 A straight line $g(x)$ approximated uniformly by a sequence of concave curves $f^{(n)}(x)$.

Proof: The existence of such a $f(x)$ can be proven by constructing a sequence of smooth and concave (or convex) curves $\{f^{(n)}(x),\ n=1,2,\ ,\}$ satisfying $f^{(n)}(x_0) \equiv g(x_0)$ and $f^{(n)}(x_1) \equiv g(x_1)$ and converging to $g(x)$ over I.

Note that $g(x)$'s arc length $S_I(g)$ is a constant. The significance measure sequence $\{R_I(f^{(n)},g)\}$ and the deviation sequence $\{D_I(f^{(n)},g)\}$ tend to zero at the same time. So $\{f^{(n)}(x)\} \rightarrow g(x)$ over I is a uniform convergence. Both 1st order derivatives of $f^{(n)}(x)$ and $g(x)$ over I will satisfy:

$$f_x^{(n)} \rightarrow g_x \equiv k \tag{A.4.1}$$

Also note that $f^{(n)}(x)$ is concave over I, both 2nd order derivatives over I will satisfy:

$$f_{xx}^{(n)} \rightarrow g_{xx} \equiv 0 \tag{A.4.2}$$

When $k = 0$, $sgn(g_x) = sgn(k) \equiv 0$ over I. Since $f^{(n)}(x)$ is concave over I, $sgn(f_x^{(n)}) = 0, \pm 1$ and $sgn(f_{xx}^{(n)}) \neq 0 \equiv sgn(g_{xx})$ for all n. So, when $f^{(n)}(x)$ converges to $g(x)$, the magnitudes of $f^{(n)}(x)$'s slope and curvature also converges to that of $g(x)$, but the signs of slope and curvature do not.

When $k = const \neq 0$, $sgn(k)$ is invariant over the interval I. From (A.4.1), for any sufficient small number $\varepsilon > 0$, there must be a *finite* integer $N = N(\varepsilon) > 0$ when $n > N$

$$k-\varepsilon < f_x^{(n)} < k+\varepsilon \qquad \text{(A.4.3)}$$

leading to $sgn(f_x^{(n)}) \equiv sgn(k)$ being invariant over I. Hence take $f(x) = f^{(N+1)}(x)$, all the above conclusions are available.

Q.E.D.

Corollary 5.1: *Take a parabolic curve $a+bx+cx^2$ (a, b, c parameters) as the concave curve $f(x)$ in Lemma 5 1 When the significance measure over I is sufficient small,*
1) parameter c will be close to zero, parameter b will be close to k, and
2) the parabolic curve's slope sign will be invariant over I if $k \neq 0$.

Proof: When the significance measure tends to zero, from Lemma 5.1 there will be

$$(b+2cx)^{(n)} \equiv f_x^{(n)} \to g_x \equiv k \qquad \text{(A.4.4)}$$

Because x varies over I, there must be

$$c^{(n)} \to 0 \qquad \text{and} \qquad b^{(n)} \to k \qquad \text{(A.4.5)}$$

the conclusion 1) follows.

From Lemma 5.1, when $k \neq 0$, the conclusion 2) follows.

Q.E.D.

Lemma 5.2: *Within a finite number of recursions of polygonal approximation, a smooth concave curve can be piecewise approximated by a set of folded straight lines such that the concave curve and the folded line are close in each section up to the degree of curvature magnitude but not the curvature sign*

Proof: Apply the recursive polygonal approximation to the concave curve C. After a finite number of recursions the significance measure of concave curve to straight line over each section will be sufficiently small and the approximation produces a set of polygonal segments $L = \{L_{k,k+1}\}$, each of them has non-zero slope. (If a horizontal segment appears, replace it with a folded straight line whose vertex locates at an intermediate point of the curve C). From Lemma 5.1, for each line segment there will be a smooth and concave curve segment $\hat{C}_{k,k+1}$ which is close to the line segment $L_{k,k+1}$ up to the the degree of curvature magnitude but not the curvature sign

If $C_{k,k+1}$ falls into the area surrounded by $\hat{C}_{k,k+1}$ and $L_{k,k+1}$, the desired approximation achieves, so the polygonal approximation for $C_{k,k+1}$ finishes; otherwise, it continues.

Note that the recursive polygonal approximation can make the folded line uniformly converge to the curve C, leading to the significance measure of the curve to the line in each segment tend to zero. Thus the deviation of the curve C to the line L in each segment, as a higher order infinitesimal quantity, tends to zero faster than the line length does. Hence, the polygonal approximation of the curve C must finish in a finite number of recursions where every $C_{k,k+1}$ segment falls into the area between $\hat{C}_{k,k+1}$ and $L_{k,k+1}$. The conclusion thus follows.

Q.E.D.

Lemma 5.3: *Within a finite number of recursions, a smooth concave curve can be piecewise approximated by a set of smooth parabolic curves such that both concave and parabolic curves are close in every section up to the degree of the curvature sign.*

Proof: From Corollary 5.2, within a finite number of recursions, a set of folded straight lines can be obtained, which approximates the smooth concave curve in every section up to the degree of the curvature magnitude but not the curvature sign.

