Linköping Studies in Science and Technology
Thesis No. 1594

A Variational Approach to Image Diffusion in Non-Linear Domains

Freddie Åström

Linköping University
INSTITUTE OF TECHNOLOGY

Department of Electrical Engineering
Linköpings universitet, SE-581 83 Linköping, Sweden

Linköping June 2013

A Variational Approach to Image Diffusion in Non-Linear Domains

Cover image of vegetables from the database "McGill Calibrated Colour Image Database", `http://tabby.vision.mcgill.ca/html/welcome.html` (May 2013). The image was corrupted with noise and filtered using the method in Paper B.

Computer Vision Laboratory
Department of Electrical Engineering
Linköping University
SE-581 83 Linköping
Sweden

ISBN 978-91-7519-606-0 ISSN 0280-7971 LIU-TEK-LIC-2013:28

Abstract

Image filtering methods are designed to enhance noisy images captured in situations that are problematic for the camera sensor. Such noisy images originate from unfavourable illumination conditions, camera motion, or the desire to use only a low dose of ionising radiation in medical imaging. Therefore, in this thesis work I have investigated the theory of partial differential equations (PDE) to design filtering methods that attempt to remove noise from images. This is achieved by modeling and deriving energy functionals which in turn are minimized to attain a state of minimum energy. This state is obtained by solving the so called Euler-Lagrange equation. An important theoretical contribution of this work is that conditions are put forward determining when a PDE has a corresponding energy functional. This is in particular described in the case of the structure tensor, a commonly used tensor in computer vision.

A primary component of this thesis work is to model adaptive image filtering such that any modification of the image is structure preserving, but yet is noise suppressing. In color image filtering this is a particular challenge since artifacts may be introduced at color discontinuities. For this purpose a non-Euclidian color opponent transformation has been analysed and used to separate the standard RGB color space into uncorrelated components.

A common approach to achieve adaptive image filtering is to select an *edge stopping* function from a set of functions that have proven to work well in the past. The purpose of the edge stopping function is to inhibit smoothing of image features that are desired to be retained, such as lines, edges or other application dependent characteristics. Thus, a step from ad-hoc filtering based on experience towards an application-driven filtering is taken, such that only desired image features are processed. This improves what is characterised as *visually relevant features*, a topic which this thesis covers, in particular for medical imaging.

The notion of what are relevant features is a subjective measure may be different from a layman's opinion compared to a professional's. Therefore, we advocate that any image filtering method should yield an improvement not only in numerical measures but also a visual improvement should be experienced by the respective end-user.

Acknowledgments

This work is the product of discussions and interaction of many people which I have encountered throughout my studies. In particular I would like to take the opportunity to thank

- **Michael Felsberg**, supervisor, for excellent guidance, support, and for being a true source of inspiration.
- **George Baravdish**, co-supervisor, a supportive and encouraging mathematician who can explain the most difficult concepts in tangible words.
- **Reiner Lenz**, co-supervisor, for helpful insights on color image processing.

Furthermore I would like to thank all other members of the computer vision laboratory for providing a pleasant working environment. In particular I mention **Per-Erik Forssén**, **Rudolf Mester**, and **Vasileios Zografos** for involvement in my studies, also **Johan Wiklund** for keeping my computer happy.

I want to further thank my **family** for encouragement and unconditional support through this work.

Especially, I give my gratitude to **Hafsa Åström** who always encouraged and supported me through difficult times.

This research has received funding from the Swedish Research Council through grants for the projects *Non-linear adaptive color image processing*, *Visualization-adaptive Iterative Denoising of Images*, and from the ECs 7th Framework Programme (FP7/2007-2013), grant agreement 247947 (GARNICS). Also, this work has been conducted within (in collaboration with) the center for Medical Image Science and Visualization (CMIV) at Linköping University, Sweden. CMIV is acknowledged for provision of financial support and access to leading edge research infrastructure.

Freddie Åström June 2013

Contents

Part I

Background Theory

Chapter 1

Introduction

1.1 Motivation

The technology for modern consumer cameras has advanced to the point of producing very high quality images under favourable illumination and atmospheric conditions. However, poor light conditions or camera motion during acquisition still poses major problems in terms of image quality. The problem is relevant since most cellphone cameras use cheap sensors designed for portability. Thus, this thesis investigates the problem of image enhancement, with a particular emphasis on denoising of color images.

In the left column of Figure 1.1 two such noisy images are illustrated. The noise is modeled as an additive component and it is clear from the images that the noise is unwanted artifact since it reduces the impression of having a high quality image. If this noise would not be present, as in the enhanced version in the right column, colors and image structure appear more visually appealing.

The approach taken to image enhancement is based on theory from partial differential equations (PDE). The approach is motivated since PDEs are often used to model physical phenomena such as heat convection and wave propagation. Particularly the framework of heat convection can be applied in image processing since it describes a mass transportation system. The physical interpretation of utilizing the process of heat convection in image processing is that image structures can then be used to guide the transportation of mass. This results in an adaptive filtering process which suppresses noise and preserves image features. Thus the topic of this thesis is about defining PDE-based filtering methods that suppress noise but preserve structures such as lines and edges contained in an image.

Noisy Enhanced

Figure 1.1: Example of images from [15] with additive noise, and the corresponding filtered version on the right obtained using the method in Paper B.

1.2 Outline

This section gives an overview of the two main parts of this thesis. The first part consists of background theory and the second part includes three publications.

1.2.1 Outline Part I: Background Theory

In this part the primary tools and theoretical models to achieve adaptive filtering are formulated. In particular, the diffusion framework and its numerical approximation are considered. Also, theory on the model used for color image processing is presented.

1.2.2 Outline Part II: Included Publications

Preprint versions of three publications are included in Part II. The full details and abstract of these papers, together with statements of the contributions made by the author, are summarized below.

Paper A: Color Persistent Anisotropic Diffusion of Images

> F. Åström, M. Felsberg, and R. Lenz. Color Persistent Anisotropic Diffusion of Images. In *SCIA'11*, volume 6688 of *LNCS*, pages 262–272, Springer, Heidelberg, 2011.

Abstract:
Techniques from the theory of partial differential equations are often used to design filter methods that are locally adapted to the image structure. These techniques are usually used in the investigation of gray-value images. The extension to color images is non-trivial, where the choice of an appropriate color space is crucial. The RGB color space is often used although it is known that the space of human color perception is best described in terms of non-euclidean geometry, which is fundamentally different from the structure of the RGB space. Instead of the standard RGB space, we use a simple color transformation based on the theory of finite groups. It is shown that this transformation reduces the color artifacts originating from the diffusion processes on RGB images. The developed algorithm is evaluated on a set of real-world images, and it is shown that our approach exhibits fewer color artifacts compared to state-of-the-art techniques. Also, our approach preserves details in the image for a larger number of iterations.

Contribution:
The main novelty in this paper was to suggest a color model which reduce artifacts introduced when filtering image regions that contain sharp discontinuities of color intensities. The author contributed to the findings, performed the experiments and main part of the writing.

