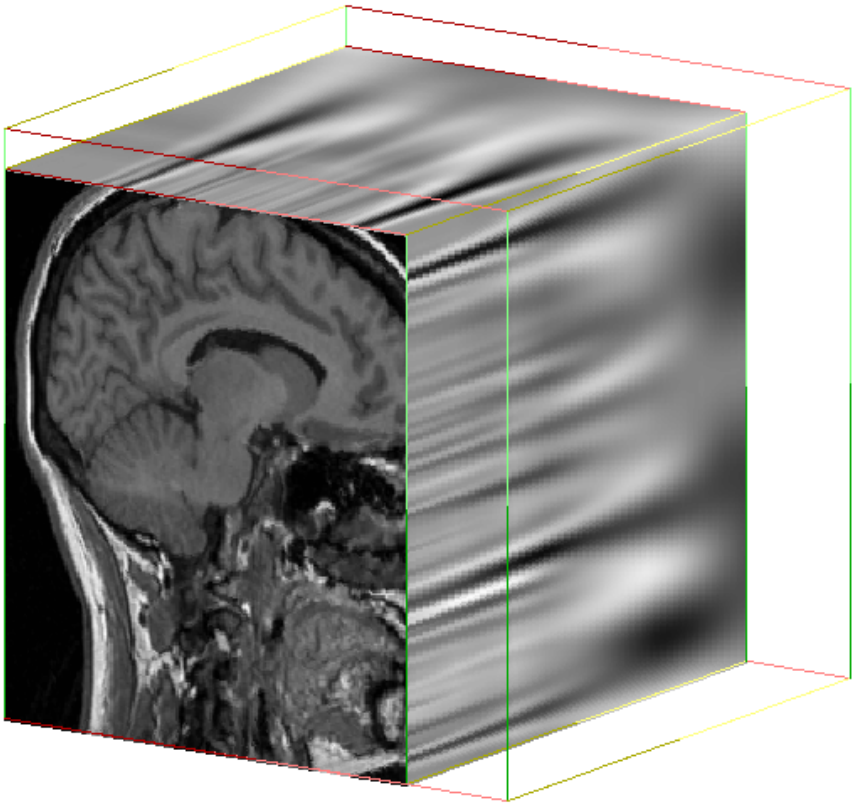
Introduction to Scale-Space Theory: Multiscale Geometric Image Analysis

Tutorial VBC ’96, Hamburg, Germany

Fourth International Conference on Visualization in Biomedical Computing



Bart M. ter Haar Romeny, PhD
Utrecht University, the Netherlands
bart@cv.ruu.nl

22 September 1996

Technical Report No. ICU-96-21, Utrecht University

1 Introduction

Multiscale image analysis has gained firm ground in computer vision, image processing and modeling biological vision. However, the approaches have been characterised by a wide variety of techniques, many of them chosen ad hoc. Scale-space theory, as a relatively new field, has been established as a well founded, general and promising multiresolution technique for image structure analysis, both for 2D, 3D and time series.
The rather mathematical nature of many of the classical papers in this field has prevented wide acceptance so far. This tutorial will try to bridge that gap by giving a comprehensible and intuitive introduction to this field. We also try, as a mutual inspiration, to relate the computer vision modeling to biological vision modeling. The mathematical rigor is much relaxed for the purpose of giving the broad picture. In appendix A a number of references is given as recommended reading[1]. This work is based on the contributions of many, as mentioned in the acknowledgements.

2 The multiscale nature of things

In **mathematics** objects have no scale. We are familiar with the notion of points, that really shrink to zero, lines of zero width. No metrical *units* are involved in mathematics, as in physics. Neighborhoods, like necessary in the definition of differential operators, are taken into the limit to zero, so we can really speak of *local operators*.

In **physics** objects live on a *range* of scales. We need an instrument to do an observation (our eye, a camera) and it is the range that this instrument can see that we call the scale range. To expand the range of our eye we have a wide armentarium of instruments available, like microscopes and telescopes. The scale range known to humankind spans about 50 decades, as is beautifully illustrated in the book (and movie) "Powers of Ten" [Mor85]. The scale range of an instrument is bounded on two sides: the *inner scale* is the smallest detail seen by the smallest aperture (e.g. the CCD element, a cone or rod); the *outer scale* is the coarsest detail that can be discriminated, i.e. it is the whole image (field of view). Dimensional units are *essential*: we express any measurement in these units, like meters, seconds, candelas, ampères etc.

In **front-end vision** the apparatus (starting at the retina) is equipped just to extract multiscale information. Phychophysically is has been shown that the threshold modulation depth is constant (within 5%) over more than two decades, so the visual system must be equipped with a large range of sampling apertures. There is abundant electrophysiological evidence that the receptive fields[2] (RF's) in the retina come in a wide range of sizes [HW62, HW79, Hub88, Pap94, BI94, Zek93, SMKL93].

In any image analysis there is a *task*: Here the notion of scale is often an essential part of the question too: "Do you want to see the leaves or the tree"?

3 Linear Scale-Space Theory

[1] These tutorial notes do in no way pretend to be complete. These notes are an excerpt of a forthcoming introductory book on scale-space theory and front-end vision [tHR97], appearing Spring or Summer 1997.

[2] It is not so that every receptor in the retina (rod or cone) has its own fiber in the optic nerve to further stages. In a human eye there are about 150.10^6 receptors and 10^6 optic nerve fibres. Receptive fields form the elementary 'apertures' on the retina: they consist of many cones (or rods) in a roughly circular area projecting to a single (ganglion) output cell, thus effectively integrating the luminance over a finite area.

3.1 Physics of Observation

If we start with taking a close look at the observation process, some elementary questions come up:
• What is an image anyway?
• What do we *mean* with the 'structure of images' (Koenderink 1984 [Koe84])?
• How good should a measurement (observation) be?
• How accurately can we measure?
• What are the best apertures to measure with?
• How do we incorporate the notion of scale in the mathematics of observation?
• Does the visual system make *optimal* measurements?

Any physical observation is done through an aperture. By necessity this aperture has to be *finite* (would it be zero no photon would come through). We can modify the aperture considerably by using instruments, but never make it of zero width. There is no such thing as a physical 'point'.
We consider here physical observations that are *uncommitted*, i.e. there is no model involved. Later we will fully incorporate the notion of a model, but in the first stages of observation *we know nothing*. This is an important notion, it will lead to the establishment of *linear* scale-space theory [FtHRKV94b, tHRF93, LtHR94, Lin94]. It is a natural requirement for the first stage, but *not* for further stages, where extracted information, knowledge of model and/or task comes in etc. We then come into the important realm of *nonlinear* scale-space theory, which will be discussed in section 7.
A single constant size aperture function may be sufficient in a controlled physical application (e.g. fixed resolution device), in the most general case *no a priori size* is determined. In some way control is needed.
Counterintuitive may seem the notion that noise is always part of the observation. It can only be extracted from the observation if we have a model of the observed entity, or both. Very often this is not considered explicitly. We e.g. assume the object comes from the 'blocks world' so it has straight contours, but often this is not known.
In mathematics the smallest distance can be considered by taking the limit to zero, but in physics this reduces to the finite aperture separation distance (sampling distance). Therefore we may foresee serious problems when we deal with such notions as differentiation (especially for high order), and sub-pixel accuracy.

To compute any type of representation from the image data, information must be extracted using certain *operators* interacting with the data. Basic questions then are: What operators to use? Where to apply them? How should they be adapted to the task? How large should they be?
In the next paragraph we consider the aperture function as an operator: we will search for *constraints* to pin down the exact specification of this operator. As we will see, for an unconstrained front-end there is a unique solution for the operator: the Gaussian kernel.

3.2 The aperture function of an uncommitted front-end

What is the best aperture function? How can it be derived? To answer these questions, we first establish the requirements and constraints appropriate for the physical situation at hand.
Let's give an example where things go wrong: if we take a square aperture, and make larger copies of it to sample at larger scales, we get the tesselation effect seen in figure 1. Koenderink coined this *spurious resolution* [Koe84], i.e. the emergence of details that were not there before. As a general rule we want the information content only to decrease with larger aperture.
The following line of reasoning is due to Florack et al. [FtHRKV94b]. The requirements can be stated as *axioms*, or postulates for an uncommitted visual front-end. In essence, as stated before, it is the mathematical formulation for "we know nothing, we have no preference whatsoever":

Figure 1: Spurious resolution due to square apertures. Sagittal MR sampled at 3 different aperture sizes, expressed in the pixelsize of the left image: size 1×1 pixel (left); size 4×4 pixels (middle); size 8×8 pixels (right).

- linearity (no knowledge, no model, no memory)
- spatial shift invariance (no preferred location)
- isotropy (no preferred orientation)
- scale invariance (no preferred size, or scale)

The constraints for the physical sampling system are derived by realizing that the description of *any* physical system must be described in a way independent of a particular choice of coordinate system. If we change the coordinate system, then the description must describe the same system. We want to be *invariant* under the actions of a particular coordinate transformation. There are many *groups* of coordinate transformations (see section 4 for an overview of most relevant groups). We deal with medical images mostly, and we consider the group of orthogonal coordinate transformations: translations and rotations. Further, we want no particular knowledge or model involved, so the system must be linear. Any knowledge or model would just *be* the nonlinearity, as we will see later in the nonlinear scale-space theories.

We call the construct a scale-space, when a true simplification of the image occurs. The constraint of scale invariance is the most important. It enables the action of two operators to be equavalent to the same action done by one operator. This known as the *semigroup* property. A consequence is that we have one way traffic here: it is impossible to reverse the action, and find the two original operations again.

3.3 Scale and Dimension

Every *physical* unit has a *physical dimension*, and it is this that mostly discriminates physics from mathematics. It was Baron Jean-Baptiste Fourier who already in 1822 established the concept of dimensional analysis. This is indeed the same mathematician Fourier, so famous for his Fourier transformation. He described the concept of dimension in his memorable work entitled *"Théorie analytique de la chaleur"* [Fou55] as follows:

> It should be noted that each physical quantity, known or unknown, possesses a *dimension* proper to itself and that the terms in an equation cannot be compared one with another unless they possess the same *dimensional exponent*.

A physicist always checks first, when he inspects a new formula whether the dimensions are correct. It is impossible to add meters to meters/second.