Now, approximate each concave curve segment $C_{k,k+1}$ with a parabolic curve $P_{k,k+1}$ passing through three points of the curve segment, i.e., both end points of the curve segment and the maximum deviation point from the straight line segment $L_{k,k+1}$, as shown in Figure A.4 2.

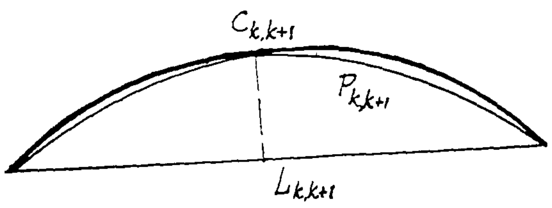


Figure A.4 2. Using parabolic fitting to correct the result of polygonal fitting.

Because the parabolic segment $P_{k,k+1}$ and the concave curve segment $C_{k,k+1}$ have the same concavity, they have the same curvature sign. Moreover, the maximum deviation of $P_{k,k+1}$ from $C_{k,k+1}$ is smaller than the maximum deviation of $C_{k,k+1}$ from $L_{k,k+1}$. Hence, the connected parabolic curve segments $\{P_{k,k+1}\}$ and the concave curve C will be close each other up to the degree of the curvature sign.

Q.E.D.

Lemma 5.4: *On a smooth surface, the sign of the Gaussian curvature K explicitly depends only on 2nd order derivatives of the surface whereas the mean curvature H also involves the 1st order derivatives.*

Proof: From the definitions of K and H in (2.2.12),(2.2.13) and (2.4.1) in Chapter 2, there will be

$$E = 1+z_x^2 > 0, \quad G = 1+z_y^2 > 0 \quad \text{and} \quad EG-F^2 = 1+z_x^2+z_y^2 > 0 \tag{A.4.6}$$

Hence

$$sgn(K) = sgn(LN-M^2) = sgn(z_{xx}z_{yy}-z_{xy}^2) \tag{A.4.7}$$

and

$$sgn(H) = sgn(EN-2FM+GL) = sgn(Ez_{yy}-2Fz_{xy}+Gz_{xx}) \tag{A.4.8}$$

Q.E.D.

Lemma 5.5: *When the parameter of the first fundamental form $F = 0$ (cf. (2.2.12)), the surface must be a developable one where $sgn(K) \equiv 0$ and $sgn(H)$ will be related only to the sign of the 2nd order derivative.*

Proof: From $F \equiv z_x z_y = 0$, it follows that

$$z_x = 0 \,,\, z_y \neq 0 \quad \text{or} \quad z_y = 0 \,,\, z_x \neq 0 \text{ or } \quad z_x = z_y = 0 \tag{A.4.9}$$

Therefore, it is always true that

$$z_{xy} = z_{yx} = 0 \tag{A.4.10}$$

$$z_{xx} = 0 \quad or \quad z_{yy} = 0 \tag{A.4.11}$$

Referring to (A.4.6)-(A.4.8), the above result leads to a developable surface where

$$sgn(K) = sgn(z_{xx}z_{yy} - z_{xy}^2) \equiv 0 \tag{A.4.12}$$

$$sgn(H) = sgn(z_{xx}) \qquad or \qquad sgn(H) = sgn(z_{yy}) \tag{A.4.13}$$

Q.E.D.

Lemma 5.6: *On a smooth surface when the parameter of the second fundamental form $M = 0$ (cf. (2 2 13)), the calculation of derivatives can be decomposed into x- and y- directions*

Proof: From $M = \frac{z_{xy}}{\sqrt{1 + z_x^2 + z_y^2}} = 0$, there is $z_{xy} = z_{yx} = 0$. So $z(x,y)$ can be separated as

$$z = \tilde{z}(x) + \overline{z}(y) \tag{A.4.14}$$

Hence, the derivatives can be obtained separately as

$$z_x = \tilde{z}_x(x)\ ,\ \ z_{xx} = \tilde{z}_{xx}(x)\ ,\ \ z_y = \overline{z}_y(y)\ ,\ \ z_{yy} = \overline{z}_{yy}(y) \tag{A.4.15}$$

Q.E.D.

Theorem 5.1: *(Feasibility:) To ensure that a recursive piecewise fitting can approximate a surface up to the degree of the signs of K and H within a finite number of steps,*

1) the maximal order of polynomial basis function should be 2, and

2) the feasible form of the fitting bi-polynomial is

$$z(x,y) = a_0+a_1x+a_2y+a_3x^2+a_4y^2 \tag{A.4.16}$$

Proof: Conclusion 1) is first proven as follows.

The bi-polynomial basis function set is $\{1, x, y, x^2, y^2, xy, x^3, y^3, x^2y, xy^2,...\}$. Obviously, a curved surface cannot be approximated up to the signs of K and H by a set of planar facets, so the maximal order of basis functions should not be lower than 2.