Paper B: On Tensor-Based PDEs and their Corresponding Variational Formulations with Application to Color Image Denoising

> F. Åström, G. Baravdish, and M. Felsberg. On Tensor-Based PDEs and their Corresponding Variational Formulations with Application to Color Image Denoising. In *ECCV 2012: 12th European Conference on Computer Vision, 7-12 October, Firenze, Italy*, 2012.

Abstract:
The case when a partial differential equation (PDE) can be considered as an Euler-Lagrange (E-L) equation of an energy functional, consisting of a data term and a smoothness term is investigated. We show the necessary conditions for a PDE to be the E-L equation for a corresponding functional. This energy functional is applied to a color image denoising problem and it is shown that the method compares favorably to current state-of-the-art color image denoising techniques.

Contribution:
This paper introduces a novel functional derived from the PDE equations of standard non-linear diffusion schemes. The case when a tensor-based PDE can be expressed as an energy functional is investigated. The author contributed to the derivation of the main theorem, corollary and the proposition. Also the author performed the main part of writing and performed all experiments.

Paper C: Targeted Iterative Filtering

> F. Åström, M. Felsberg, G. Baravdish, and C. Lundström. Targeted Iterative Filtering. In *Fourth International Conference on Scale Space and Variational Methods in Computer Vision (SSVM 2013), 2-6 June 2013, Schloss Seggau, Austria*, volume 7893 of *LNCS*, pages 1–11, Springer, Heidelberg, 2013.

Abstract:
The assessment of image denoising results depends on the respective application area, *i.e.* image compression, still-image acquisition, and medical images require entirely different behavior of the applied denoising method. In this paper we propose a novel non-linear diffusion scheme that is derived from a linear diffusion process in a value space determined by the application. We show that application-driven linear diffusion in the transformed space compares favorably with existing nonlinear diffusion techniques.

Contribution:
This paper is the first paper which introduces a diffusion framework resulting in a non-linear filtering method which is application-driven rather than data-driven. The author contributed to the findings of the necessary condition for existence of solution, the transformation of statistical moments using a non-linear mapping function, the main part of writing and performed all experiments.

Other Publications

The following publications by the author are related to the included papers.

F. Åström and R. Köker. A parallel neural network approach to prediction of Parkinson's Disease. *Expert systems with applications*, 38(10):12470–12474, 2011. (Methodology based on my MSc work)

F. Åström, M. Felsberg, and R. Lenz. Color Persistent Anisotropic Diffusion of Images. In *SSBA Symposium 2011, 17-18 March, Linköping*, 2011. (Early version of Paper A)

F. Åström, M. Felsberg, G. Baravdish, and C. Lundström. Visualization Enhancing Diffusion. In *SSBA Symposium on Image Analysis*, Stockholm, Sweden, 2012. (Early version of Paper C)

Chapter 2

Variational calculus based image filtering

The framework of partial differential equations (PDE) is a theoretically well founded approach in computer vision. PDEs are frequently used in applications such as optical flow, deblurring, image enhancement, depth estimation from stereo, and image restoration. The common denominator for these approaches is the formulation of energy functionals and the differences lie in the modeling of the problem.

In this work we have considered the problem of image restoration, where the artifact to be removed is described as an additive component to the image signal. Particularly, we are interested in the process described as image diffusion. In physics the term diffusion is used to describe the transportation of mass or heat across e.g. cell membranes. The diffusion process is described by Fick's law

$$J = -D\nabla u \tag{2.1}$$

where ∇u refers to the concentration gradient, J is the mass flow and D is a diffusivity constant describing the rate of flow [22]. This process continues as long as the concentration gradient is non-zero, which is as long as the substance on each side of the membrane wall is not equally distributed. With the motivation of Fick's law researchers try to model image restoration using this physical process; the analogy being that by enforcing a mass transportation within an image will yield a regularization of the image values to attain a state of equilibrium. By this process artifacts in the image would be suppressed. Artifacts describe "unwanted" structure, however what structures are considered as artifact is application dependent and thus it will be more accurately specified in following sections.

Figure 2.1 is an overview of the different cases of image diffusion and their possible corresponding variational formulation. As will be established in the first section of this chapter, the PDE of linear diffusion is derived from a corresponding functional. The case of non-linear diffusion can be derived for certain edge stopping functions, however, due to spatial averaging this connection cannot be made in the general case, thus this is indicated by the dashed arrow. Variational formulation

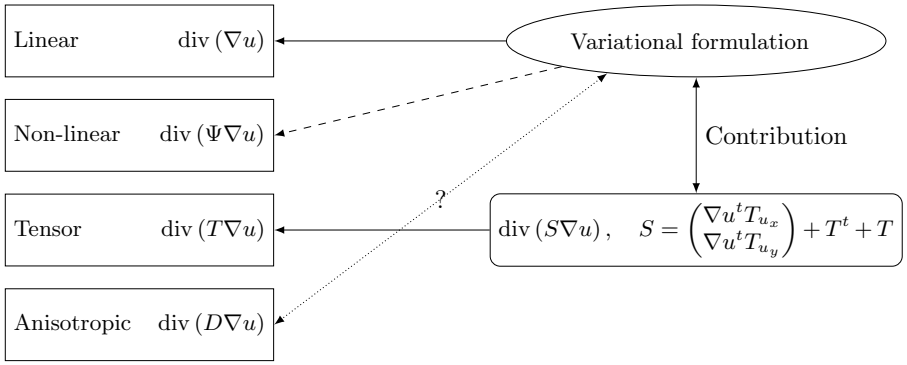


Figure 2.1: Overview of existing variational formulations in image diffusion, see text for details.

for general anisotropic diffusion has not yet been made. However, if the tensor D is expressed using a first order approximation, then it has been shown that a corresponding variational formulation exists [11].

The primary motivation to establish a connection between a variational formulation and a tensor-based PDE scheme is that image diffusion can then be understood as an energy minimization problem possibly enabling further advances in the diffusion community. Thus in Paper B, we derive the corresponding Euler-Lagrange equation for a tensor-based energy functional and investigate for which conditions we can find a functional given a tensor-based PDE. The connection is indicated by the thick line with double arrowhead in Figure 2.1.

2.1 Linear diffusion

The variational approach to image diffusion is to model an energy functional $E(u)$ consisting of a data term and a regularization term $R(u)$, i.e.

$$E(u) = \frac{1}{2}\int_\Omega (u-u^0)^2\,dx + \lambda R(u)\ , \tag{2.2}$$

where $x \in \Omega$ and u^0 denotes the observed image. The constant λ is a positive scalar which determines the amount of the regularization. The domain Ω is a grid described by the image size in pixels, and $\nabla = \partial_x = \left(\partial_{x_1}, \cdots, \partial_{x_d}\right)^t$ is the gradient operator, and $\dim(\nabla) = d$ is the number of dimensions. If $R(u)$ is selected as

$$R(u) = \frac{1}{2}\int_\Omega |\nabla u|^2\,dx\ , \tag{2.3}$$

one obtains linear diffusion which is the model describing the mass transportation process (2.1), the most fundamental diffusion method.