The law of scale invariance indicates that we have a full freedom of reparametrization:

Proposition 1 (Law of Scale Invariance) *Physical laws must be independent of the choice of fundamental parameters.*

Dimensional analysis is an elegant tool to find out basic relations between physical entities, or even to *solve* a problem. It is often a method of first choice, when no other information is available. It is often quite remarkable how much one can deduct by just using this technique. We will use dimensional analysis to establish the expression defining the basic linear isotropic scale-space kernel.
First we set up a matrix of all dimensional units and physical variables. We consider the N-dimensional case, in the Fourier domain (this will turn out to lead to easier expressions). The variables are: σ the 'width' of the aperture function in the spatial domain, $\vec{\omega}$ the spatial frequency vector, L_0 the luminance distribution of the outside world, L the luminance after observation through the aperture. The isotropy axiom states that there is no preference for direction, so only the length of the vector $\vec{\omega}$ matters: $\sigma\vec{\omega} \Rightarrow \sigma\omega$, which is a scalar.

	σ	ω	L	L_0
length	$+1$	-1	0	0
intensity	0	0	$+1$	$+1$

Note that we have 2 fundamental physical quantities and 4 derived quantities. We can make a number of dimensionless combinations. Dimensional analysis states that the dimensionless quantities should be functions of each other, which may be raised to arbitrary power. There is a theorem in physics, the Pi Theorem, that states how many dimensionless quantities can maximally be formed, given the matrix of unit dimensions and physical variables: this maximum number is the number of physical variables minus the rank of the matrix. In our case: 4-2=2.
For the two independent dimensionless quantities we choose $\frac{L}{L_0}$ and $\sigma\omega$, so

$$\frac{L}{L_0} = G((\sigma\omega)^p)$$

G is the function (operator, kernel) to be determined.
Spatial shift invariance implies convolution, i.e. we scan the aperture over any possible location of the image:

$$L(\vec{x},\sigma) = L_0 * G(x;\sigma)$$

with $*$ defined as the convolution operation

$$f * g = \int_{-\infty}^{\infty} f(u)g(x-u)du.$$

In the Fourier domain convolution becomes multiplication:

$$L(\omega;\sigma) = L_0(\omega) \cdot G(\omega;\sigma)$$

When we study the asymptotic behaviour of the system at the border scales, we find at the inner scale that the image is not scaled at all:

$$\lim_{\sigma\omega\downarrow 0} G(\sigma\omega) \to 1$$

and at the outer scale complete spatial averaging occurs:

$$\lim_{\sigma\omega\to\infty} G(\sigma\omega) \downarrow 0 \tag{1}$$

The most important constraint is the requirement that the operation must be *self-similar*: the observation through an aperture increases the inner scale, and when we observe this observation again, the inner scale is again increased with the scale of the aperture. The total scaling of such a *cascade* of scalings must be consistent with performing just a singe rescaling. So, a concatenation of two rescalings $G(\sigma_1\omega)$ and $G(\sigma_2\omega)$ should be a rescaling $G(\sigma_3\omega)$ with the effective scale parameter $\sigma_3 = \sigma_2 \oplus \sigma_1$. Note that $\oplus$ need not to be ordinary addition. In mathematical terminology, the operator $\oplus$ and the group of positive numbers constitute a *commutative semigroup*. A result from the theory of semigroups is that

$$G(\omega\sigma_1)G(\omega\sigma_2) = G((\omega\sigma_1 + \omega\sigma_2)^p)$$

A general solution to this constraint is:

$$G(\omega\sigma) = \exp((\alpha\omega\sigma)^p)$$

Here we see the emergence of the exponential function. If we require the D spatial dimensions to be *separable*, we have

$$G(\omega\sigma) = \Pi_{i=1}^{D} G(\omega_i\sigma_i)$$

This makes σ to a real length, so we are able to determine the effective σ from the effective scalings along the coordinate axes, i.e.. scaling in each direction separately is the same as an isotropic scaling with $G(\omega\sigma)$. This fixes $p = 2$, because the length of the total projection vector is calculated by the Pythagoras formula from the magnitude of the constituent projection vectors $(\vec{\omega}\sigma \cdot \hat{e}_i)\hat{e}_i$. We demand a *real* solution, so α^2 is real. Because of (1) α^2 is negative. We are free to choose $\alpha^2 = -\frac{1}{2}$. So finally we get the *normalized Gaussian kernel* in the Fourier domain:

$$G(\vec{\omega}, \sigma) = \exp\left\{-\frac{1}{2}\sigma^2\omega^2\right\}$$

which is in the spatial domain

$$G(\vec{x}, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}^D} \exp\left\{-\frac{\vec{x}\cdot\vec{x}}{2\sigma^2}\right\}$$

So we have derived the Gaussian kernel from a set of axioms as the *unique* kernel for an uncomitted front-end. Note that this is only true for the choice of uncommittment: when we relax that requirement, other kernels are possible [PFMvG93, FtHRKV92].

Convolution with a Gaussian necessarily increases the inner scale: the Gaussian is the *operator* that transforms the inner scale of the image. The *cascade property* states that it is the same if one reaches a final certain scale in a single step from the input image by a given Gaussian aperture, or apply a sequence of Gaussian kernels, to reach the same scale.
The stack of images as a function of increasing inner scale is coined a linear 'scale-space' (Witkin 1983 [Wit83], Koenderink 1984 [KvD94]), see figure 2.
It was first realized by Koenderink [Koe84] that the generating equation of a linear scale-space is the linear *diffusion equation*:

$$\frac{\partial L}{\partial s} = \vec{\nabla}\cdot\vec{\nabla}L = \Delta L = L_{xx} + L_{yy} \tag{2}$$

(for 2D), stating that the derivative to scale equals the divergence of the gradient of the luminance function, which is the Laplacean, the sum of the second partial derivatives. The Gaussian is the Green's function of the diffusion equation [BCHSL92]. When the diffusion is equal for all directions, i.e. the sigma's of the Gaussian are equal, we call the process *isotropic*. When the diffusion is equal for every point of the image, we call the process *homogeneous*. Because of the diffusion equation, the process of generating a multiscale representation is also known as *image evolution*.

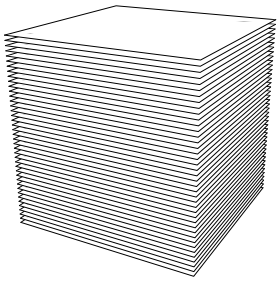


Figure 2: A scale-space of an image is the stack of images at all scales between the inner and outer scale of the image.

The derivation of the diffusion equation has been done in many ways [BWBD86], among which we mention:
• Causality [Koe84]: coarser scales can only be 'caused by' what happened at finer scales;
• Maximum principle [HG84]: at any scale change the maximal luminance at the coarser scale is always lower then the maximum intensity at the finer scale, the minumum is always larger.
• No new extrema at larger scales [BWBD86, Lin90]: this holds *only* for one-dimensional signals.
• Physics of luminance diffusion: the decrease of the luminance with time (or scale, which is here equivalent) is equal to the divergence of a *flow*, i.e. equal to the divergence of the luminance gradient [Koe84].

Note that the s in the diffusion equation has the dimension of the *squared* spatial dimension, so it takes the role of the Gaussian variance. The relation to the standard deviation is: $\sigma^2 = 2s$.

3.4 Gaussian derivatives and regularization

The Gaussian kernel is now established as the unique scale-space operator to change scale. There is an important additional result: one of the most useful results in linear scale-space theory is that the spatial derivatives of the Gaussian are solutions of the diffusion equation too, and together with the zeroth order Gaussian they form a *complete family of differential operators*.
If we want to take the derivative of an *observed* image, i.e. the convolution[3] of the outside world image with the aperture function, we get the following result due to the fact that we may commute the differential and the convolution operators:

$$\frac{\partial}{\partial x}(L * G) = L * \frac{\partial G}{\partial x}$$

So the derivative is found by convolving the image with the derivative of a Gaussian. This means that the derivative is given at a given *scale*: We call the derivative operator the *scaled derivative operator* or *Gaussian derivative operator*. Differentiation and scaling is intrinsically connected: it is impossible to differentiate discrete data without increasing the inner scale, i.e. blurring a little. This is a natural consequence of a very important property of this family of differential operators: *regularization* of the differentiation process.

Differentiation is *ill-posed*. What does this mean? For any physical process one wants the output to vary only slightly, when an operator applied to the data is varied slightly: take for example a small high frequency disturbance on our luminance:

$$\tilde{L}(x) = L(x) + \varepsilon \cos \omega x$$

[3] Also denoted as the *contraction* of the source (the image) and the operator, in Florack 1996 [Flo96].

The Gaussian derivative is given by

$$\tilde{L}_x(x,s) = L_x(x,s) + \varepsilon\omega e^{-\omega^2 t/2}\cos\omega x$$

We see that disturbances can be made arbitrarily small provided that the derivative of the signal is computed at sufficiently coarse scale s in scale-space.

Of course the notion of increasing the inner scale with any differentiation is counterintuitive. It is important to realize that the operator is regularized, and not the data. The data should never be modified. This rule however is often seen to be violated: many approaches exist to regularize the data: cubic splines, thin plate approximations, deformable templates, graduated nonconvexity, blurring etc. In essence the outcome may be the same or similar, the philosophy differs. A good review of scale-space operators in the context of regularization is given by Nielsen [NFD96]. Florack [FtHRKV94a] showed that the problem of ill-posedness was solved by Schwartz [Sch51, Sch66], who showed that the derivative of a *distributions* is acquired by convolving the distribution with the derivative of a (any) smooth *test function*. In our case the discrete image is the distribution, the test function is the smooth aperture function, i.e. the (infinitely differentiable) Gaussian kernel. So images can be considered 'regular tempered distributions' in the mathematical language of distribution theory. A clear overview of the regularization properties of scale-space kernels is given by Nielsen et al. [NFD96].

The notion of self-similarity of the kernels becomes especially clear when we express them in so-called *natural coordinates*, i.e. dimensionless spatial units. We then make all measurements relative to the scale at hand: we can consider the scale of the operator as the natural yardstick to measure spatial properties at that scale. Expressed in the natural coordinate $\tilde{x}$, where

$$\tilde{x} \mapsto \frac{x}{\sigma}$$

the Gaussian derivatives get a factor in front:

$$\frac{\partial^n G}{\partial \tilde{x}^n} \mapsto \sigma^n \frac{\partial G}{\partial x^n}$$

The limiting case, when the scale of the operator shrinks to zero, leads to the familiar mathematical differential operator. The zero-th order operator, the Gaussian kernel itself, shrinks to the well-known delta Dirac function (also called the Kronecker delta function). So the Gaussian kernel itself can be seen as the *identity operator*, it does nothing but changing the inner scale.
This intrinsically regularizing framework now enables the extraction of in principle any order of derivative for discrete data. As it turns out, for higher order derivatives we need larger scales. There we encounter a limit: there is a fundamental relation between the order of differentiation, the scale of the operator and the accuracy of the result. This will be discussed in section 4.4. The operators can equally apply to the *temporal domain* with a *temporal scale*. See for a discussion on temporal scale [Koe88, LF96]. Another important concern is: what scale to take? The idea of scale selection will be discussed in section 5.1.