Suppose a basis functions of 3rd order, say, x^3, is appended as a term to the above fitting bi-polynomial (A.4.16). It becomes:

$$z(x,y) = a_0+a_1x+a_2x^2+a_3y+a_4y^2+a_6x^3 \tag{A.4.17}$$

The second order derivative z_{xx} will be

$$z_{xx} = 2a_2 + 6a_6x \tag{A.4.18}$$

As x varies over $[x_0, x_1]$ and the signs of coefficients a_2 and a_6 are unknown before fitting, the invariance of $sgn(z_{xx})$ in the region cannot be ensured, nor can the invariance of $sgn(K)$ be ensured, thus x^3 is not suitable to be selected as a basis fitting function, nor is y^3 or any term of higher order than 3.

Now conclusion 2) is proven. Calculating the derivatives of $z(x,y)$ in (A.4.16) gives

$$z_x = a_1+2a_2x \,, \quad z_y = a_3+2a_4y \tag{A.4.19}$$

$$z_{xx} = 2a_2 \,, \quad z_{yy} = 2a_4 \,, \quad z_{xy} = 0 \tag{A.4.20}$$

Thus, once a satisfactory surface approximation has been obtained, the fitting coefficients a_2 and a_4 are determined and the sign of the Gaussian curvature is invariant in the region:

$$sgn(K) = sgn(z_{xx}z_{yy}) = sgn(a_2a_4) \tag{A.4.21}$$

Meanwhile, the sign of the mean curvature H will be

$$sgn(H) = sgn(Ez_{yy}+Gz_{xx}) = sgn(a_4E+a_2G) \tag{A.4.22}$$

Indeed, it may vary over the region. However, as the 2nd order derivative $z_{xy} \equiv 0$ in (A.4.16), from lemma 5.6, the calculations of surface derivatives can be split into the calculations of the curve derivatives in the x- and y-directions respectively. From lemma 5.1 and 5.2, when the significance measure to the approximated curve is small, z_x and z_y are close to constants. So, both $E = 1+z_x^2$ and $G = 1+z_y^2$ will be close to *positive* constants in each section, and $sgn(H)$ becomes invariant then.

On the other hand, if the 2nd order basis function xy is appended as a term to the fitting bi-polynomial $z(x,y)$ in (A.4.16), $z(x,y)$ and its derivatives become

$$z(x,y) = a_0+a_1x+a_2x^2+a_3y+a_4y^2+a_5xy \tag{A.4.23}$$

$$z_x = a_1+2a_2x+a_5y \,, \quad z_y = a_3+2a_4y+a_5x \tag{A.4.24}$$

$$z_{xx} = 2a_2 \,, \quad z_{yy} = 2a_4 \tag{A.4.25}$$

The sign of the Gaussian curvature K remains the same as (A.4.21), but the sign of the mean curvature H changes to

$$sgn(H) = sgn(Ez_{yy}+Gz_{xx}-2Fz_{xy}) = sgn(a_4E+a_2G-a_5F) \tag{A.4.26}$$

Compared with (A.4.22), the term F may vary over the region, leading to $sgn(H)$ changing. A simple way to guarantee $sgn(H)$ being invariant is to require that $F \equiv z_xz_y$ is nearly constant in the region. This can be achieved by making both z_x and z_y be nearly constant in the region. As x and y vary over the region, this suggests $a_5 = 0$ in (A.4.24), and in (A.4.23) as well. That is, (A.4.23) returns to (A.4.16) where the xy term is excluded.

Q.E.D.

Corollary 5.2: *Using the bi-polynomial in (A 4 16) to approximate a smooth surface up to the degree of the signs of K and H, the derivatives can be calculated in x- and y- directions independently, i e , the 2-D surface fitting can be decomposed into 1-D parabolic curve fittings.*

Proof: From (A.4.16), $z_{xy} = z_{yx} = 0$. Thus, from lemma 5, the conclusion is obtained.

Q.E.D.

Appendix 5.

Other tests during region growing

Best fitting test

At each scale level the spline-smoothing function gives a root mean square error estimate. Related to this estimate, a threshold is assigned to all the surface points as the initial value of the fitting error tolerance. During the region fitting if the fitting error is less than the tolerance, it becomes the current value of the tolerance of this region. During later region growing, a candidate point is accepted only if its fitting error is less than the current tolerance of the region. So the segmentation is guided by the best fitting of the region

C^0 test

To be accepted into the growing region, the deviation of a candidate point from the predicted position should be less than a given threshold of the fitting error.

For a single scale segmentation the threshold can be determined based on the noise estimate of the raw data, such as the root mean square error or the maximal absolute error.

For segmentation at multiple scales, there are both noisy influences and surface distortions. The distortion continues when the noise vanishes So a simple noise estimation is not enough. The error from the distortion cannot be ignored. However, nothing is known about the distortion before a proper segmentation has been obtained. Hence, both noise and distortion errors are estimated together at each scale level. The estimate is the root mean square error of the difference between the data surface and its spline-smoothed version.

C^1 test

With only a C^0 fitting error test there is no guarantee as to which is the right surface extrapolation. For instance, at the top of a fairly flat roof, region growing of either fold could easily go some distance into the other fold instead of stopping right at the top. This "blind growing" risk yields a wrong location for the boundary A curvature extrema line can be obtained by a separate calculation as in [Ponce and Brady 1987] to prevent the region from growing beyond this line.