In order to find the solution u^* which minimizes (2.2), one searches for stationary points by computing the Euler-Lagrange (E-L) equation

$$\frac{\partial E(u)}{\partial u} = 0 \quad \text{in} \quad \Omega, \quad \nabla u \cdot n = 0 \quad \text{on} \quad \partial\Omega \ , \tag{2.4}$$

where n is the normal vector on the boundary $\partial\Omega$, and $\cdot$ denotes the scalar product. The interpretation of the boundary condition is that there is no heat transportation across the boundary resulting in a preservation of the average intensity level of the image. This boundary condition is a Neumann boundary condition.

To compute the E-L equation, the Gâteaux derivative in variational calculus is used in connection to Green's first identity. The variational derivative of the regularization term of $E(u)$ is then defined as

$$\partial R \cdot v = \lim_{\varepsilon \to 0} \frac{R(u + \varepsilon v) - R(u)}{\varepsilon} \ , \tag{2.5}$$

where v is a test function not equal to zero. Now using (2.3)

$$\begin{aligned} \partial R \cdot v &= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_\Omega |\nabla(u + \varepsilon v)|^2 - |\nabla(u)|^2 \, dx \\ &= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_\Omega (|\nabla u|^2 + 2\varepsilon \nabla u^t \nabla v + \varepsilon^2 |\nabla v|^2) - |\nabla u|^2 \, dx \\ &= \int_\Omega 2 \nabla u^t \nabla v \, dx \ , \end{aligned}$$

with Green's first identity [20]

$$\int_\Omega (2\nabla u^t) \nabla v \, dx = \int_{\partial\Omega} v (2\nabla u \cdot n) \, dS - \int_\Omega v \, \mathrm{div}\,(2\nabla u) \, dx \ ,$$

where $S = \text{bdy } \Omega$, (bdy = boundary), we obtain

$$\partial R \cdot v = - \int_\Omega \mathrm{div}\,(2\nabla u)\, v \, dx \ ,$$

where we have used the Neumann condition $\nabla u \cdot n|_{\partial\Omega} = 0$, and $\mathrm{div}\,(\cdot)$ is the divergence operator. From (2.4) and since $v \neq 0$ we obtain the condition that

$$-2 \, \mathrm{div}\,(\nabla u) = 0 \iff -2\Delta u = 0 \ ,$$

where Δ is the Laplace operator. Then the u^* which minimizes $E(u)$ is obtained by solving the Euler-Lagrange equation

$$\begin{cases} u - u^0 - \lambda \, \mathrm{div}\,(\nabla u) = 0 & \text{in} \quad \Omega \\ \nabla u \cdot n = 0 & \text{on} \quad \partial\Omega \ . \end{cases} \tag{2.6}$$

Since Δ is a rotationally symmetric operator, it does not adapt to any content such as lines and edges in the image u^0, which results in smoothing of image structures. The solution of the heat equation given at a time t has been shown to relate to the framework of scale-space representation [14], where the finest scale is given at $t = 0$ and the coarsest scale is obtained at $t \to \infty$. However, to avoid trivial solutions (2.6) is implemented using an iterative scheme and is stopped before $t \to \infty$. More on numerical aspects of implementing diffusion methods will be described in chapter 4.

2.2 Non-linear diffusion

With the aim to adapt the diffusion filtering to structures in the image, Perona and Malik [17] proposed a non-linear scalar diffusion process. The novelty in this scheme is to reduce the filtering at image structures with large gradients such as lines and edges. A functional which models this behavior can be described using the regularization term

$$R(u) = \int_\Omega \Phi(|\nabla u|)\, dx \ , \tag{2.7}$$

where $\Phi : \mathbb{R} \to \mathbb{R}$. In order to minimize $R(u)$, the variational derivative is computed using

$$\partial R \cdot v = \left. \frac{\partial R(u + \varepsilon v)}{\partial \varepsilon} \right|_{\varepsilon = 0} \ ,$$

which is an equivalent form to (2.5). Concentrating on the regularization term we obtain

$$\begin{aligned} \partial R \cdot v &= \int_\Omega \frac{\partial}{\partial \varepsilon} \Phi(|\nabla(u + \varepsilon v)|)|_{\varepsilon=0}\ dx \\ &= \int_\Omega \frac{\Phi'(|\nabla(u + \varepsilon v)|)}{|\nabla(u + \varepsilon v)|} (\nabla u \cdot \nabla v + \varepsilon |\nabla v|^2)|_{\varepsilon=0}\, dx \ , \end{aligned}$$

which results in the variation in v as

$$\partial R \cdot v = \int_\Omega \frac{\Phi'(|\nabla u|)}{|\nabla u|} \nabla u \cdot \nabla v \, dx \ ,$$

using Greens identity, and Neumann boundary conditions, the E-L equation is obtained as

$$\left\{ \begin{array}{rcll} u - u^0 - \lambda \operatorname{div}(\Psi(|\nabla u|)\nabla u) & = 0 & \text{in} & \Omega \\ \nabla u \cdot n & = 0 & \text{on} & \partial\Omega \ , \end{array} \right. \tag{2.8}$$

where

$$\Psi(|\nabla u|) = \frac{\Phi'(|\nabla u|)}{|\nabla u|} \ . \tag{2.9}$$

If Ψ is chosen as a strictly decreasing function such that if $|\nabla u| \to 0$ then $\Psi \to 1$ and $|\nabla u| \to \infty$ then $\Psi \to 0$ one obtains an inhibition of the filtering process for large image gradients and in homogeneous regions the filtering will be isotropic. The corresponding function Φ in $E(u)$ can be obtained by integration of (2.9), i.e.

$$\Phi(|\nabla u|) = \int_\Omega \Psi(|\nabla u|)|\nabla u| \, d|\nabla u| \ . \tag{2.10}$$

Two popular selections of diffusivity functions Ψ, with the corresponding Φ-function computed according to (2.10), are

$$\Psi(|\nabla u|) = \frac{1}{1 + (|\nabla u|/k)^2} \implies \Phi(|\nabla u|) = \frac{k^2}{2} \log\left(1 + \left(\frac{|\nabla u|}{k}\right)^2\right) + C \ , \tag{2.11}$$

and

$$\Psi(|\nabla u|) = \exp(-(|\nabla u|/k)^2) \implies \Phi(|\nabla u|) = -\frac{k^2}{2}\exp\left(-(|\nabla u|/k)^2\right) + C \ , \quad (2.12)$$

where k is an edge-stopping parameter fixed to suppress the flux at edges and lines in the image. The constant C should be chosen such that Φ is non-negative. The parameter k can be selected according to methods such as histogram analysis, decision theory [24], robust statistics [18] or noise estimation methods [9]. Black and Rangarajan [8] show that for certain Ψ-functions there exists families of corresponding Φ-functions.