First we take a look at the application of differential geometry and tensor analysis to discrete image data.

4 Geometric Structure and Invariance

The set of partial derivatives to order N in a particular point is called the local *N-jet*, and the set of scaled (i.e. Gaussian) derivatives the multiscale local N-jet. The only relevant entities to consider are properties that are

not changed when we change our coordinate system, i.e. which are *invariant* under a particular coordinate transformation. These local image properties, expressed as some (mostly polynomial) combination of the N-jet elements, are called the local invariants. An example is the inner product of the gradient vector with itself, so we get the *length* of the gradient. Length, in Euclidean geometry, may be defined without reference to coordinate axes, so the dot product of two vectors is a geometric invariant. As an intuitive notion realise the fact that the invariant property is 'attached' to the object, not to the coordinate frame. Hermann Weyl (1885-1955), famous for his contributions to invariant theory [Wey46], stated it clearly: "Every invariant has a geometric meaning".

Coordinate transformations are expressed as elements of *groups*: some important groups of coordinate transformations in computer vision are:

- the group of orthogonal coordinate transformations: rotations and translations of an orthogonal coordinate frame;
- the affine coordinate transformations: linear transformations of the coordinate axes: rotations, translations, shear and expansion;
- the similarity group: the orthogonal group expanded with scaling;
- the perspective transformations: projective coordinate transformations;

Coordinate transformations are characterized by the matrix of the transformations. When the determinant of this matrix is unity, we 'create no extra volume', we call the group *special*. In the following we will focus primarily on medical imaging, where the special group of orthogonal transformations is of most interest.

4.1 Tensors

Tensors play a central role in coordinate independent physics. As Einstein's philosophers' stone, the absolute differential calculus, tensor analysis gives us an easy way to construct the invariants to our need.
Tensors are actually lists of lists of lists of ... etc. A vector is a one-tensor, a single list of numbers. A two-tensor is represented by a matrix, and its physical meaning should be seen as an operator, transforming a vector in a vector. Tensors are equipped with indices, one for each list-index. An excellent introduction to tensor analysis which serves as concise introduction to scale-space theory is Simmonds 1995 [Sim95].
We consider spatial partial derivatives as tensors. E.g. the one-tensor L_i denotes the gradient vector of the (e.g. 3D) luminance function $L(x,y,z)$, consisting of the first order derivatives with respect to x, y and z:

$$L_i \equiv \left\{ \frac{\partial L}{\partial x}, \frac{\partial L}{\partial y}, \frac{\partial L}{\partial z} \right\}$$

In the same fashion the matrix of second order partial derivatives (the *Hessian matrix*) is defined as the two-tensor L_{ij}. The first order differential *nabla operator* is denoted as $\nabla_i \equiv \partial_i$.
There are two constant tensors:

- the Kronecker delta tensor δ_{ij}, defined as the unit or symmetry operator. This tensor has the value 1 when the indices have the same value, and zero otherwise.
- the Levy-Civita epsilon tensor $\varepsilon_{ij\ldots}$, defined as the anti-symmetry operator. The elements of this tensor take the value of the sign of the permutations of the indices $(ij\ldots)$.

Examples in 2D:

$$\delta_{ij} \equiv \left\{ \begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right\} \qquad \varepsilon_{ij} \equiv \left\{ \begin{array}{cc} 0 & 1 \\ -1 & 0 \end{array} \right\}$$

There is much to say about tensors, and, as it turns out, the theory simplifies substantially when we consider Cartesian frames only: consider for example what happens to the image gradient L_i when we apply a linear transformation to the displacement vector $\tilde{x}_i = a_{ij} x_j$ (and thus $x_i = a_{ij}^{-1} \tilde{x}_j$). The chain rule says that $\tilde{\nabla}_i L = a_{ji}^{-1} L_j$. This shows that in general, x_i and L_i transform differently: x_i is said to be a contravariant vector, and L_i a covariant vector. Contravariant vectors are given upper indices, so x_i becomes x^i. The good news is that for Cartesian frames the notion of covariant (lower indices) and contravariant (upper indices) tensors (see for a clear explanation [Sim95] or [MTW73]) disappears, because of the fact that the frame is orthogonal: $a_{ji}^{-1} = a_{ij}$. We will only use lower indices.

Invariant properties are acquired when we *contract* tensors, i.e. form inner products in such a way that scalar values crank out. This is done by *summing over paired indices.* The summation symbol is often omitted and the contraction is known as the *Einstein convention.* E.g. in 3D indices run over *x*, *y* and *z*:

$$L_i L_i \equiv \sum_{i=x,y,z} L_i L_i = L_x^2 + L_y^2 + L_z^2 \quad \text{squared gradient}$$

$$L_{ii} \equiv \sum_{i=x,y,z} L_{ii} = L_{xx} + L_{yy} + L_{zz} \quad \text{Laplacean}$$

where L_x denotes $\frac{\partial L}{\partial x}$. Of course, contraction can also incur the constant tensors. Loosely spoken, contraction with the Kronecker delta tensor gives an inner product, with the Levy-Civita tensor an outer product.
In this way an endless number of invariant geometric properties can be constructed, and measured in each point of the image at a particular scale or range of scales. For clear reasons, such contracted invariants are called *manifest invariants*. Many papers have appeared applying the successful application of (often complex) differential geometry to images [FtHRKV93, tHRFSV94, TG95, WM93, MAS92].
Invariant theory had a major development about a century ago, and is now again fully actual in many important computer vision areas (see the excellent review book [MZ92]). It was shown by Hilbert [Hil90] that *any* invariant of finite order can be expressed as a polynomial expression of a set of *irreducible invariants*. This set is typically very small, e.g. for 2D to second order the list of irreducible invariants numbers only five, i.e.

$$L,\ L_i L_i,\ L_{ii},\ L_i L_{ij} L_j \text{ and } L_{ij} L_{ji}.$$

The irreduclible invariants form in each pixel a *basis* of structural descriptors, and as such are promising in texture description and pattern classification [SM96].

4.2 Gauge coordinates

So far we considered so-called *extrinsic geometry*, i.e. we used an external coordinate frame. Formulas markedly clean up when we are able to define *locally* a coordinate frame, such that one or more of the partial derivatives is zero (in physics this is known as the gauge condition). So we adapt in each point the local frame to the object. The easiest example in 2D is to line one axis of the coordinate frame with the isophote (line of constant intensity) in each point which give a new axis: *v*, and the other, orthogonal axis then aligns with the gradient, giving a new axis: *w*. An infinitesimal deviation along the *v* axis gives no variation in the luminance, as we move along the isophote: $L_v \equiv 0$. Also by this definition, L_w gives the direction where the luminances changes the most: the gradient direction.

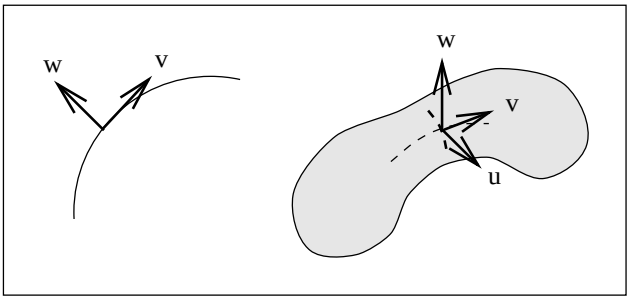


Figure 3: Local gauge coordinates in 2D and 3D (*w* points in the normal, i.e. gradient direction). In 3D, *u* and *v* denote the principal cu

Beause of the *"fixing of the gauge"* we have frozen the extra degree of freedom that we had for our coordinate system: a rotation. Therefor *any* differential expression with partial derivatives with respect to the gauge coordinates *v* and/or *w* is an orthogonal invariant.
Of course we can go from one representation to the other:

$$\frac{\partial}{\partial v} = \frac{L_i \varepsilon_{ij} \partial_j}{\sqrt{L_k L_k}} \ , \ \frac{\partial}{\partial w} = \frac{L_i \delta_{ij} \partial_j}{\sqrt{L_k L_k}}$$

Note the symmetry and the use of the δ and ε tensors. As an example how convenient the gauge coordinates are, let us derive the expression for isophote curvature by studying the local isophote $L(v, w(v) = c$ (where c is a constant) in the point P by implicit differentiation and application of the chain rule, knowing that $L_v \equiv 0$:

$$L_v + L_w w' = 0 \ \ \rightarrow \ \ w'(P) = -\frac{L_v}{L_w} \ \ w'(P) = 0$$

$$L_{vv} + 2L_{vw}w' + L_{ww}w'^2 + L_w w'' = 0$$

we find for the isophote curvature $\kappa \equiv w''$:

$$\kappa(P) = -\frac{L_{vv}}{L_w} \tag{3}$$

a convenient and concise equation. The expression $L_{vv}L_w^2$ turns out to be a good *corner detector*, and is invariant under affine transformations.
In 3D one can study the curvature of the surface in all compass directions starting from the point in study, and it turns out that the largest and the smallest curvatures are found in perpendicular directions: these are the *principal curvatures*, along the *principal directions*. The product of the two is termed the Gaussian curvature, and is an important invariant feature. The sign of the Gaussian curvature e.g. indicates whether the surface is convex or concave. On local flat surface points the Gaussian curvature is zero, the connections of these points forming the *parabolic lines*.
In figure 4 the Gaussian curvature of the inner surface of the left ventricle of a dog is shown for 8 frames of a time sequence, showing the convex and concave parts of the surface as well as the parabolic lines clearly.
The grid formed by the lines (integral paths) of just the maximal or minimal curvature directions on a surface (the *extremal mesh* has succesfully been used in 3D image matching (registration) applications [TG95].
To make structure analysis independent of e.g. brightness and contrast variations ('putting up your sunglasses'), invariance under general intensity transformations forms an important class. The structural properties of isophotes (curvature and its higher order derivatives) form the natural invariants under this requirement.