On the other hand, the existence of a roof patch prevents both sides of the roof from merging since a roof patch on the KH sign image always expands as the scale increases. The expansion can be captured by tracking across scale levels for such an indication. In this sense, the C^1 test is an unnecessary step, when a patch is semantically known as the roof patch.

C^2 test

Note that the least square fitting function is an analytic function and thus is a global function, whose performance in any small seed region is sufficient to determine the behaviour at a remote point. Hence, with the specially selected basis function set (*i.e.*, by imposing restrictions to the solution space), the C^0 or C^1 test could effectively prevent a region from blind growing in the discrete case. Therefore, the C^2 test seems unnecessary in practice.

Appendix 6.

Updating the Householder transformation for solving the least squares fitting equation in region growing

Due to noise, the shape of a region on the KH image is often irregular and may have holes inside. Fitting surface data on such a region is a difficult problem.

Traditionally, the least squares method is used to get the solution via the normal equation·

$$\mathbf{A}^T\mathbf{A}\mathbf{w} = \mathbf{A}^T\mathbf{z} \qquad \text{(A.6.1)}$$

which is derived from the fitting equation called the over-determined equation:

$$\mathbf{A}\mathbf{w} = \mathbf{z} \qquad \text{(A.6.2)}$$

where

$$\mathbf{A} \equiv \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \\ a_{n+1,1} & a_{n+1,2} & \cdots & a_{n+1,n} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix}, \quad \mathbf{w} \equiv \begin{bmatrix} w_1 \\ w_2 \\ \cdot \\ \cdot \\ \cdot \\ w_n \end{bmatrix}, \quad \mathbf{z} \equiv \begin{bmatrix} z_1 \\ z_2 \\ \cdot \\ \cdot \\ \cdot \\ z_n \\ z_{n+1} \\ \vdots \\ \vdots \\ z_m \end{bmatrix} \qquad \text{(A.6.3)}$$

$$[a_{i,1}\ a_{i,2}\ a_{i,3} \cdots\ a_{i,11}] \equiv [1\ x\ y\ xy\ x^2\ y^2\ xz\ yz\ z^2\ x^2y^2\ x^4\ y^4]_i \qquad \text{(A.6.4)}$$

and i denotes the i-th point in the region, m is the size of the current region and n is the size of the current basis function set, its range is $3 \leq n \leq 11$

Note that solution of the normal equation may be a seriously ill-posed numerical computation problem because the operation $\mathbf{A}^T\mathbf{A}$ not only introduces round-off errors but also squares the condition-number of the coefficient matrix. Bartles and Jezioranski [1979] proposed an efficient and general method to treat surface fitting on irregular regions, which can avoid these difficulties But it is preferable to use the least squares fitting method because it is probably the simplest method with fast performance on an irregular region despite not being perfect. The Householder transformation [Golub and Van Loan 1983][Wilkinson and Reinsch 1971][Press *et al* 1988] can be used to get the least squares solution directly from the fitting equation The normal equation is no longer required. Other methods such as QR and SVD can be adopted too (cf. [Golub and Van Loan 1983] etc.).

In spite of the simplicity and fast performance of least squares fitting, this computation is still intensive for the region growing process. During growing, least squares fitting will be exercised again and again to get the current fitting coefficients and fitting residuals. The latter is an indication for accepting or rejecting the tested candidate points. So, some intermediate results are temporally saved to reduce computation as described below.

To solve the fitting equation $Aw = z$ directly, the Householder transformation matrix **H** is applied to both sides of the fitting equation to eliminate all elements under the diagonal of the matrix **A**, so as to get the upper matrix $\mathbf{B} = \mathbf{HA}$ and the vector $\mathbf{y} = \mathbf{Hz}$ as the intermediate results, where

$$\mathbf{Bw} = \mathbf{y} \tag{A.6.5}$$

or

$$\begin{bmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,n} \\ & b_{2,2} & \cdots & b_{2,n} \\ & & \ddots & \vdots \\ & & & b_{n,n} \\ & & & b_{n+1,n} \\ & & & \vdots \\ & & & b_{m,n} \end{bmatrix} \mathbf{w} = \begin{bmatrix} y_{1,n} \\ y_{2,n} \\ \vdots \\ y_{n,n} \\ y_{n+1,n} \\ \vdots \\ y_{m,n} \end{bmatrix} \tag{A.6.6}$$

The residual is determined by (A 6 6) as $r^2 = b_{n+1,n}^2 + b_{n+2,n}^2 + \cdots + b_{m,n}^2$, and the least squares solution $\mathbf{w} = [\, w_1 \;\; w_2 \;\; \cdots \;\; w_n \,]^T$ is obtained from further solving the upper sub-matrix equation.

The growing continues successively. When a phase of growing is finished, r row vectors corresponding to r accepted points will be appended to the top of the matrix **B**. The vector **y** is updated as well. Only an $(n+r)\times n$ upper sub-matrix of the $(n+m)\times n$ updated matrix **B** is recomputed to get a new intermediate result. This treatment significantly reduces the computation of the updating when the size of the growing region m is large (*i.e* , the size of the matrix **B** is large).