Here we make the remark that Ψ is an ad-hoc selection of the diffusivity function. In Paper C we consider the case where the diffusivity function is determined based on prior knowledge of important image structures. In this case the filtering is transferred from a gradient guided filter to a value guided filter. For this application-driven approach we establish necessary and sufficient conditions for the existence of the solution, details can be found in the supplementary material of Paper C.

2.3 Tensor-based diffusion

Despite an increased level of details is retained in the non-linear Perona and Malik [17] diffusion, it also preserves noise at lines and edges. A diffusion process that steers the filtering process parallel to the image structure, still preserving edges, was introduced by Weickert [22]. This leads to a tensor-based non-linear diffusion scheme which uses the structure tensor [7, 10]

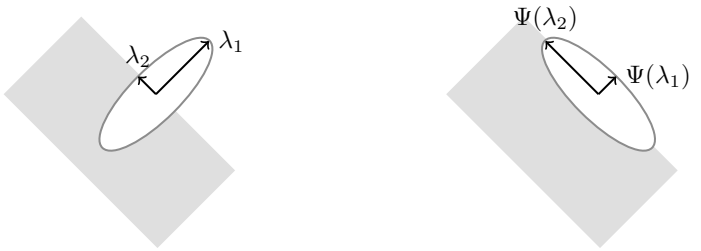
$$T(u(x,y)) = \begin{pmatrix} \int w(x-\xi)(\partial_x u(\xi))^2 \, d\xi & \int w(x-\xi)\partial_x u(\xi)\partial_y u(\xi) \, d\xi \\ \int w(x-\xi)\partial_x u(\xi)\partial_y u(\xi) \, d\xi & \int w(x-\xi)(\partial_y u(\xi))^2 \, d\xi \end{pmatrix} , \quad (2.13)$$

where w is a Gaussian function. The interpretation of the tensor T is that it describes a second moment matrix computed by integrating the image gradient over the domain of u. In practice it is useful to estimate the local variance by applying a windowing function chosen as the Gaussian function w. If w is selected as

$$w(x) = \delta(x) = \begin{cases} 1, & x = 0 \\ 0, & \text{otherwise} \ , \end{cases} \quad (2.14)$$

then T is a tensor of rank 1 with eigenvalues $\lambda_1 = |\nabla u|^2$ and $\lambda_2 = 0$ computed using (2.17) below. The implication of a rank one tensor is that only changes orthogonal to the image structure are described [9]. The structure tensor is illustrated in Figure 2.2 (a) and it can be seen from its eigensystem that the tensor is aligned orthogonal to the image structure, in this case an edge. The resulting effect of using this tensor is that the edge is blurred.

To align the structure tensor with the image structure, its eigenvalues are remapped using a monotonically decreasing function $\Psi(s)$ such that $\Psi(0) = 1$ and



(a) Structure tensor (b) Diffusion tensor

Figure 2.2: Illustration of a tensor in $\mathbb{R}^2$ mapped to be aligned with the image structure.

$\Psi(s) = 0$ as $s \to \infty$, conditions fulfilled by e.g. (2.11) or (2.12). Then the diffusion tensor can be computed as [9]

$$D(T) = E^t \begin{pmatrix} \Psi(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & \Psi(\lambda_d) \end{pmatrix} E \ , \tag{2.15}$$

where E consists of the eigenvectors and λ_i, $i = 1, ..., d$ the eigenvalues of T. The obtained diffusion tensor is shown in Figure 2.2 (b). The resulting update scheme is given by the PDE

$$\begin{cases} u - u^0 - \lambda \operatorname{div}(D(T)\nabla u) & = 0 \quad \text{in} \quad \Omega \\ \nabla u \cdot n & = 0 \quad \text{on} \quad \partial\Omega \ . \end{cases} \tag{2.16}$$

The remapping of the structure tensor described in (2.15) requires the positive semi-definite structure tensor to be decomposed into eigenvalues and eigenvectors. If T is a two dimensional and a symmetric matrix such that

$$T = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \ ,$$

where $a, b, c \in \mathbb{R}$, there always exists v, λ such that v is an eigenvector and λ an eigenvalue in the eigensystem

$$Tv = \lambda v \ .$$

Using the relations

$$\begin{aligned} \det(T) &= ac - b^2 \\ \operatorname{tr}(T) &= a + c \ , \end{aligned}$$

one can show that the eigenvalues are

$$\lambda_{1,2} = \frac{1}{2}\left(\operatorname{tr}(T) \pm \sqrt{\operatorname{tr}(T)^2 - 4\det(T)}\right) \ , \tag{2.17}$$

and λ_1 is chosen such that $\lambda_1 > \lambda_2$. The eigenvector $v = \left(v_1, v_2\right)^t$ associated with λ_1 can then be obtained by solving the equation system

$$\left\{ \begin{array}{ccccc} (a-c-\alpha)v_1 & + & 2bv_2 & = & 0 \\ 2bv_1 & + & (c-a-\alpha)v_2 & = & 0 \end{array} \right. , \tag{2.18}$$

where $\alpha = \sqrt{(a+c)^2 - 4(ac-b^2)}$. This gives the orthonormal eigenvectors v and w such that

$$v = \begin{pmatrix} 2b \\ c-a+\alpha \end{pmatrix} \qquad \text{and} \qquad w \perp v \ ,$$

where w is the eigenvector corresponding to λ_2 and the normalization to unit length has been left out [22].

Chapter 3

Image color representation

Up to this point of the thesis, the presented framework has not been limited to any particular image representation model such as spectral images, color images, video sequences or grayscale images. Considering these different image modalities, particular image representations can be described as special cases of the more general theoretical results.

Color images are usually represented as components of the RGB (red, green and blue) color space. An example of a color image is depicted in Figure 3.3 (a). Figure 3.1 illustrates the color spectrum of visible light experienced by humans. The color range is from violet (left) to red (right) and the wavelengths change from short to long. The RGB color space is sampled from this spectrum of visible light. Commonly, image data is represented using an 8bit representation giving a total of 256 quantization levels. Since a color image commonly contains 3 components one obtains a total of 3×255 values to represent the equiluminant colors from pure violet to pure red. If we are interested in one color component and we vary its value from 0 to 255, one obtains a transition from dark to light (i.e. from the color black (0) to white (1)). Obviously, higher quantization levels can represent a larger number of spectral components. Due to the structure of the light spectrum, color components of the RGB color space are highly correlated [19]. The implication of this relation, for image filtering methods, is that by modifying one color component, artifacts may be introduced at boundaries with sharp color gradients. However, the most basic way of viewing an image is to model it as a grayscale image. To construct such a grayscale image from an RGB image, a simple approach is to average the color components. If a color opponent transform

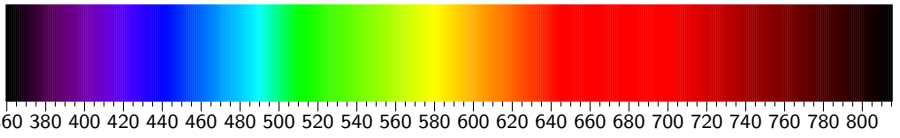


Figure 3.1: Color spectrum. Example collected from the LaTeX xcolor package.