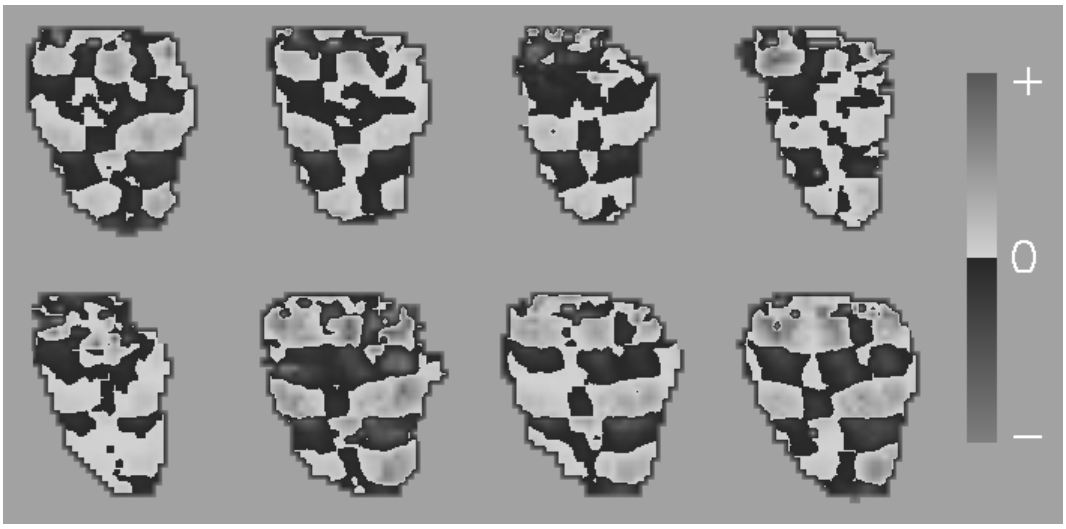


Figure 4: Left: Gaussian curvature of the left ventricle of a canine heart. Eight frames of an MRI sequence, 16 gated frames/second. The scale to the right indicates the sign of the Gaussian curvature: positive for the elliptic areas (convex or concave), negative for the hyperbolic areas (saddlepoint-like). The 'flat' points of zero curvature form the parabolic lines between these areas. From [NDtHRV95].

4.3 Implementation

Gaussian derivatives are calculated by convolution of the image function with the derivative of the Gaussian. This is most easily handled in the Fourier domain, where a convolution reduces to a regular product, and the *n*-th order derivative operator reduces to multiplication with a factor $(i\omega)^n$:

$$F\left(\frac{\partial G}{\partial x}\right) = i\omega_x F(G)$$

$$F\left(L * G_{x_1 \ldots x_n}\right) = F(L).(i\omega_1)\ldots(i\omega_n)F(G)$$

where F denotes the Fourier transformation.

4.4 Accuracy

There exist a fundamental relation between the order of differentiation of an operator, its scale and the required accuracy [tHRNWF94]. For smaller scale σ the Fourier transform of the kernel increases in width, at a certain scale giving rise to *aliasing*. In theory this occurs at all scales due to the infinite extent of the exponential function, but it becomes apparent at smaller scales. The 'leaked' information is folded back, in theory even from all further periods as well, so the amplitude no longer represents the accurate value of 'pure' differentiation.
We consider the powerspectrum, i.e. the square of the signal. The aliasing error can be defined as the relative integrated energy of the aliased frequencies over the total energy:

$$\begin{aligned} error(n,\sigma) &= \frac{\Delta E_n(\sigma,\omega)}{E_n(\sigma,\omega)} \\ &= \frac{\int_{\pi}^{\infty} \omega^{2n} e^{-\sigma^2\omega^2} d\omega}{\int_{0}^{\infty} \omega^{2n} e^{-\sigma^2\omega^2} d\omega} \end{aligned}$$

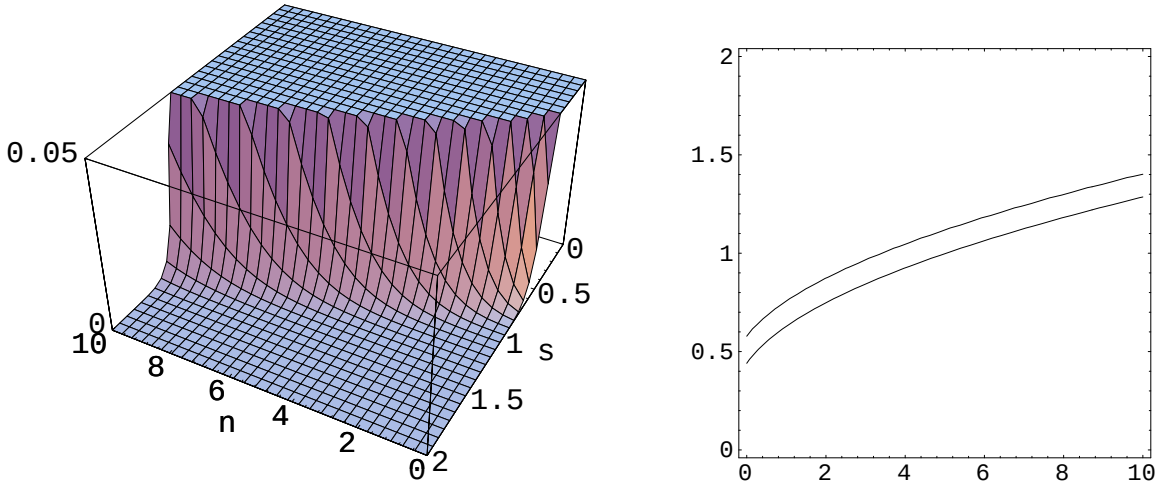


Figure 5: Left: Error of aliasing as a function of differential order (n) and scale (s) of the Gaussian derivative operator, expressed in pixels. Orders (n) up to 10, scales (s) up to 2 pixels. Top level is equal to 5% error. Right: Maximum differential order that can be extracted with Gaussian derivative operators without excessive aliasing as a function of scale of the Gaussian derivative kernels expressed in pixels. Horizontal: orders up to 10, vertical: scales up to 2 pixels. Allowed error: 5% (upper line), 1% (lower line). From [tHRNWF94].

$$= \frac{\Gamma(\frac{1}{2}+n) - \Gamma(\frac{1}{2}+n, 0, 4\pi^2)}{\Gamma(\frac{1}{2}+n)}$$

where $\Gamma(n)$ is the Euler gamma function, and $\Gamma(n, z_0, z_1)$ is the generalized incomplete gamma function $\Gamma(n, z_0) - \Gamma(n, z_1)$. Figure 5 shows the aliasing error as a function of scale σ and order n for orders up to 10. Note that indeed for higher order a larger scale must be selected.
The error allowed is a matter of choice, dependent on the application. The error as a function of σ and n is rather steep, indicating that the maximum allowed order for a given σ (or the reverse) is rather independent of the particular choice of the errorlevel.
This error occurs due to limits in the *inner scale*, i.e. the kernelsize approaches the resolution of the measurement. A similar line of reasoning can be put up for the error due to limits in the *outer scale*, when the kernel gets too wide in the *spatial domain*.

5 Multiscale shape description and segmentation

The differential properties described so far are stricty *local*. Nothing is said about the relations e.g. two neighboring points have. This is the well known problem of *perceptual grouping*. A number of approaches are proposed to attack this fundamental and important problem: Pizer et al. [PBC+94, PEMF96, MPPG96] suggest the use of *symmetry* in shapes. The study of the medial axis (which is a result of processing multilocal information: the boundaries) over scale gives an important shape descriptor, enabling the quantification, matching and deformation studies of shape. The multiscale 'medialness' or ridge function is coined the *core* of an object. It carries the information of the 'width' of the object.

Intuitively it is easy to imagine that image structure has a hierarchical structure over scale. The decreasing information in the scale-space 'stack' gives rize to a graph-like structure of *singularities* (points where

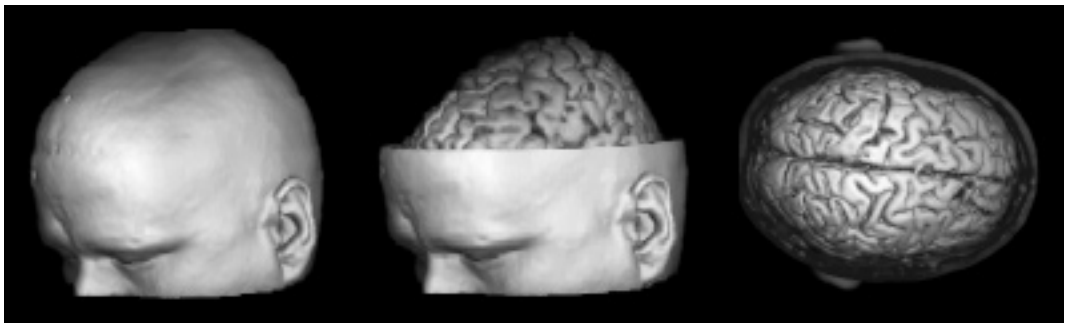


Figure 6: Exploiting the *deep structure* in scale-space to link 'similar' voxels in segments: hyperstack segmentation of MRI brain data with a tumor in the right frontal lobe. From [VKV95].

the gradient vanishes, e.g. intensity extrema, saddle points) and catastrophes (points where higher order derivatives vanish). The structure of this tree, also referred to as the *deep structure*, may be exploited to *link* pixels according to certain criteria to create *segments*. This is exploited in the study of the *hyperstack* [VKV94, VNV96, NVV96]. It enables the *top-down* hierarchical analysis of image structure: the number of required segments determines at what level in scale-space the down projection to the ground level starts. Figure 6 shows a segmentation of the human cortex applying this idea.
Another important 'grouping' paradigm is the use of multiscale image features in *geometrical reasoning* or statistical grouping paradigms [Kar95], like the Hough transform.

5.1 Scale Selection

The scale so far is a *free* parameter: if we don't know what scale to choose, just calculate them all. For a very first stage (like a retina) this may seem a logical strategy, for later stages there must be some *scale selection* for a number of reasons: efficiency of representation, and matching the captured data to the *task* (the leaves or the tree). There are many ways to select a most appropriate scale. Lindeberg pioneered this field by considering the ranking of the volumes objects span in scale-space, the *scale-space primal sketch.* An operator gives maximal output if its size is tuned best to the object. Recently other approaches have been proposed: the study of the variation of the information content (as the logarithm of the entropy of the voxel intensities) over scale [Jäg95], or the condition number of the coefficients when sets of simultaneous equations occur, e.g. in multiscale optic flow [NDV95, NM96].