Bibliography:

[AHLB67] Ahlberg, J.H., Nilson, E N., and Walsh, J.L., *Theory of Splines and Its Applications*, Academic Press, New York, 1967.

[AMES77] Ames, W.F., *Numerical Methods for Partial Differential Equations*. Academic Press, New York, 1977, 2nd edition, pp. 47.

[ASAD84] Asada, H. and Brady, M., *The Curvature Primal Sketch*, Technical Report AIM-758, MIT, AI Lab., 1984.

[BABA86] Babaud, J., Witkin, A P., Baudin, M and Duda, R.O., *Uniqueness of the Gaussian Kernel for Scale-Space Filtering*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-8, No 1, 1986.

[BALL82] Ballard, D.H., and Brown, C M., *Computer Vision*, Prentice-Hall Inc., Englewood Cliffs, New Jorsey, USA, pp 2-6, 341-344.

[BARL79] Barlow, H. and Reeves, B., *The Versatility and Absolute Efficiency of Detecting Mirror Symmetry in Random Dot Displays*, Vision Research, 1979, 19, 783-793.

[BART85] Bartles, R.H, and Jezioranski, J J., *Least-square fitting using orthogonal multi-nomials*, ACM Trans. Mathematical Software, **11**, 3 (Sept), 1985, 201-217.

[BERT87] Bertero, M., Poggio, T. and Torre, V., *Ill-posed Problems in Early Vision*, MIT A.I. Memo N0. 924, May 1987.

[BESL86a] Besl, P.J., and Jain, R C., *Invariant Surface characteristics for three dimensional object recognition in range images*, Computer Vision, Graphics, Image Processing **33**, *1 (January), 33-80*

[BESL86b] Besl, P.J., *Surfaces in Early Range Image Understanding*, PhD Dissertation, Electrical Engineering and Computer Science Department (RSD-TR-10-86), University of Michigan, 1986

[BLAC86] Black, A., and Zisserman, A., *Weak Continuity Constraints in Computer Vision*, Report CSR-197-86, Dept of Computer Science, Edinburgh University, UK, 1986.

[BRAD82] Brady, M., *Computational approaches to image understanding* ACM computing Survey **14**, 1 (Mar), 3-17.

[BRAD84] ________, *Representing Shape*, Proc. of IEEE 1984 International Conference on Robotics, pp. 256-265.

[BRAD85] Brady, M., Ponce, J , Yuille, A. and Asada, H , *Describing Surface*, MIT A.I. Memo No. 822, 1985 pp. 19-20.

[BRAK78] Brakk, K.A., *Mathematical Notes*, Princeton University Press, Princeton, NJ, 1978. Vol. **20**

[BROU84] Brou, P., *Finding the orientation of objects in vector maps*. Int. J. Rob. Res. 3(4), 1984. pp. 89-175.

[BRUC89] Bruce, V. and Burton M., *Computer Recognition of Faces*, Handbook of Research on Face Processing, A.W. Yong and H.D. Ellis (Eds.), Elsevier Science Publishers B.V. North-Holland, 1989. pp. 487-506.

[CAI87a] Cai, L.-D., *Diffusion Smoothing on Dense Range Data*, DAI WP-200, Department of A.I., University of Edinburgh, 1987.

[CAI87b] ________, *Some Notes On Repeated Averaging Smoothing,* DAI RP-337, Department of A.I., University of Edinburgh, 1987. Also in Proc. of BPRA 4th Int. Conf. on Pattern Recognition, Cambridge, U.K., 1988, *Lecture Notes in Computer Science, J.Kittler (Ed), Vol. 301*, Springer-Verlag, pp. 596-605.

[CAI88a] ________, *A "Small Leakage" Model for Diffusion Smoothing of Image Data,* DAI (resubmitted) RP-418, Department of A.I., University of Edinburgh, 1988. Also in Proc. of the 11th IJCAI Conference, Detroit, Michigan, USA, August, 1989, pp. 1585-1590.

[CAI88b] ________, *A Diffusion Smoothing Approach to Sculptured Surfaces.* DAI RP-406, Department of A.I., University of Edinburgh, 1988 Also in Proc. of IMA 3rd Int. Conf. on Mathematics of Surfaces, Oxford, U.K., September, 1988, pp. 267-286.

[CAI88c] ________, *Spline Smoothing A Special Case of Diffusion Smoothing,* DAI RP-435, Department of A.I., University of Edinburgh, 1988. Also in Proc. of the 5th AVC conference, Reading, UK, July, 1989, pp. 273-276.

[CAI88d] ________, *An Estimate of the Relationship between Zero Thresholds of Gaussian Curvature and Mean Curvature*. DAI RP-450, Dept of A.I., University of Edinburgh, January 1989. Also forthcoming in Proc. of the 4th IMA Conference on the Mathematics of Surfaces, Bath, UK, September, 1990.