(described in the following section) is used then the first component which is seen in Figure 3.3 (e) corresponds to the gray value representation of (a). For more information on image datasets used in this thesis see chapter 5.2.

3.1 Filtering vector-valued images

When vector-valued images are considered for image processing, primarily two approaches are possible. Either each component is considered separately and independently, or there is some (non-)linear combination of the color components. The work conducted in Paper A investigates techniques to decorrelate the RGB components. In particular PCA (principal components analysis) and a color opponent transform similar to PCA are used [13]. The same color opponent model is used in Paper B in the numerical evaluation of the proposed diffusion approach. Diffusion filtering in the RGB color space is possible, e.g. Weickert [23] defined anisotropic diffusion in the RGB color space by weighting the computed diffusion tensor for each channel equally. However, in Paper A, this approach was shown to be suboptimal in terms of perceived image quality.

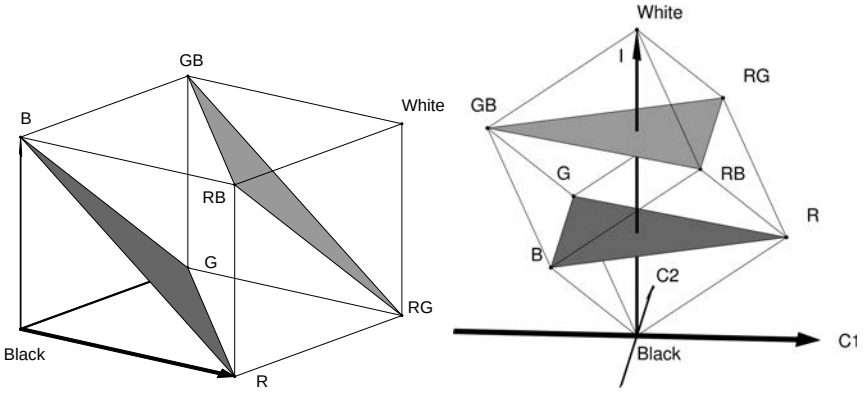
To illustrate the information contained in an RGB color image, Figure 3.3 (a) shows an image which is mainly composed by the primary colors green and red, but also cyan, a mixture of the blue and green color. In the same figure, red, green and blue color components respectively are shown in (b)-(d). The color opponent transform is shown in Figure 3.3 (e)-(g) and is defined as [13]

$$I = \frac{1}{\sqrt{3}}(R + G + B) \ , \tag{3.1}$$

$$C1 = \sqrt{\frac{2}{3}}(R - \frac{1}{2}G - \frac{1}{2}B) \ , \tag{3.2}$$

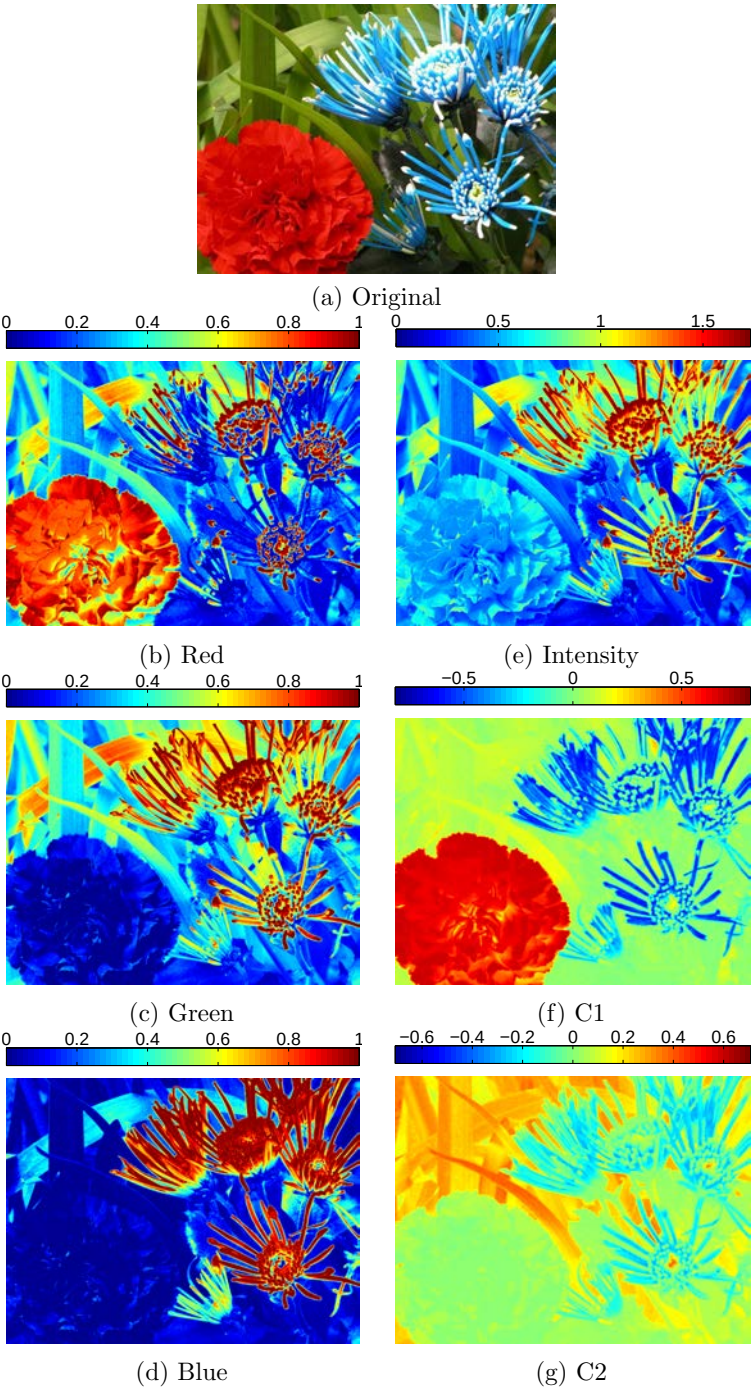
$$C2 = \frac{1}{\sqrt{2}}(G - B) \ , \tag{3.3}$$

where I denotes the intensity and $C1$ and $C2$ are two components orthogonal to the intensity axis. The color transformation is illustrated using a vector diagram in Figure 3.2. In the left part of the figure, the RGB color cube is depicted. The dark-gray hyperplane indicated by the vectors R, G and B represents the case when the vector sum of the components is constant 1, then the transform to the color opponent plane is given on the right shown with the corresponding labels. The light-gray hyperplane composed of the vectors RB, RG and GB shows the case when the sum of the components is constant 2, the corresponding color opponent transformation is also shown in the right figure. The interpretation of the color transformation is that i.e. the color red is primarily described by a large contribution in $C1$, and that green and blue are opponent colors in the sign of $C2$. Figure 3.3 (e)-(g) illustrates the color transform in comparison to the RGB components. The color map was chosen for sake of visualization.