6 The human visual front-end

There are many neurophysiological and psychophysical data indicating the multiscale analysis by the human visual front-end system. Front-end is here defined as the preattentive stage. The cones and rods in the retina form more or less circular *receptive fields* (RF) of a striking range of sizes. The 'center-surround' sensitivity profile closely resembles a Laplacean of a Gaussian[4]. On the front page a scale-space of the Laplacean is depicted of a sagittal MR of the brain (resolution 256x256 pixels, scale range 256 levels from 1-15 pixels, note the tree structure with increasing scale), i.e. here we see the stack of images sampled by the human retina. An excellent introduction to the early visual system is the Nobelprice-winning work of Hubel and Wiesel [HW62] and the highly recommended book [Hub88].

[4]Speculation: The diffusion equation shows the equivalence between the Laplacean and the derivative to scale, i.e. the variation of the luminance when the aperture size is varied. Is this last property the important measured parameter?

The receptive fields in the retina project to the lateral geniculate nucleus (LGN), a peanut shape kernel in the thalamus. From there projections run to the *visual cortex*. The receptive field profiles in the cortex have a close resemblance to the Gaussian derivatives [You86], which have even been proposed [KvD87, Koe87, KVD88] to give a good *taxonomy* for the many kind of cells found, such as simple cells, complex cells, hyper complex cells, end-stopped cells etc.

The multiscale 'stack' provides an elegant model for the distribution of RF sizes over the retina [LF94]. It is well known that the acuracy decreases with eccentricity from the fovea. When the paradigm of self-similarity is taken into account for the *processing capacity* per scale, all scales should be treated in a similar manner, with the same amount of 'wetware'. This leads to the same number of retinal RF's per scale, tesselated e.g. in a roughly circular area. The smallest RF's together lay out the foveal region, while larger RF sets must occupy a larger area, but again centered around the fovea. This superposition of RF's 'mats' elegantly explains the observed linear decrease with excentricity of many RF-size related phenomena, as acuity, velocity detection thresholds etc.

The Gaussian model only holds for the linear, uncommitted front-end. As soon as the possibility exist to use extracted geometrical information, kernels can be *tuned* to a particular task, leading to *nonlinear* kernels. One example is the framework in which the conductance term becomes a function of differential properties of the image, i.e. of the N-jet J^N:

$$\frac{\partial L}{\partial s} = \vec{\nabla} \cdot c(J^N) \vec{\nabla} L$$

It is interesting to consider the front-end connections in this scale-space framework. In figure 7 a schematic is given of the initial *optical tract*. The first stage, retina to LGN, is completely uncommitted. Here we cannot have something else as a linear stage. The RF's in the LGN show a center-surround structure. Often not or barely mentioned, there is a massive projection of fiber *backwards* from primary cortex to LGN [SK90, She93]: about 50% of all LGN synapses derive from corticogeniculate neurons, while only 10-20% are from the retina, a striking minority! Local inhibitory cells take care of about 25% of the synaptic input. These strong feedback channels, together with the multiscale oriented differential N-jet representation the cortical columns are capable of, provide input to the idea that the front-end visual system may modify the RF's as a function of the image itself. Recent accurate measurements of RF's show that indeed many RF's show strongly varying sensitivity profiles over time[5] [DOF95]. Furthermore, the visual system can be considered a *cascade* of levels (named V1, V2 etc.), from which higher levels information regarding the task, model or memory ('expectancy') may be used to tune the front-end processes.

[5] See the actual measured receptive field data at the interesting Web site of the Visual Neuroscience Lab of Freeman, Ohzawa et al.: **http://totoro.berkeley.edu/teaching/AA_teaching_aids.html**.

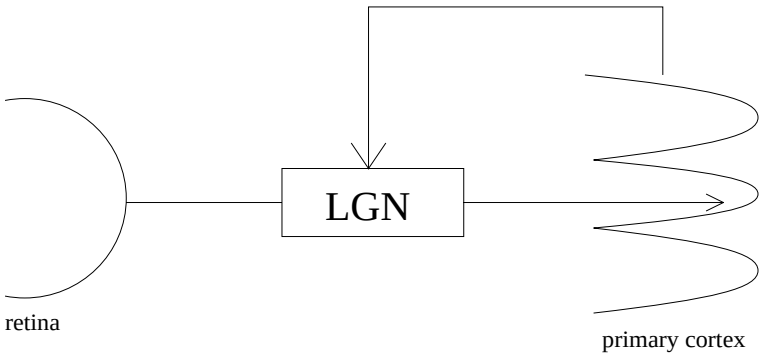


Figure 7: Schematic drawing of the initial optic tract.

7 Nonlinear Scale-Spaces

In the isotropic scale-space all structure is blurred, also the regions of particular interest like edges. Another drawback is the reduced localisation accuracy. Linear scale-space is the paradigm for the *uncommitted* front-end, *because* there is no preference: all information is processed in the same way. Can we relax this constraint, and specify the *task* at hand (e.g. preserving important edges) into the front-end processing, while keeping the properties of a scale-space? It turns out that this is possible, and we enter the broad realm of the nonlinear scale-spaces[6].

There are two essential properties for a multiresolution representation to be a scale-space: the image should be simplified, and it should be possible to establish *relations over scale*. For, only then we able to analyse the deep structure[7]. The retina also measures all scales in a parallel fashion, i.e. *simultaneously*. In this section we summarize the current state-of-the-art with respect to nonlinear multiresolution representations which have the essential scape-space properties. Especially the work by Alvarez, Guichard, Lions and Morel [AGLM93]), Weickert [Wei96a], Van den Boomgaard and Smeulders [vdBS94], and Jackway and Deriche [JD96] should be mentioned, having established many of the results that follow.

7.1 Luminance conserving scale-spaces

In scale-spaces which preserve the luminance relations between points over scale can be established based on their grey value. An important class to consider are the formulations where the change of luminance over scale (the evolution of the image) can be expressed as a *divergence of a flow*, where the flow in general is a function of the local image structure, i.e. the local jet of derivatives:

$$\frac{\partial L}{\partial s} = \nabla \cdot F(L_i, L_{ij}, \ldots)$$

Linear scale-space is an example of such a formulation, the flow is then given by the gradient (see equation (2)). From the divergence theorem (with appropriate boundary conditions, i.e. no flow across the boundaries) we can see that this evolution indeed conserves the luminance:

$$\frac{\partial}{\partial s} \int_V L \; dV = \oint_{\partial V} F \cdot \mathbf{N} = 0$$

[6] This section is based on a more detailed description by Niessen et al. in [NVWV96].

[7] Wavelet representations e.g. lack the scale-space property, for the representation is by *orthogonal* wavelets, which are independent, so no direct relations can be established.

where **N** is the normal to the boundary.

The function F gives us a way to adapt the diffusion to local image structure. It is here that we can bring in knowledge of the task to steer the diffusion process. Some important examples are:

Gradient dependent diffusion: $F = c(\|\nabla L\|)\nabla L$.
This case, where the flow is some scalar function of the gradient, was first proposed by Perona and Malik [PM90]. Their paper gave much impetus to the field. They proposed F to be a decreasing function of the gradient, e.g.

$$c(\vec{x},t) \;=\; c\left(\|\nabla L(\vec{x},t)\|\right) \;=\; e^{-\left(\frac{\|\nabla L\|^2}{k^2}\right)}$$

so that the nonlinear diffusion equation becomes

$$\frac{\partial L}{\partial s} \;=\; \nabla \cdot (c\,\nabla L)$$

As an analogy to heat conduction this can be seen as decribing the evolution of the temperature distribution of a non-uniformly heated surface where the *conductivity* varies: where there is a high gradient, there is a *thermal isolator*, i.e. there is little diffusion. This diffusion, where the conductivity is a function of some image property, is coined *inhomogeneous diffusion*. However, this evolution equation has no smooth solutions, it is ill-posed [NS92, Wei96a, Kic96]. This was solved by Catté et al. [CLMC92] by taking regularized (i.e. Gaussian) derivatives at a certain scale for the determination of the gradient. This has the advantage that insignificant edges, e.g. from high frequency noise, are not kept. Whitaker and Gerig [WG94] proposed extensions of this scheme to vector-valued images, like MRI T_1 and T_2 images.

Tensor dependent diffusion: $F = \mathbf{D}(\mathbf{J})\nabla L$.
Here the diffusion is no longer a function of a scalar entity as above, but a function of a tensorial entity. An example is the *structure tensor* [Wei94], i.e. the Gaussian blurred second moment matrix, e.g. in 2D:

$$\mathbf{J} = G(x,y;\sigma_{int}) * \left\{ \begin{array}{cc} L_x L_x & L_x L_y \\ L_y L_x & L_y L_y \end{array} \right\}$$

where σ_{int} is called the *integration scale*. This tensor is often recognized in texture analysis. It captures the *orientation* of the local structure, when integrated over the volume determined by the integration scale, because orientation is scale dependent. The eigenvectors of this tensor give the principal directions of the structure. A diffusion tensor **D** which is a function of the structure tensor **J** can be designed such that diffusion is done *along* edges and not across. It enhances the coherence in textured images, and it reduces noise at edges, whereas the Perona and Malik equation reduces the diffusion at edges altogether, thus giving limited noise reduction at edges. It was shown by Weickert [Wei96b, Wei96a] that this representation exhibits genuine scale-space properties: image information is reduced, local extrema are not enhanced.

7.2 Geometric flows

This approach to nonlinear scale-spaces considers the evolution of *curves and surfaces* as a function of their geometry. Much of this field originated in the study of propagating fronts in physical flows e.g. in flames [OS88]. 2D curves are best understood and will be discussed here. The curves in images are the isophotes, i.e. curves of points of equal luminance. Curves are fully described by their curvature as a function of curve arclength, so we may state

$$\frac{\partial C}{\partial s} \;=\; g(\kappa)\vec{N}$$

where κ denotes isophote curvature of the curve C, and $\vec{N}$ denotes the outward normal. Note that only the evolution in the normal direction is considered: motion along the tangential direction does not change the curve. We consider the much-studied case $g = \alpha + \beta\kappa$ which leads to a geometrical scale-space interpretation. For $\alpha = 0$, so $g = \kappa$, we get the well known *Euclidean shortening flow*. We can see the geometric interpretation if we rewrite the equation in (v,w)-gauge coordinates (see figure 3). Isophote curvature, as we have seen in eq. (3) is given by $\kappa(P) = -\frac{L_{vv}}{L_w}$, so we get for the evolution:

$$\frac{\partial L}{\partial s} = L_{vv} \tag{4}$$

which can be interpreted as a diffusion equation which only diffuses in the local isophote direction. Alvarez [AGLM92, ALM92] took this property as a starting point to derive the equation.
This evolution process has strong smoothing properties. It was shown by Grayson [Gra87] that non-convex curves evolve into convex curves, and by Gage and Hamilton [GH86] that convex curves shrink to curves with constant curvature, i.e. to round *points*, in finite time. The shrinking is responsible for the name 'shortening flow'.