[CAI89a] ________, *Approximating a surface up to curvature signs using depth data alone,* DAI RP-452, Department of A.I., University of Edinburgh, 1989. Also in Proc. of the SPIE/SPSE Symposium on Curves and Surfaces in Computer Vision and Graphics, Santa Clara, California, February 11-16, 1990.

[CANN83] Canny, J.F, *Finding Edges and Lines in Images,* MIT A.I. Tech. Report 720, 1983. pp. 72-75.

[CHER57] Chern, S.S., *A Proof of the uniqueness of Minkowski's problem for convex surfaces* Am. J. Math. **79**, 949-950.

[CHER67] ________, *Curves and Surfaces in Euclidean Spaces,* Studies in Global Geometry and Analysis, MAA Studies in Mathematics, The Mathematical Association of America, 1967.

[CLEN60] Clenshaw, C.W., *Curve Fitting with a Digital Computer,* Comput. J., **2**, 1960, pp. 170-171.

[CLAR86] Clark, H.H., and Wilkes-Gibbs, D., *Referring as a collaborative process,* Cognition, **22** (1986), 1-39.

[COUR53] Courant, R. and Hilbert, D., *Methods of Mathematical Physics,* Vol. 1. Wiley (Interscience), New York, 1953.

[CRAN47] Crank, J. and Nicolson, P., *Proc Camb Phil Soc*, **32**, 50, 1947.

[DAVI87] Davies, E.R., *Design of optimal Gaussian operators in small neighbourhoods,* Image and Vision Computing, Vol. 5, No. 3, August, 1987. pp. 199-205.

[DAVI75] Davis, P.J., *Interpolation and Approximation.* Dover, New York, 1975.

[DEO74] Deo, N., *Graph Theory with Applications to Engineering and Computer Science,* Prentice-Hall Inc., Englewood Cliffs, N.J., 1974.

[DoCA76] Do Carmo, M.P., *Differential Geometry of Curves and Surfaces,* Prentice Hall, Englewood Cliffs, 1976.

[DOUG56] Douglas, J.Jr. and Rachford, H.H.Jr., *Trans. Am Math Soc.,* **82**, 421, 1956.

[DOUG61] Douglas, J.Jr., *J. Soc. Ind. Appl. Math.*, **9**, 433, 1961.

[DUDA73] Duda, R.D., and Hart, P.E., *Pattern Classification and Scene Analysis,* New York: Wiley, 1973.

[ELFF76] Elffers, J., *Tangram The ancient Chinese shapes game*. New York: McGraw-Hill, 1976.

[ESHE86] Eshera, M.A., and Fu, K.S., *An Image Understanding System Using Attributed Symbolic Representation and Inexact Graph-Matching*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-8, No. 5, 1986.

[FAUG83] Faugeras, O.D , Hebert, M. and Pauchon, E., *Segmentation of range data into planar and quadric patches* in Proc. of Computer Vision and Pattern Recognition Conference (Arlington, Va.), 1983. pp. 8-13.

[FAUX79] Faux, I.D. and Pratt, M.J., *Computational Geometry for Design and Manufacture,* Ellis Horwood, Chichester, U.K 1979.

[FENG78] Feng, K. *et al., Numerical Computation Methods* (in Chinese), Defence Industry Press, China, 1978. pp 489-509.

[FISC83] Fischler, M.A , and Bolles, R.C., *Perceptual Organization And The Curve partitioning problem*, Proc. of 8th IJCAI conference, 1983, pp. 1014-1018.

[FISH89] Fisher, R.B., *From Surfaces to Objects*, John Wiley & Sons, Chichester, UK. 1989. pp 40-42.

[GILB83] Gilbarg, D. and Trudinger, N., *Elliptic Partial Differential Equations of Second Order*, Springer-Verlag, Berlin, West Germany, 1983

[GOLD88] Goldgof, D.B., Lee, H. and Huang, T.S., *Motion analysis of nonrigid surfaces*, in Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1988, pp. 375-380

[GOLU83] Golub, G H., and Van Loan, C.F., *Matrix Computations*, John Hopkins University Press, Baltimore, USA, 1983.

[GONZ87] Gonzalez, R.C., and Wintz, P., *Digital Image Processing*, Addison-Wesley Publishing Company, Reading, Massachusetts, USA, 1987.

[GOOD68] Goodman, J.W , *Introduction to Fourier Optics*, McGraw-Hill, New York, 1968.

[GOUR85] Gourlay, A.R., *Implicit convolution,* Image and Vision Computing, Vol. 3 No. 1, 1985, pp. 15-23

[HARA81] Haralick, R.M. and Watson, L.T., *A Facet Model for Image Data.* Computer Graphics Image Processing **15**, 1981, 113-129.

[HARA83] Haralick, R.M., Watson. L.T., and Laffley. T.J., *The topographic primal sketch* , Int. J. Rob. Res. 2(1)· 50-72, 1983.

[HEBE82] Hebert, M. and Ponce, J., *A New Method For Segmenting 3D Scenes Into Primitives* In Proceedings of 6th International Conference Pattern Recognition, Munich, West Germany, Oct. 19-22, 1982. pp. 836-838.

[HELD83] Heldreth, F.C., *The Detection of Intensity Changes by Computer and Biological Vision Systems*, Computer Vision, Graphics, and Image Processing **22**, 1-27, 1983.

