DIPLOMA THESIS

Computer Engineering
May 2002

Smoothing of Matrix-Valued Data

Thomas Brox
Computer Vision, Graphics, and Pattern Recognition Group
Department of Mathematics and Computer Science
University of Mannheim
68131 Mannheim, Germany

Referees

Prof. Christoph Schnörr
Prof. Joachim Weickert

For
My parents and Ilka

I want to thank Joachim Weickert and Christoph Schnörr for their support.

I hereby certify that the work reported in this diploma thesis is my own and that work performed by others is appropriately cited.

Signature of the author:

Mannheim, 6th May 2002 Thomas Brox

Abstract

During the last decade diffusion methods became more and more popular in the fields of image processing and computer vision. They are used for smoothing and regularization in cases where discontinuity preserving properties are wanted. Since the first discontinuity preserving smoothing operator has been presented, a lot of generalizations were made in order to fit new applications. One of those generalizations was the extension from scalar-valued to vector-valued data. This diploma thesis regards a further generalization towards matrix-valued data as well as its application in the field of optic flow estimation. More specifically, it deals with the following items:

- The diffusion of matrix-valued data is derived from conventional diffusion methods.

- It is then applied to a special matrix: the frequently used *structure tensor* [FG87]. Some experiments are shown and it turns out that some modifications to the original technique make the whole process for this specific application case more robust and easier to handle. After all, an extension of the conventional linear structure tensor to a nonlinear structure tensor is obtained.

- This nonlinear structure tensor again opens a set of new applications. Actually it can be applied in all cases where a linear structure tensor is used. We concentrate here on optic flow estimation. The classic method of Lucas-Kanade [LK81, Luc84] as well as its spatio-temporal counterpart of Bigün et al. [BG88, BGW91] is improved by using the new nonlinear structure tensor.

- Since Lucas-Kanade and Bigün are only special cases of another, more general optic flow estimation technique, the CLG technique [WBS01], the nonlinear structure tensor can also be applied to improve this method. This leads to a framework unifying a whole set of differential optic flow estimation techniques. In this framework all those techniques only differ in the kind of smoothing at certain processing steps.

- Experiments demonstrate the improvements that can be achieved with the new technique in the field of optic flow estimation and show also the performance in comparison to other contemporary algorithms.

Contents

1 Introduction

Smoothing of image data plays a fundamental role in the field of image processing and computer vision. In fact it is so fundamental that hardly any method in computer vision can do without smoothing. This is because smoothing allows an exchange of information between neighboring parts of an image, thus extending the effect of local data to a wider area. Besides that information exchange, smoothing has also a regularizing effect. So smoothing can turn an ill-posed problem with a non-unique solution into a well-posed problem.

When talking of images one might first think of two-dimensional structures of scalar-valued data. Such structures are very common in form of grey-value images. But structures need not to be necessarily two-dimensional. For example a video sequence is a three-dimensional structure and especially in medical applications there are often images consisting of volume data.
The same way image data need not be necessarily scalar-valued. Taking a color image for example, each pixel consists of a color vector and is therefore vector-valued. In medical imaging there also exist matrix-valued images. They are called *DT-MRI images* and are the product of a recent image acquisition technique that measures the diffusion characteristics of water molecules in tissues.
However, there are further applications to smooth matrix-valued data besides input images themselves. Many methods in computer vision work with matrices created from the original image data. The frequently used *structure tensor* [FG87] for instance is a matrix derived from the image gradient.

The most common way to smooth arbitrary data is its convolution with a *Gaussian kernel* [Iij59]. This method is both fast and simple, and in the first place it causes no problems when being extended to data of arbitrary dimension or value. Yet, in recent years nonlinear diffusion methods have become more and more popular and seem to replace Gaussian smoothing in many fields of application. The reason is their ability to preserve discontinuities in the data, thus avoiding the blurring effect known from Gaussian smoothing.
All diffusion methods have their source in the work of Perona-Malik [PM87] which proposed isotropic diffusion for grey-value images, where smoothing stopped in the area of discontinuities. Later Weickert [Wei94] proposed an anisotropic version[1] which had the ability to smooth along discontinuities but not across them. Gerig et al. [GKKJ92] demonstrated how to use dif-

fusion for vector-valued data, and Tschumperlé and Deriche [TD01] presented the first time a technique to diffuse matrix-valued data without simply diffusing each matrix channel separately. In parts this diploma thesis is based on a technical report [WB02] presenting an anisotropic version of the isotropic technique from [TD01].
While Tschumperlé and Deriche applied their technique to the smoothing of DT-MRI images, where primarily denoising capabilities are wanted, this work concentrates on the structure tensor. For this matrix field there are further demands besides denoising capabilities. To meet the demands of various algorithms that make use of the structure tensor it must primarily be ensured to close structures of a certain scale. The conventional *linear structure tensor* obtained by Gaussian smoothing meets these demands, yet it is interesting whether a *nonlinear structure tensor* obtained by nonlinear diffusion can perform better, especially in respect to discontinuity preservation.