If $\beta = 0$ the curve only evolves in the normal direction, so inwards or outwards. This closely resembles the erosion and dilation processes in mathematical morphology (see section 7.3).
If α and β are nonzero, we get the so-called *reaction-diffusion* or entropy scale-space [Kim91], with $\beta\kappa$ the diffusive and α the reaction term. Kimia et al. [KTZ94] studied this scale-space for the decomposition of shapes into parts as a multiscale shape description.

Sapiro and Tannenbaum [ST93] presented the theory for *affine invariant scale-space* evolutions. With Olver [OST94] these authors generalized this idea by showing that scale-spaces for curve evolutions invariant under specific groups of transformations, such as the affine and projective group, could be derived by taking derivatives with respect to invariant *arclength*.

Nonlinear diffusion equations and the task: Nonlinear scale-spaces are a flexible framework to tune the evolution process to the task. E.g. to emphasize significant edges the geometric evolutions have been combined with image gradient dependent speed terms [AGLM92, ROF92, STYK94, NtHRFV96]:

$$\frac{\partial L}{\partial s} = h(L_w) * L_{vv}$$

where h is a decreasing function of the image gradient L_w, so there is less smoothing in neighborhoods of large contrast. In figure 8 an example is given of this enhanced shortening flow to denoise CT slices with edge preservation for an interactive anatomical atlas on CD-ROM [WBHH$^+$94].

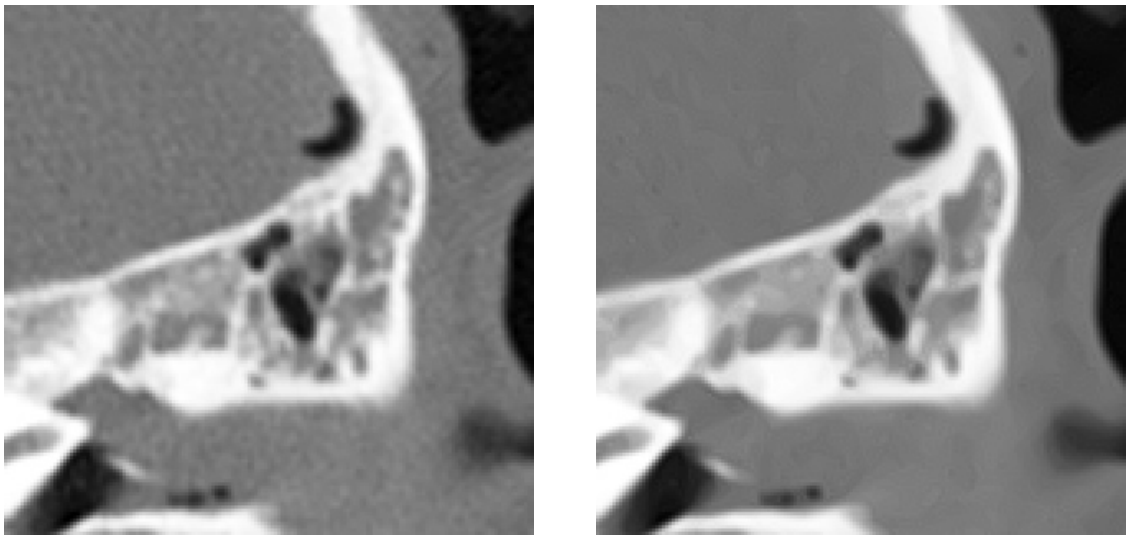


Figure 8: Modified Euclidean shortening flow applied to a CT slice of the inner ear region. Note the edge preserving smoothing. Application: postprocessing of images for an interactive CD-ROM anatomical atlas [WBHH+94].

Osher and Sethian [OS88] published a now classical scheme whereby the level curves are represented by *characteristic functions*, with a finite value of different sign in- and outside the curve, and zero at the levelset. This can be used in image segmentation [CCCD93, MSV95, MSV96], where the zero level set then is supposed to stop at image boundaries. These levels sets can split and merge, making this implementation flexible and popular.

7.3 Morphological scale-spaces

In mathematical morphology a scale-space is represented by the stack of images eroded or dilated with a *structuring element* of increasing size. The flexibility in changing the shape of this element is related to the flexibility in the PDE approach. It has been shown that dilations and erosions with a convex structuring element satisfy the differential semi-group property, and as a particular case, that normal motion flow is formally equivalent to erosion with a circular element. See for extensive discussion of morphological scale-spaces [JD96, BM92, SKS+93, vdBS94].

7.4 Variational approach: functional minimization scale-spaces

Mumford and Shah [MS85] initialized a new field by introducing *energy minimizing functionals*: the image evolution is driven by a functional specified as

$$E(f,K) = \beta \int_{\Omega} (f-g)^2 dx + \int_{\Omega - K} |\nabla f|^2 \, dx + \alpha \, |K|$$

where g the observed grey value image, f the approximation of g, Ω the image domain, K the (closed) set of edges, and $|K|$ is the total length of the edgeset K. α and β are positive parameters. The penalty term $(f-g)$ takes care that the image does not deviate too much from the original. Many modifications have been proposed on this scheme. Rudin [ROF92] took the nonlinear *total variation* as the functional to remove noise in forensic image analysis.

8 Conclusions

Scale-space theory gives us a *complete* family of regularized differential operators, emphasising the intimite relation between differentiation and scale. When the process is fully uncommitted, the isotropic (linear) diffusion equation describes the process of *generating* the scale-space. When there is committment, i.e. evolution process is some function of image- or task-structure, we get nonlinear scale-spaces. Many formalisms have appeared in literature, often seeming entirely different, but it has sometimes been possible to show the equivalence of the approaches.

The field is in rapid development, and this tutorial could only briefly touch the main ideas. In many fields related to medical imaging scale-space theory is now applied [tHR96a], like object segregation and shape representation [PBC$^+$94], heart motion by multiscale optic flow [FN94, NDF$^+$95, NM96], texture analysis [LG93], tumor detection [Kar95], enhancing MRI data [GKKJ92], 2D/3D image matching [vdEMV92] and many others.

A section with recommended references for further reading is included below.

9 Acknowledgements

The author gratefully acknowledges all members of the Utrecht University 3D Computer Vision group (”Tools of Geometry in Vision” team: **http://www.cv.ruu.nl**) on which work most of this tutorial material is based (see also [tHR96b], and the members of the EC-US ”Diffusion” collaboration **http://www.cv.ruu.nl/Collaborations/**.

From July 2-4, 1997 the ***First International Conference on Scale-Space Theory in Computer Vision*** will be held in Utrecht, the Netherlands, as a follow-up to the successful workshops of the Diffusion consortium 1993-1996.

See the WWW page **http://www.cv.ruu.nl/Conferences/ScaleSpace97.html**.

A Recommended reading

Linear scale-space theory:
• Jan Koenderink’s classical paper in Biological Cybernetics: ”The Structure of Images”, [Koe84]
• Tony Lindeberg’s book: ”Scale-Space Theory in Computer Vision”, [Lin94]
• Luc Florack’s monograph: ”The Structure of Scalar Images”, [Flo96]
• Romeny/Florack’s tutorial bookchapter on the relation between scale-space and front-end vision: ”A Multiscale Geometric Model of Human Vision”, [tHRF90]

Nonlinear scale-space theory:
• The EC-US collaboration ”Diffusion” overview book: ”Geometry-Driven Diffusion in Computer Vision”, [tHR94]

• Pietro Perona and Jitendra Malik's classical PAMI paper on nonlinear diffusion: "Scale-Space and Edge Detection Using Anisotropic Diffusion", [PM90]
• The classical paper on morphological nonlinear scale-spaces of Luis Alvarez and the CEREMADE group: "Axiomes et équations fondamentales du traitement d'images", [AGLM92]
• Osher and Sethian's paper on front propagation: "Fronts propagating with curvature dependent speed", [OS88]

Human visual perception:
• David Hubel's work in a Scientific American series: "Eye, Brain and Vision", [Hub88]
• Semir Zeki's book: "A Vision of the Brain", [Zek93]

Tensors and differential geometry:
• Simmonds clear text on tensors: "A Brief on Tensor Analysis", [Sim95]

References

[AGLM92] L. Alvarez, F. Guichard, P. L. Lions, and J. M. Morel. Axiomes et équations fondamentales du traitement d'images. *C. R. Acad. Sci. Paris*, 315:135–138, 1992. 17, 17, 20

[AGLM93] L. Alvarez, F. Guichard, P. L. Lions, and J. M. Morel. Axioms and fundamental equations of image processing. *Arch. for Rational Mechanics*, 123(3):199–257, September 1993. 15

[ALM92] L. Alvarez, P.-L. Lions, and J.-M. Morel. Image selective smoothing and edge detection by nonlinear diffusion. II. *SIAM J. Num. Anal.*, 29(3):845–866, June 1992. 17

[BCHSL92] J. V. Beck, K. D. Cole, A. Haji-Sheikh, and B. Litkouhi. *Heat Conduction using Green's Functions*. Hemispere Publishing Corporation, London, 1992. 5

[BI94] P. Buser and M. Imbert. *Vision*. The MIT Press, London, England, 1994. 1

[BM92] R. W. Brockett and P. Maragos. Evolution equations for continuous-scale morphology. In *Proc. IEEE Int. Conf. on Acoust., Speech, Signal Processing*, pages 125–128, 1992. 18

[BWBD86] J. Babaud, A. P. Witkin, M. Baudin, and R. O. Duda. Uniqueness of the Gaussian kernel for scale-space filtering. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 8(1):26–33, 1986. 6, 6

[CCCD93] V. Caselles, F. Catte, T. Coll, and F. Dibos. A geometric model for active contours in image processing. *Numer. Math*, 66:1–31, 1993. 18

[CLMC92] F. Catté, P.-L. Lions, J.-M. Morel, and T. Coll. Image selective smoothing and edge detection by nonlinear diffusion. *SIAM J. Num. Anal.*, 29(1):182–193, February 1992. 16

[DOF95] G. C. DeAngelis, I. Ohzawa, and R. D. Freeman. Receptive field dynamics in the central visual pathways. *Trends Neurosci.*, 18:451–458, 1995. 14