[HOFF82] Hoffman, D.D., & Richards, W.A., *Representing Smooth Planar curves for Recognition: Implications for Figure-ground Reversal*, Proc. of USA National conference for AI, 1982, pp. 5-8.

[HORN83] Horn, B.K.P., *Extended Gaussian Images*, MIT A.I. Memo No. 740, July 1983.

[HSIU81] Hsiung, C.C., *A first course in differential geometry*, Wiley-Interscience, New York, 1981.

[HUMM84] Hummel, R.A., Kimia, B., and Zucker, S.W., *Deblurring Gaussian Blur*, McGill University, Montrel, P.Q., Canada, Tech. Rep. 83-13R, 1984.

[HUMM87] Hummel, R., and Moniot, R., *Solving Ill-conditioned Problems by Minimizing Equation Error*, In Proceedings of 1st International Conference on Computer Vision, June 8-11, 1987, London, England, pp 527-533.

[HURL86] Hurlbert, A., and Poggio, T., *Visual Attention and Computer*, **Nature** 321, 12 June, 1986, pp. 651-652.

[IKEU84] Ikeuchi, K , Horn, B.K.P., *et al*, *Picking up an object from a pile of objects*, 1st Int. Sym. on Rob. Res. *MIT Press, 1984, pp 139-162.*

[KOEN84] Koenderink, J.J., *The Structure of Images*, Biological Cybernetics **50**, 1984, pp. 363-370.

[KOEN86] Koenderink, J.J., and van Doorn, A J., *Dynamic Shape*, Biological Cybernetics **53**, 1986. pp. 383-396.

[LEDL64] Ledley, R.S., *High Speed Automatic Analysis of Biomedical Pictures*, Science, Vol. 146, No. 3641, pp. 216-223, 1964

[LI78] Li, Y.S., and Huang, Y.Q., *Numerical Approximation (in Chinese)*, Education Press, China, 1978.

[LIGH60] Lighthill, M.J., *Introduction to Fourier Analysis and Generalized Functions*, Cambridge University Press, London and New York, 1960.

[LINK85] Link, N.K., and Zucker, S.W , *Sensitivity to Corners in Flow Patterns*, Computer Vision & Robotics Laboratory, Dept of Electrical Engineering, Mcgill University, TR-85-4R, April 1985.

[LITT85] Little, J.J., *Recovering shape and determining attitude from extended Gaussian images*. Tech. Report, Univ. of British Columbia, Dept. of CS, 1985.

[LOWE87] Lowe, D.G., *Three-dimensional object recognition from single two-dimensional images*, Artificial Intelligence, Vol. 31, pp. 355-395, 1987.

[MACC86] Maccallum, K.J , and Zhang, J.-M., *Curve Smoothing Techniques Using B-splines*, The Computer Journal, Vol. 29, No. 6, 1986, pp. 564-571.

[MACL72] Maclead, I.D.G., *Comments on Techniques for Edge Detection*, in Proc. of the IEEE60, 344, 1972.

[MAKI67] Makinson, G.J., Generalized Least Squares Fit by Orthogonal Polynomials, Comm. ACM., **10**, 1967, pp. 87-88.

[MARI84] Marimont, D H., *A Representation for Image Curves*, in Proc. of the Fifth AAAI National Conference on Artificial Intelligence, Austin, Texas, USA, Augest 6-10, pp. 237-241.

[MARR76] Marr, D., *Early processing of visual information*, Phil. Trans. Roy. Soc. B275: 843-524, 1976.

[MARR77] _______, *Artificial Intelligence* — a personal view. *Artificial Intelligence*, **9**, *1977, pp. 37-48*

[MARR78] Marr, D., and Nishihara, H.K., *Representation and recognition of the spatial organization of three-dimensional shapes*, Proc. R.. Soc. Lond. B200 (1978), 269-294.

[MARR80] Marr, D. and Hildreth, F.C., *Theory of Edge Detection*, Proc. R. Soc. Lond. B207 (1980) 187-217.

[MARR82] Marr, D., *Vision*, Freeman: new York, 1982.

[MEDI84] Medioni, G., and Nevatia, R., *Description of 3-D surfaces using curvature properties*, in Proc. of Image Understanding Workshop, New Orleans, La., October 1984, DARPA, pp. 291-299.

[MOKH86] Mokhtarian, F. and Mackworth, A., *Scale-Based Description and Recognition of Planar Curves and Two-Dimensional Shapes*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-8, No. 1, 1986.

[PAVL72] Pavlidis, T., *Segmentation of pictures and maps through functional approximation*, Computer Graphics, Image processing **1**, 360-372.

[PAVL80] Pavlidis, T and Panfilov, A.V., *Algorithms for shape analysis of contours and waveforms*, IEEE Trans. Pattern Anal Machine Intell., Vol. PAMI-2, July, 1980, pp. 301-312.

[POGG84] Poggio, T., Voorhees, H., and Yuille, A.L., *A Regularized Solution to Edge Detection*, MIT AI Lab. Memo 733, 1984.