RGB Color opponent transform

Figure 3.2: Illustration of color opponent transform. For the color components $0 \leq R, G, B \leq 1$ the range of the opponent components are $0 \leq I \leq 3/\sqrt{3}$, $-\sqrt{2/3} \leq C1 \leq \sqrt{2/3}$ and $-1/\sqrt{2} \leq C2 \leq 1/\sqrt{2}$. See text for details.



(a) Original

(b) Red

(e) Intensity

(c) Green

(f) C1

(d) Blue

(g) C2

Figure 3.3: (a) Original image, (b)-(d) RGB components of (a) and (e)-(g) color opponent transform of (a). The image (a) was obtained from [16], with label identification *samplemerry_flor0003.jpg*

Chapter 4

Numerical schemes

In this chapter we consider the numerical aspects of implementing the presented diffusion methods, as well as error measures correlated with perceived image quality are discussed.

4.1 Numerical implementation

A diffusion process, such as the ones described in (2.6), (2.8) and (2.16) can be solved numerically by considering an evolution equation of the form

$$\partial_t u = A(u)u \ , \tag{4.1}$$

where A is a spatially dependent linear operator containing i.e. the derivative operators on its diagonals. An iterative discretization scheme can then be written as

$$\frac{u^{i+1} - u^i}{\lambda} = \beta A(u^i)u^{i+1} + (1-\beta)A(u^i)u^i \ , \tag{4.2}$$

where $0 \leq i$ is the number of iterations. For $\beta = 1$ (4.2) denotes a fully implicit scheme, if $\beta = 1/2$ then the scheme is a semi-implicit scheme and for $\beta = 0$ a fully explicit scheme is obtained. If $\beta > 0$ then it is required to solve an equation system of size $N^2 \times N^2$ in each iteration, where N is the number of pixels in the image. Troughout this work the explicit scheme has been used due to its simplicity. The explicit scheme reads

$$u^{i+1} = u^i + \lambda A(u^i)u^i \ , \tag{4.3}$$

and the matrix A, also denoted as stencil, can be expressed using a sparse representation making memory requirements less demanding.

Considering the PDEs (2.6), (2.8) and (2.16) they are of the type

$$\operatorname{div}(D\nabla u) = \operatorname{div}(D)\,\nabla u + \operatorname{tr}(D\Delta u) \ , \tag{4.4}$$

0	+	0
+	+	+
0	+	0

(a)

?	+	?
+	+	+
?	+	?

(b)

Table 4.1: (a) Signs of stencil when D is a scalar function. (b) Signs of stencil when D is a tensor [24].

which yield the equivalent form

$$\operatorname{div}\begin{pmatrix} d_{11}\partial_x u + d_{12}\partial_y u \\ d_{12}\partial_x u + d_{22}\partial_y u \end{pmatrix} = \partial_x(d_{11}\partial_x u) + \partial_x(d_{12}\partial_y u) + \partial_y(d_{12}\partial_x u) + \partial_y(d_{22}\partial_y u) \ . \tag{4.5}$$

From (4.4) it can be seen that both first and second order derivatives need to be approximated. A common method to approximate these derivatives is by using finite differences, introduced in the next section. An important remark is that when approximating (4.5) one can attain negative values in the stencil A due to the mixed derivatives components $\partial_x\partial_y$.

The signs of the stencil, A, obtained for D being a scalar function, as in (2.8) or a tensor as in (2.16) are illustrated in Table 4.1. The possible existence of negative factors in the stencil, indicated with ?, can lead to over- and undershoots. However if the steplength, λ, is small enough then these artifacts are rarely observed. Non-negative schemes have been formulated which are proven to yield positive entries in the stencil [22]. For example, the non-negative scheme by Weickert has been used in Paper B.

Felsberg [9] derived an autocorrelation-driven diffusion scheme resulting in the approximation $\operatorname{div}(D)\nabla u \approx 0$ in (4.4). Thus, reducing the complexity of the implementation considerably. This is also the discretization used in Paper A, and has shown to yield competitive results compared to other discretization techniques.

4.2 Finite difference operators

An important aspect of solving the PDEs introduced in this thesis is how to implement the derivatives of a function. The Taylor series in one dimension reads

$$u(x) = \sum_{k=0}^{\infty} \frac{u(a)^{(k)}}{k!}(x-a)^k \ ,$$

where it is assumed that $u \in \mathcal{C}^\infty$. However often second order approximations are sufficient ($k \leq 2$). Here, the computation of the first and second order image derivatives are considered. There are three types of finite difference operators;

∂^+ : forward operator

∂^- : backward operator

∂ : central difference operator such that $\partial = (\partial^+ + \partial^-)/2$

Given an image grid $\Omega \subset \mathbb{R}^2$ where $\Omega = [M, N]$, and M is the size of the grid in the first dimension (y-direction), and N is the size of the second dimension (x-direction). Let h be the size of one pixel given in x- and y-direction. A third order two dimensional Taylor series expansion in the x-direction and y-direction, respectively, then reads

$$u(x+h,y) = u(x,y) + hu_x(x,y) + \frac{h^2}{2}u_{xx}(x,y) + O(h^3) \tag{4.6}$$

$$u(x-h,y) = u(x,y) - hu_x(x,y) + \frac{h^2}{2}u_{xx}(x,y) + O(h^3) \tag{4.7}$$

$$u(x,y+h) = u(x,y) + hu_y(x,y) + \frac{h^2}{2}u_{yy}(x,y) + O(h^3) \tag{4.8}$$

$$u(x,y-h) = u(x,y) - hu_y(x,y) + \frac{h^2}{2}u_{yy}(x,y) + O(h^3) \ . \tag{4.9}$$

The Taylor expansion in the diagonal directions reads

$$u(x+h,y+h) = u + hu_x + hu_y + \frac{h^2}{2}(u_{xx} + 2u_{xy} + u_{yy}) + O(h^3) \tag{4.10}$$

$$u(x+h,y-h) = u + hu_x - hu_y + \frac{h^2}{2}(u_{xx} - 2u_{xy} + u_{yy}) + O(h^3) \tag{4.11}$$

$$u(x-h,y+h) = u - hu_x + hu_y + \frac{h^2}{2}(u_{xx} - 2u_{xy} + u_{yy}) + O(h^3) \tag{4.12}$$

$$u(x-h,y-h) = u - hu_x - hu_y + \frac{h^2}{2}(u_{xx} + 2u_{xy} + u_{yy}) + O(h^3) \ , \tag{4.13}$$

where the argument (x, y) has been dropped for increased clarity. From (4.6) and (4.7) it is possible to derive the forward and backward finite difference operators ∂^+ and ∂^- as a second order approximations in the x-direction

$$\partial_x^+ u = \frac{u(x+h,y) - u(x,y)}{h} + O(h^2) \tag{4.14}$$

$$\partial_x^- u = \frac{u(x,y) - u(x-h,y)}{h} + O(h^2) \tag{4.15}$$

$$\partial_x u = \frac{u(x+h,y) - u(x-h,y)}{2h} + O(h^3) \ , \tag{4.16}$$

and by using (4.8) and (4.9) the corresponding operators in the y-direction is obtained.