[Flo96] L. M. J. Florack. *The Structure of Scalar Images*. Computational Imaging and Vision Series. Kluwer Academic Publishers, Dordrecht, 1996. In preparation. 6, 19

[FN94] L. M. J. Florack and M. Nielsen. The intrinsic structure of the optic flow field. Technical Report ERCIM-07/94-R033 or INRIA-RR-2350, ERCIM, July 1994. http://www-ercim.inria.fr/publication/technical_reports. 19

[Fou55] J. Fourier. *The Analytical Theory of Heat*. Dover Publications, Inc., New York, 1955. Replication of the English translation that first appeared in 1878 with previous corrigenda incorporated into the text, by Alexander Freeman, M.A. Original work: "Théorie Analytique de la Chaleur", Paris, 1822. 3

[FtHRKV92] L. M. J. Florack, B. M. ter Haar Romeny, J. J. Koenderink, and M. A. Viergever. Families of tuned scale-space kernels. In G. Sandini, editor, *Proceedings of the European Conference on Computer Vision*, pages 19–23, Santa Margherita Ligure, Italy, May 19–22 1992. 5

[FtHRKV93] L. M. J. Florack, B. M. ter Haar Romeny, J. J. Koenderink, and M. A. Viergever. Cartesian differential invariants in scale-space. *Journal of Mathematical Imaging and Vision*, 3(4):327–348, November 1993. 9

[FtHRKV94a] L. M. J. Florack, B. M. ter Haar Romeny, J. J. Koenderink, and M. A. Viergever. Images: Regular tempered distributions. In Ying-Lie O, A. Toet, H. J. A. M. Heijmans, D. H. Foster, and P. Meer, editors, *Proc. of the NATO Advanced Research Workshop Shape in Picture - Mathematical description of shape in greylevel images*, volume 126 of *NATO ASI Series F*, pages 651–660. Springer Verlag, Berlin, 1994. 7

[FtHRKV94b] L. M. J. Florack, B. M. ter Haar Romeny, J. J. Koenderink, and M. A. Viergever. Linear scale-space. *Journal of Mathematical Imaging and Vision*, 4(4):325–351, 1994. 2, 2

[GH86] M. Gage and R. S. Hamilton. The heat equation shrinking convex plane curves. *J. Differential Geometry*, 23:69–96, 1986. 17

[GKKJ92] G. Gerig, O. Kübler, R. Kikinis, and F. A. Jolesz. Nonlinear anisotropic filtering of MRI data. *IEEE Transactions on Medical Imaging*, 11(2):221–232, June 1992. 19

[Gra87] M. Grayson. The heat equation shrinks embedded plane curves to round points. *Journal of Differential geometry*, 26:285–314, 1987. 17

[HG84] R. A. Hummel and B. C. Gidas. Zero crossings and the heat equation. Technical Report 111, New York Univ. Courant Institute of Math. Sciences, Computer Science Division, 1984. 6

[Hil90] D. Hilbert. *Theory of Algebraic Invariants*. Cambridge Mathematica Library. Cambridge University Press, 1890. Reprinted lecture notes 1897. 9

[Hub88] D. H. Hubel. *Eye, Brain and Vision*, volume 22 of *Scientific American Library*. Scientific American Press, New York, 1988. 1, 13, 20

[HW62] D. H. Hubel and T. N. Wiesel. Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex. *Journal of Physiology*, 160:106–154, 1962. 1, 13

[HW79] D. H. Hubel and T. N. Wiesel. Brain mechanisms of vision. *Scientific American*, 241:45–53, 1979. 1

[Jäg95] Martin Jägersand. Saliency maps and attention selection in scale and spatial coordinates: An information theoretic approach. In *Proc. Fifth Intern. Conf. on Computer Vision*, pages 195–202, MIT Cambridge, MA, June 20-23 1995. IEEE. Catalog number 95CB35744. 13

[JD96] P. T. Jackway and M. Deriche. Scale-space properties of multiscale morphological dilation-erosion. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 18(1):38–51, 1996. 15, 18

[Kar95] Nico Karssemeijer. Detection of stellate distortions in mammograms using scale-space operators. In *Proc. Information Processing in Medical Imaging 1995*, pages 335–346, 1995. 13, 19

[Kic96] S. Kichenassamy. Nonlinear diffusions and hyperbolic smoothing for edge enhancement. In *Proceedings of 12th International Conference on Analysis and Optimization of Systems*, volume 219 of *Lecture Notes in Control and Information Sciences*, pages 119–124, 1996. 16

[Kim91] B. B. Kimia. Entropy scale-space. In *Proc. of Visual Form Workshop*, Capri, Italy, May 1991. Plenum Press. 17

[Koe84] J. J. Koenderink. The structure of images. *Biol. Cybern.*, 50:363–370, 1984. 2, 2, 5, 6, 6, 19

[Koe87] J. J. Koenderink. Design principles for a front-end visual system. In Rolf Eckmiller and Christoph v. d. Malsburg, editors, *Neural Computers*. Springer-Verlag, 1987. Proceedings of the NATO Advanced Research Workshop on Neural Computers, held in Neuss, Fed. Rep. of Germany, September 28–October 2. 14

[Koe88] J. J. Koenderink. Scale-time. *Biol. Cybern.*, 58:159–162, 1988. 7

[KTZ94] B. B. Kimia, A. Tannenbaum, and S. W. Zucker. Shapes, shocks, and deformations, I. *International Journal of Computer Vision*, 15(3):189–224, 1994. 17

[KvD87] J. J. Koenderink and A. J. van Doorn. Representation of local geometry in the visual system. *Biol. Cybern.*, 55:367–375, 1987. 14

[KVD88] J. J. Koenderink and A. J. Van Doorn. Operational significance of receptive field assemblies. *Biol. Cybern.*, 58:163–171, 1988. 14

[KvD94] J. J. Koenderink and A. J. van Doorn. Two-plus-one dimensional differential geometry. *Pattern Recognition Letters*, 21(15):439–443, May 1994. 5

[LF94] T. Lindeberg and L. Florack. Foveal scale-space and linear increase of receptive field size as a function of eccentricity. Technical Report ISRN KTH/NA/P–94/27–SE, Dept. of Numerical Analysis and Computing Science, Royal Institute of Technology, August 1994. Submitted. 14

[LF96] T. Lindeberg and D. Fagerstrom. Scale-space with causal time direction. In *Proc. 4th European Conference on Computer Vision*, volume 1064 of *LNCS*, pages 229–240, Cambridge, UK, April 1996. Springer-Verlag. 7

[LG93] T. Lindeberg and J. Gårding. Shape from texture from a multi-scale perspective. In H. H. Nagel et al., editors, *Proceedings of the fourth ICCV*, pages 683–691, Berlin, Germany, 1993. IEEE Computer Society Press. 19

[Lin90] T. Lindeberg. Scale-space for discrete signals. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 12(3):234–245, 1990. 6

[Lin94] T. Lindeberg. *Scale-Space Theory in Computer Vision*. The Kluwer International Series in Engineering and Computer Science. Kluwer Academic Publishers, Dordrecht, the Netherlands, 1994. 2, 19

[LtHR94] T. Lindeberg and B. M. ter Haar Romeny. Linear scale-space: I. basic theory. II. early visual operations. In B. M. ter Haar Romeny, editor, *Geometry-Driven Diffusion in Computer Vision*, Computational Imaging and Vision, pages 1–38,39–72. Kluwer Academic Publishers, Dordrecht, the Netherlands, 1994. 2

[MAS92] O. Monga, N. Ayache, and P. T. Sander. From voxel to intrinsic surface features. *Image and Vision Computing*, 10(6):403–417, July/August 1992. 9

[Mor85] P. Morrison. *Powers of Ten: About the Relative Size of Things in the Universe*. W. H. Freeman and Company, 1985. 1

[MPPG96] B. S. Morse, S. M. Pizer, D. T. Puff, and C Gu. Zoom-invariant vision of figural shape: Effects on cores of image disturbances. Technical Report TR96-005, University of North Carolina, Dept. of Computer Science, 1996. To appear in revised form in Computer Vision and Image Understanding. 12

[MS85] D. Mumford and J. Shah. Boundary detection by minimizing functionals. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, San Francisco, 1985. 18

[MSV95] R. Malladi, J. Sethian, and B. Vemuri. Shape modeling with front propagation: a level set approach. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 17(2):158–174, 1995. 18

[MSV96] R. Malladi, J. A. Sethian, and B. C. Vemuri. A fast level set based algorithm for topology-independent shape modelling. *Journal of Mathematical Imaging and Vision*, 6(2/3):269–289, June 1996. 18

[MTW73] C. W. Misner, K. S. Thorne, and J. A. Wheeler. *Gravitation*. Freeman, San Francisco, 1973. 9

[MZ92] J. L. Mundy and A. Zisserman, editors. *Geometric Invariance in Computer Vision*. MIT Press, Cambridge, Massachusetts, 1992. 9

[NDF+95] W. J. Niessen, J. S. Duncan, L. M. J. Florack, B. M. ter Haar Romeny, and M. A. Viergever. Spatiotemporal operators and optic flow. In S. T. Huang and D. N. Metaxas, editors, *Physics-Based Modeling in Computer Vision*, pages 78–84. IEEE Computer Society Press, 1995. 19

[NDtHRV95] W. J. Niessen, J. S. Duncan, B. M. ter Haar Romeny, and M. A. Viergever. Spatiotemporal analysis of left ventricular motion. In M. H. Loew, editor, *Medical Imaging 95: Image Processing*, pages 250–261. SPIE Press, Bellingham, 1995. 11

[NDV95] W. J. Niessen, J. S. Duncan, and M. A. Viergever. A scale-space approach to motion analysis. In J. C. Van Vliet, editor, *Computing Science in the Netherlands 95*, pages 170–181. Stichting Mathematisch Centrum, Amsterdam, 1995. 13

[NFD96] M. Nielsen, L. M. J. Florack, and R. Deriche. Regularization, scale-space, and edge detection filters. In *Proc. Fourth European Conference on Computer Vision*, Cambridge, UK, April 14-18 1996. 7, 7

[NM96] W. J. Niessen and R. Maas. Multiscale optic flow and stereo. In J. Sporring, M. Nielsen, L. Florack, and P. Johansen, editors, *Gaussian Scale-Space Theory*, Computational Imaging and Vision. Dordrecht: Kluwer Academic Publishers, 1996. In press. 13, 19

[NS92] M. Nitzberg and T. Shiota. Nonlinear image filtering with edge and corner enhancement. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 14(8):826–833, 1992. 16