[PONC87] Ponce, J. and Brady, M., *Toward a Surface Primal Sketch*, in Three-Dimensional Machine Vision, T. Kanade (Ed.), Kluwer Academic Publishers, 1987, pp 225-227

[PONG84] Pong, T.C., Shapiro, L.G., Watson, L.T. and Haralick, R.M., *Experiments in segmentation using a facet model region grower*, Computer Vision, Graphics, and Image Processing, **25**, 1-23.

[POST78] Poston, T., and Stewart, I., *Catastrophe Theory and its Applications*, Pitman Publishing Limited, London, 1978, pp. 291, pp 351

[PORR88] Porrill, J., Pollard, S B , Pridmore, T.P., Bowen, J.B., Mayhew, J.E.W. and Frisby, J.P., *TINA a 3D vision system for pick and place*, Image and Vision Computing, vol. 6, No. 2, May, 1988, pp 91-99.

[PRAT78] Pratt, W., *Digital Image Processing*, John Wiley and Sons, New York, 1978.

[PRES88] Press, W.H *et al* , *Numerical Recipes in C*, Cambridge University Press, New York, 1988.

[ROSE76] Rosenfeld A and Kak A.C., *Digital Picture Processing*, Academic Press, New York, 1976, pp 174-178.

[ROSI88] Rosin, P.L., and West, G A.W., *Detection of Circular Arcs in Images*, in Proc. of the Fourth Alvey Vision Conference, University of Manchester, 31 August - 2 September, 1988, pp. 259-263.

[REIN67] Reinsch, C.H., *Smoothing by spline functions*, Num. Math. **10**, 177-183, 1967.

[SCOT88] Scott, G.L., Turner, S.C., and Zisserman, A , *Using a Mixed Wave/Diffusion Process to elicit the Symmetry Set* Proc. of The Fourth Alvey Vision Conference, University of Manchester, 31 August - 2 September, 1988, pp.

[SHAP81] Shapiro, L.G., and Haralick, R.M., *Structural Descriptions and Inexact Matching*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-3, No. 5, 1981, pp. 504-519.

[SNYD85] Snyder, W. and Bilbro, G., *Segmentation of Three-Dimensional Images*, in Proc. of International Conference on Robotics and Automation (St. Louis, Mo., March 25-28, 1985). IEEE-CS, New York, pp. 396-403.

[STEV81] Stevens, K.A., *The visual interpretation of visual surface pieces*, Artificial Intelligence, **17**, pp. 47-75.

[TERZ85] Terzopoulos, D., *Computing Visible-Surface Representations*, MIT A.I. Memo No. 800, 1985, pp. 12-17.

[TORR84] Torre, V., and Poggio, T., *On Edge Detection*, MIT A.I. Lab. Memo 768, 1984.

[WANG81] Wang, D.C.C., Vagnucci, A.H. and Li, C.C., *Gradient Inverse Weighting Smoothing Schema and the Evaluation of its Performance* Computer, Graphics and Image Processing **15**, 167-81, 1981.

[WARR78] Warrington E.K. and Taylor, A M., *Two categorical stages of object recognition* Perception 7, 1978, 695-705.

[WATT82] Watt, R.J., and Andrews, D.P., *Contour curvature analysis: hyperacuities in the discrimination of detailed shape*, Vision Research, **22**, pp. 449-460.

[WILK71] Wilkinson, J.H. and Reinsch, C., *Linear Algebra*, Vol. II of *Handbook for Automatic Computation* New York: Springer-Verlag, 1971.

[WILS79] Wilson, H.R., and Bergen, J.R., *A Four Mechanism Model for Threshold Spatial Vision*, Vision Research **19**, 19-32, 1979

[WIST77] Winston, P.H., *Artificial Intelligence*, Addison-Wesley Publishing Company, USA, 1977, pp. 18-26.

[WITK83a] Witkin A.P., *Scale-space Filtering*, in Proc. of International Joint Conference on Artificial Intelligence, 1983, pp 1019-1022.

[WITK83b] Witkin A.P. and Tenenbaum J.M., *On the Rule of Structure in Vision*, in Human and Machine Vision, Beck, Hope and Rosenfeld (Eds.), Academic Press, 1983, pp 481-543.

[YANG88] Yang, H.S., *Range Image Segmentation and Classification via Split-and Merge Based on Surface Curvature*, in Proc. of BPRA 4th Int. Conf. on Pattern Recognition, Cambridge, U.K , 1988, *Lecture Notes in Computer Science, J Kittler (Ed), Vol 301*, Springer-Verlag. pp. 58-67.

[YUIL83] Yuille A.L. and Poggio T., *Scaling theorems for zero-crossings*, MIT AI Memo No. 722, 1983.

[YUIL88] Yuille, A.L., *The creation of structure in dynamic shape*, in Proc. of 2nd International Conference on Computer Vision, June 8-11, 1988, London, England, pp. 527-533.

[ZUCK83] Zucker, S.W., *Computation and Psychophysical Experiments in Grouping: Early Orientation Selection* in Human and Machine Vision, Back, Hope and Rosenfeld Eds. New York: Academic Press, 1983.