With the central differences, the second order derivatives are computed by summing (4.6) and (4.7), respectively (4.8) and (4.9)

$$u_{xx}(x,y) = \frac{u(x+h,y) - 2u(x,y) + u(x-h,y)}{h^2} + O(h^3)$$

$$u_{yy}(x,y) = \frac{u(x,y+h) - 2u(x,y) + u(x,y-h)}{h^2} + O(h^3)$$

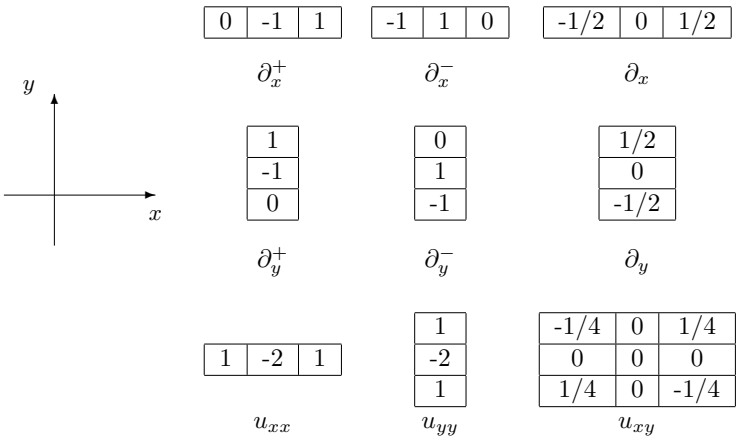


Figure 4.1: Convolution masks, where $h = 1$ and the coordinate system on the left indicates the respective x and y-orientation.

and the mixed-derivative filter is obtained by summing (4.10) and (4.13) and subtracting (4.11) and (4.12) which result in the third-order approximation

$$u_{xy}(x,y) \approx \frac{u(x+h,y+h) - u(x+h,y-h) - u(x-h,y+h) + u(x-h,y-h)}{4h^2} \,. \tag{4.17}$$

In practice it is useful to write the finite difference operators as convolution masks. Figure 4.1 illustrates the convolution masks derived in this section as well as the first order derivative of the y-component.

Chapter 5

Evaluation

5.1 Image quality

Image quality is a subjective measure, and it is application dependent what is considered a good result. In this thesis, image quality has been assessed both qualitatively and quantitatively. Note that a qualitative assessment as "visually appealing" can be biased since it is a subjective measure and what appears as a good result for one person may not be generalized to others. Hence, the purpose of image quality measurements is to not only attempt to model what is visually appealing, but it is also to model an objective benchmark measure which enables a comparison between two different competing methods.

Throughout this work the considered noise component of an image is assumed as a linear additive component which is normally-distributed i.e.

$$u^0 = u_0 + \eta \ , \tag{5.1}$$

where u^0 is the image, u_0 is the noise free image and $\eta \sim \mathcal{N}(0, \sigma^2)$ is the noise component.

Mean Squared Error

A quantitative measure that is commonly used to represent image quality is the mean squared error (MSE). It is computed by summing the squared difference between the noise-free image and the filtered image $\hat{u}$ i.e.

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^{N} ((u_0)_i - \hat{u}_i)^2 \ , \tag{5.2}$$

where N is the number of pixels in the image. Another popular measure, based on the MSE, is the peak-signal-to-noise-ratio (PSNR),

$$\text{PSNR} = 10 \log_{10} \left(\frac{1}{MSE} \right) \ , \tag{5.3}$$

and $0 \leq u \leq 1$. However, as shown in [21] the perceptual quality does not correlate very well with the MSE and the PSNR metrics. With this motivation we often impose larger importance to the structural similarity index (SSIM) as a measure for image quality, explained in the next section.

Structural Similarity Index

The Structural SIMmilarity index (SSIM) is an error measure which measures image similarity. It is designed to be an objective image quality metric more consistent with how the human visual system perceives image quality. The fundamental difference between MSE and SSIM is that MSE is a metric which measures image quality in terms of absolute strength of the average error signal, whereas SSIM is a measure which compares local statistical properties of the signal thus giving an error estimate in terms of structural differences [21]. The SSIM error value is computed as

$$\mathrm{SSIM}(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \ . \tag{5.4}$$

In (5.4) x and y represent the two different images that are compared, one is the noisy image and the other one is the noise free image. The SSIM measure is designed based on the three components: luminance estimated by the mean value μ, the signal contrast which is estimated from the standard deviation σ and the structure similarity is computed according to the correlation coefficient between the images x and y. C_1 and C_2 are two normalisation constants.

In practice, the statistical properties are computed in local 8×8 pixel sized windows. However this may introduce block artifacts [21], hence it is recommended that a small Gaussian filter is used as a weight function. Furthermore, rather than applying the SSIM index measure directly on the image, it is suggested that the image scale is fitted to depend on view distance and image resolution. The authors of SSIM have made the implementation publicly available on their webpage[1]. This is also the implementation of SSIM that was used in this thesis.

5.2 Image datasets

In the image analysis community certain images are popular for algorithm benchmarks. This section gives a brief overview of the image datasets that have been used in the included publications (Paper A, B, C).

Road sign dataset

The dataset was collected in [12] and the primary aim with the dataset is road sign recognition. The image file type is *.jpg* and the image size is 1280×960 pixels. The collected images were acquired from a camera mounted behind the windshield of a car. Regions of interest for image diffusion are edges with different color tone and intensity. In paper A, images *image00756.jpg* and *image00312.jpg* were used.

[1] `www.cns.nyu.edu/~lcv/ssim/` (April 2013)

Berkeley Image Segmentation dataset

The Berkeley dataset [15] was originally created for the purpose of evaluating image segmentation and boundary detection algorithms. The dataset consists of 300 images of size 481×321 pixels, available both in RGB and grayscale versions. In this thesis the dataset has been used mainly for providing images with high-frequency components, in particular the images {*87065, 175043, 208001, 8143*}*.jpg* were used in Paper B.

McGill Calibrated Colour Image Database

This is a dataset with about 850 images and each image has the resolution 786×576 pixels. 9 different categories of images are available and they include for example flowers, texture, animals, fruits and landscapes scenes [16]. In this thesis the original version of the cover image (*samplepippin_jtalon0040.jpg*) as well as the image illustrating the color opponent transformation in section 3.1 (*samplemerry_flor0003.jpg*) have been obtained from this dataset.

Computed Tomography dataset

Paper C uses a Computed Tomography (CT) dataset obtained through a cooperation with CMIV, Center of Medical Imaging and Visualization, Linköping, Sweden. The database consists of 400 CT images which were captured post-mortem. The scans have been acquired using a tube voltage of 120 kV, tube current of 127 mAs and slice thickness of 1 mm.