[NtHRFV96] W. J. Niessen, B. M. ter Haar Romeny, L. M. J. Florack, and M. A. Viergever. A general framework for geometry-driven evolution equations. *International Journal of Computer Vision*, 1996. In press. 17

[NVV96] W. J. Niessen, K. L. Vincken, and M. A. Viergever. Comparison of multiscale representations for a linking-based image segmentation model. In *Proceedings IEEE Workshop on Mathematical Methods in Biomedical Image Analysis*, pages 263–272, San Francisco, 1996. 13

[NVWV96] W. J. Niessen, K. L. Vincken, J. Weickert, and M. A. Viergever. Nonlinear multiscale representations for image segmentation. *Computer Vision and Image Understanding*, 1996. Submitted. 15

[OS88] S. Osher and S. Sethian. Fronts propagating with curvature dependent speed: algorithms based on the Hamilton-Jacobi formalism. *J. Computational Physics*, 79:12–49, 1988. 16, 18, 20

[OST94] P. Olver, G. Sapiro, and A. Tannenbaum. Differential invariant signatures and flows in computer vision: A symmetry group approach. In Bart M. ter Haar Romeny, editor, *Geometry-Driven Diffusion in Computer Vision*, Computational Imaging and Vision, pages 255–306. Kluwer Academic Publishers, Dordrecht, 1994. 17

[Pap94] T. V. Papathomas, editor. *Early Vision and Beyond*. The MIT Press, London, England, 1994. 1

[PBC+94] Stephen M. Pizer, Christina A. Burbeck, James M. Coggins, Daniel S. Fritsch, and Bryan S. Morse. Object shape before boundary shape: Scale-space medial axes. *Journal of Mathematical Imaging and Vision*, 4(3):303–313, 1994. 12, 19

[PEMF96] S. M. Pizer, D. Eberly, B. S. Morse, and D.S. Fritsch. Zoom-invariant vision of figural shape: The mathematics of cores. Technical Report TR96-004, University of North Carolina, Dept. of Computer Science, 1996. To appear in revised form in Computer Vision and Image Understanding. 12

[PFMvG93] E. J. Pauwels, P. Fiddelaers, Th. Moons, and L. J. van Gool. An extended class of scale-invariant and recursive scale-space filters. Technical Report KUL/ESAT/MI2/9316, Catholic University Leuven, 1993. 5

[PM90] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 12(7):629–639, July 1990. 16, 20

[ROF92] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992. 18

[Sch51] L. Schwartz. *Théorie des Distributions*, volume I, II of *Actualités scientifiques et industrielles; 1091,1122*. Publications de l'Institut de Mathématique de l'Université de Strasbourg, Paris, 1950–1951. 7

[Sch66] L. Schwartz. *Théorie des Distributions*. Hermann & Cie, 1966. 7

[She93] S. M. Sherman. Dynamic gating of retinal transmission to the visual cortex by the lateral geniculate nucleus. In D. Minciacchi, M. Molinari, G. Macchi, and E. G. Jones, editors, *Thalamic Networks for Relay and Modulation*, pages 61–79. Pergamon Press, Oxford, 1993. 14

[Sim95] J. G. Simmonds. *A Brief on Tensor Analysis*. Undergraduate Texts in Mathematics. Springer-Verlag, 1995. Second Edition. 8, 9, 20

[SK90] S. M. Sherman and C. Kock. Thalamus. In G. M. Shepherd, editor, *The Synaptic Organization of the Brain*, pages 246–278. Oxford University Press, New York, 1990. Third Edition. 14

[SKS+93] G. Sapiro, R. Kimmel, D. Shaked, B. B. Kimia, and A. M. Bruckstein. Implementing continuous-scale morphology via curve evolution. *Pattern Recognition*, 26(9):1363–1372, 1993. 18

[SM96] C. Schmidt and R. Mohr. Object recognition using local characterization and semi-local constraints. Technical report, INRIA, 1996. 9

[SMKL93] R. Shapley and D. Man-Kit Lam, editors. *Contrast Sensitivity*. The MIT PRess, London, England, 1993. 1

[ST93] G. Sapiro and A. Tannenbaum. Affine invariant scale-space. *International Journal of Computer Vision*, 11:25–44, 1993. 17

[STYK94] G. Sapiro, A. Tannenbaum, Y. You, and M. Kaveh. Experiments on geometric enhancement. In *International Conference on Image Processing*, pages 472–475. IEEE, 1994.

[TG95] J.-P. Thirion and A. Gourdon. Computing the differential characteristics of isodensity surfaces. *Computer Vision, Graphics, and Image Processing: Image Understanding*, 61(2):109–202, March 1995. 9, 10

[tHR94] B. M. ter Haar Romeny, editor. *Geometry-Driven Diffusion in Computer Vision*. Kluwer Academic Publishers, Dordrecht, 1994. 19

[tHR96a] B. M. ter Haar Romeny. Applications of scale-space theory. In J. Sporring, M. Nielsen, L. Florack, and P. Johansen, editors, *Gaussian Scale-Space Theory*, Computational Imaging and Vision. Dordrecht: Kluwer Academic Publishers, 1996. In press. 19

[tHR96b] B. M. ter Haar Romeny. Scale-space research at Utrecht University. In M.-O. Berger, R. Deriche, I. Herlin, J. Jaffré, and J.-M. Morel, editors, *Proc. 12th International Conference on Analysis and Optimization of Systems: Images, Wavelets and PDE's*, volume 219, pages 15–30, CEREMADE / INRIA, Paris, France, June 26-28 1996. Springer, London. Lecture Notes in Control and Information Sciences. 19

[tHR97] B. M. ter Haar Romeny. *Front-End Vision and Multiscale Image Analysis: Introduction to Scale-Space Theory*. Kluwer Academic Publishers, Dordrecht, the Netherlands, 1997. In preparation. 1

[tHRF90] B. M. ter Haar Romeny and L. M. J. Florack. A multiscale geometric model of human vision. Technical Report 90-14, Computer Vision Research Group (3DCV), Utrecht Biophysics Research Institute (UBI), 1990. 19

[tHRF93] B. M. ter Haar Romeny and L. M. J. Florack. A multiscale geometric model of human vision. In William R. Hendee and Peter N. T. Wells, editors, *Perception of Visual Information*, chapter 4, pages 73–114. Springer-Verlag, Berlin, 1993. Second edition 1996. 2

[tHRFSV94] B. M. ter Haar Romeny, L. M. J. Florack, A. H. Salden, and M. A. Viergever. Higher order differential structure of images. *Image and Vision Computing*, 12(6):317–325, July/August 1994. 9

[tHRNWF94] B. M. ter Haar Romeny, W. J. Niessen, J. Wilting, and L. M. J. Florack. Differential structure of images: Accuracy of representation. In *Proc. First IEEE Internat. Conf. on Image Processing*, pages 21–25, Austin, TX, November, 13-16 1994. IEEE. 11, 12

[vdBS94] R. van den Boomgaard and A. W. M. Smeulders. The morphological structure of images, the differential equations of morphological scale-space. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(11):1101–1113, November 1994. 15, 18

[vdEMV92] P. A. van den Elsen, J. B. A. Maintz, and M. A. Viergever. Geometry driven multimodality matching of brain images. *Brain Topography*, 5:153–158, 1992. 19

[VKV94] K. L. Vincken, A. S. E. Koster, and M. A. Viergever. Probabilistic segmentation of partial volume voxels. *Pattern Recognition Letters*, 15(5):477–484, 1994. 13

[VKV95] K. L. Vincken, A. S. E. Koster, and M. A. Viergever. Probabilistic hyperstack segmentation of MR brain data. In N. Ayache, editor, *Computer Vision, Virtual Reality and Robotics in Medicine, Proc. CVRMed'95*, volume 905 of *Lecture Notes in Computer Science*, pages 351–357. Springer-Verlag, Berlin, 1995. 13

[VNV96] K. L. Vincken, W. J. Niessen, and M. A. Viergever. Blurring strategies for image segmentation using a multiscale linking model. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR'96*, pages 21–26, San Francisco, CA, 1996. IEEE Computer Society Press. 13

[WBHH+94] W. A. van Wolferen, T. S. Boschma, T. W. A. Hofstee-Hooftman, R. A. H. Matthijssen, J. J. Doorn, E. B. H. M. Polman, K. Zuiderveld, F. W. Zonneveld, J. A. den Boer, and G. J. R. Maat. *Temporal bone & Posterior cranial fossa*, volume 1-II. Elsevier, Amsterdam, the Netherlands, 1994. Hillen, B., Editor, CD-ROM/CD-I. 17, 18

[Wei94] J. Weickert. Scale-space properties of nonlinear diffusion filtering with a diffusion tensor. Technical Report 110, Laboratory of Technomathematics, Univ. of Kaiserslautern, Germany, 1994. Submitted. 16

[Wei96a] J. Weickert. *Anisotropic Diffusion in Image Processing*. PhD thesis, Dept. of Mathematics, University of Kaiserslautern, Germany, January 1996. 15, 16, 16

[Wei96b] J. Weickert. Theoretical foundations of anisotropic diffusion in image processing. In W. Kropatsch, R. Klette, and F. Solina, editors, *Theoretical Foundations of Computer Vision*, volume 11 of *Computing Supplement*, pages 221–236. Springer, Wien, 1996. 16

[Wey46] H. Weyl. *The Classical Groups, their Invariants and Representations*. Princeton University Press, Princeton, NJ, 1946. 8

[WG94] R. Whitaker and G. Gerig. Vector-valued diffusion. In B. M. ter Haar Romeny, editor, *Geometry-Driven Diffusion in Computer Vision*, Computational Imaging and Vision, pages 93–134. Kluwer Academic Publishers B.V., 1994. 16

[Wit83] A. P. Witkin. Scale-space filtering. In *Proc. International Joint Conference on Artificial Intelligence*, pages 1019–1023, Karlsruhe, Germany, 1983. 5

[WM93] J. Weber and J. Malik. Robust computation of optical flow in a multi-scale differential framework. In *Proc. Fourth International Conference on Computer Vision*, pages 12–20, 1993. 9

[You86] R. A. Young. The Gaussian derivative model for machine vision: Visual cortex simulation. Publication GMR-5323, General Motors Research Labs, Computer Science Dept., 30500 Mound Road, Box 9055, Warren, Michigan 48090-9055, July 7 1986. 14

[Zek93] S. Zeki. *A Vision of the Brain*. Blackwell Scientific Publications, Oxford, 1993. 1, 20