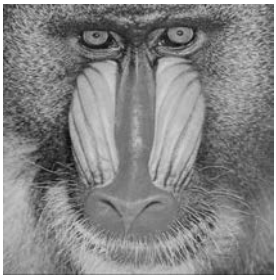
Commonly used test images

Figure 5.1 depicts some commonly used grayscale images that are used for algorithmic evaluation and testing. This dataset includes the famous cameraman, Lena and boat images among others. Paper A, makes use of the color version of Lena and Baboon.


Cameraman


Barbara


Boat


Baboon


Lena


House

Figure 5.1: Commonly used grayscale test images

Chapter 6

Concluding remarks

6.1 Results

In this thesis I have investigated image diffusion methods and presented three major contributions. In Paper A and B, it is shown that a color opponent transformation of the RGB color space yields superior results when considering image diffusion methods. The color opponent transformation is a non-Euclidean representation composed of average intensity and two components representing color balance of the RGB space.

Moreover, Paper B introduced necessary and sufficient conditions stating when a tensor-based PDE equation can be described by a corresponding energy functional. Also, it is shown that the derived Euler-Lagrange equation yields comparable results compared to state-of-the-art image denoising methods for images containing large quantities of high frequency content such as leafs, branches and pebbles.

Finally, it has been argued that the assessment of image denoising methods has to be done with respect to the intended end-user. Based on this argument an application-driven iterative filtering method was presented in Paper C for the application of medical visualization. By letting the gradient in the regularization term of the energy functional act on a visualization function rather than the value domain of the image, a non-linear diffusion scheme is derived.

6.2 Future work

So far the image filtering framework has been restricted to 2D images, where each pixel represents a scalar or vector/color intensity value. In principle, each intensity value can be considered as a sample drawn from some underlying stochastic process. Initial experiments show that utilizing knowledge of statistical properties to construct segmentation maps, obeying homogeneous regions in an image, it is possible to drive the diffusion based on density estimates. This approach is in line with the arguments presented in Paper C, namely that by estimating density

functions from the image content one obtains an application-dependent, but yet data-driven, denoising method.

From the famous saying "One image says more than a thousand words", one might say that "One video says more than a thousand images". Thus, a natural continuation of this work is to consider the case of video denoising. The step to video sequences is not trivial, two dimensional images are often assumed to primarily have spatially correlated noise, whereas video sequences also have temporally correlated noise components. We hypothesize that by propagating a segmentation map through a video sequence, one may use the previous segmentation map to recursively define density estimates to drive the image diffusion.

With regards to the tensor-based functional in Paper B, it would be desirable to investigate in more detail the case when a PDE modeled using a general tensor can be described by a corresponding functional.

Bibliography

[1] F. Åström, G. Baravdish, and M. Felsberg. On Tensor-Based PDEs and their Corresponding Variational Formulations with Application to Color Image Denoising. In *ECCV 2012: 12th European Conference on Computer Vision, 7-12 October, Firenze, Italy*, 2012.

[2] F. Åström, M. Felsberg, G. Baravdish, and C. Lundström. Visualization Enhancing Diffusion. In *SSBA Symposium on Image Analysis*, Stockholm, Sweden, 2012.

[3] F. Åström, M. Felsberg, G. Baravdish, and C. Lundström. Targeted Iterative Filtering. In *Fourth International Conference on Scale Space and Variational Methods in Computer Vision (SSVM 2013), 2-6 June 2013, Schloss Seggau, Austria*, volume 7893 of *LNCS*, pages 1–11, Springer, Heidelberg, 2013.

[4] F. Åström, M. Felsberg, and R. Lenz. Color Persistent Anisotropic Diffusion of Images. In *SCIA'11*, volume 6688 of *LNCS*, pages 262–272, Springer, Heidelberg, 2011.

[5] F. Åström, M. Felsberg, and R. Lenz. Color Persistent Anisotropic Diffusion of Images. In *SSBA Symposium 2011, 17-18 March, Linköping*, 2011.

[6] F. Åström and R. Köker. A parallel neural network approach to prediction of Parkinson's Disease. *Expert systems with applications*, 38(10):12470–12474, 2011.

[7] J. Bigün and G. H. Granlund. Optimal orientation detection of linear symmetry. In *Proceedings of the IEEE First International Conference on Computer Vision*, pages 433–438, London, Great Britain, June 1987.

[8] M. J. Black and A. Rangarajan. On the unification of line processes, outlier rejection, and robust statistics with applications in early vision. *International Journal of Computer Vision*, 19(1):57–91, 1996.

[9] M. Felsberg. Autocorrelation-driven diffusion filtering. *IEEE Transactions on Image Processing*, 20(7):1797–1806, july 2011.

[10] W. Förstner and E. Gülch. A fast operator for detection and precise location of distinct points, corners and centres of circular features. In *ISPRS Intercommission, Workshop, Interlaken, pp. 149-155.*, 1987.

[11] K. Krajsek and H. Scharr. Diffusion filtering without parameter tuning: Models and inference tools. In *2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2536–2543, 2010.

[12] F. Larsson, M. Felsberg, and P.-E. Forssen. Patch contour matching by correlating fourier descriptors. In *Digital Image Computing: Techniques and Applications, 2009. DICTA '09.*, pages 40–46, dec. 2009.

[13] R. Lenz and P.L. Carmona. Hierarchical S(3)-Coding of RGB Histograms. In *Selected papers from VISAPP 2009*, volume 68, pages 188–200. Springer, 2010.

[14] T. Lindeberg. *Scale-Space Theory in Computer Vision*. Kluwer international series in engineering and computer science: Robotics: Vision, manipulation and sensors. Springer, 1993.

[15] D. Martin, C. Fowlkes, D. Tal, and J. Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Proc. 8th Int'l Conf. Computer Vision*, volume 2, pages 416–423, July 2001.

[16] A. Olmos and F. A. A. Kingdom. A biologically inspired algorithm for the recovery of shading and reflectance images. *Perception*, 33(12):1463 – 1473, 2004.

[17] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions, PAMI*, 12:629–639, 1990.

[18] H. Scharr, M.J. Black, and H.W. Haussecker. Image statistics and anisotropic diffusion. In *Ninth IEEE International Conference on Computer Vision, 2003*, pages 840 –847 vol.2, oct. 2003.

[19] G. Sharma. Color fundamentals for digital imaging. In G. Sharma and R. Bala, editors, *Digital Color Imaging Handbook*, Electrical Engineering & Applied Signal Processing Series, chapter 1. Taylor & Francis, 2010.

[20] W.A. Strauss. *Partial Differential Equations: An Introduction*. Wiley, 2007.

[21] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004.

[22] J. Weickert. *Anisotropic Diffusion In Image Processing*. ECMI Series, Teubner-Verlag, Stuttgart, Germany, 1998.

[23] J. Weickert. Coherence-enhancing diffusion of colour images. *Image and Vision Computing*, 17(3-4):201–212, 1999.

[24] J. Weickert. Nonlinear diffusion filtering. In B. Jähne, H. Haussecker, and P. Geissler, editors, *Handbook of Computer Vision and Applications: Signal processing and pattern recognition*, number 2 in Handbook of Computer Vision and Applications. Academic Press, 1999.