The true value of a new nonlinear structure tensor can only be validated by applying it to an already known algorithm using the linear structure tensor. *Optic flow estimation* techniques seem to be a good choice. Optic flow is the displacement field describing the movement of each pixel in two successive frames of an image sequence. There exist numerous methods that estimate optic flow. The most widely used techniques are differential methods which compute optic flow from spatial and temporal derivatives. One representative of this class is the classic method of *Lucas and Kanade* [LK81, Luc84] or its spatio-temporal counterpart proposed by Bigün [BG88, BGW91]. They use a linear structure tensor to cope with a problem that all optic flow techniques have in common: Optic flow can only be estimated in the direction of the image gradient but not perpendicular to it. This so-called *aperture problem* is mostly solved by smoothness assumptions. Lucas and Kanade assume the optic flow to be constant within a neighborhood of a certain size. Smoothing the data within this neighborhood solves the aperture problem and optic flow can be computed. Though with Gaussian smoothing the data has no influence on shape and size of the neighborhood. In comparison a nonlinear smoothing method could adapt the neighborhood to the data. This is done by replacing the linear structure tensor in the Lucas-Kanade method by a nonlinear structure tensor. While based on the method of Horn and Schunck [HS81] a lot of discontinuity preserving optic flow estimation techniques have been presented [AELS99, ADK99, BA91, Coh93, HB93, KTB96, Nag83, Nes93, PGPO94, Sch94, SH89, WS01a], the lack of an appropriate technique to smooth matrix-valued data nonlinearly prevented corresponding extensions to the Lucas-Kanade and Bigün method. Only Nagel and Gehrke [NG98] made an approach by using adaptive Gaussian filters. Yet, a method using diffusion techniques is an alternative with less parameters and a better theoretical foundation.

1. Unfortunately, in the literature the term “anisotropic” is often used for nonlinear isotropic diffusion. For example Perona-Malik named their technique anisotropic though it is not anisotropic in respect of the direction of smoothing but only in its magnitude. In Chapter 2 the difference between isotropic and anisotropic diffusion will be regarded in detail.

In [WBS01] it was shown by Weickert et al. that the local methods of Lucas-Kanade and Bigün can be combined with the global method of Horn-Schunck. Their CLG method therefore combined the benefits of both philosophies. Since the CLG method again uses the structure tensor inherited from the Lucas-Kanade and Bigün method, a nonlinear structure tensor can also be applied to this technique. Furthermore, as the nonlinear structure tensor has discontinuity preserving properties, it makes sense to extend the CLG technique by using also a discontinuity preserving technique for the Horn-Schunck part. Thus, based on the ideas of the CLG method and the new nonlinear structure tensor, a framework can be formulated covering the differential methods of Lucas-Kanade and Bigün as well as Horn-Schunck and all its discontinuity preserving extensions. It is shown that all these methods only use different smoothing techniques at three succeeding processing levels.

This work consists of five further chapters. Chapter 2 deals with the topic of diffusion. Here matrix-valued diffusion is presented in the context of former diffusion techniques. In Chapter 3 the findings about matrix-valued diffusion are used to construct a nonlinear structure tensor. In this connection also the conventional linear structure tensor is analyzed to get a better understanding what a nonlinear structure tensor should do. Chapter 4 deals with optic flow estimation. The existing methods of Lucas-Kanade and Bigün as well as Horn-Schunck and CLG are reviewed and extended by the new nonlinear structure tensor from the previous chapter. This results in a general differential optic flow estimation technique. Chapter 5 shows some tests in order to verify the performance of the nonlinear structure tensor in the context of optic flow estimation in comparison to its linear counterpart. Furthermore, it presents the results that can be achieved with the general technique and compares them to those of other optic flow estimation methods from the literature. The sixth chapter is a summary concluding this work and giving an outlook on possible future research in this field.

2 Diffusion

This chapter is mainly based on a lecture about partial differential equations [Wei01]. It is a review of existing diffusion techniques and the foundation for the next chapter where diffusion methods are used to develop a nonlinear structure tensor.

Diffusion is a phenomenon we can observe in our everyday's life. When turning on the heating on a cold winter day, the warmth of the heating spreads in the room, or when we fill two liquids in a vessel they mix[1]. Assumed we do not disturb the system, concentration differences of the two liquids will slowly be equilibrated and the underlying process can be described by a diffusion equation based on two principles: equilibration of concentration differences and conservation of mass.

The equilibration of concentration differences is described by Fick's law:

$$j = -D \cdot \nabla u.$$

The concentration gradient ∇u creates a flux j with a symmetric, positive definite *diffusion matrix* D.

Conservation of mass can be described in terms of the continuity equation

$$\partial_t u = -\operatorname{div} j.$$

Both equations together lead to the diffusion equation

$$\partial_t u = \operatorname{div} (D \cdot \nabla u).$$

In image processing this model of diffusion can be used to smooth an image. Instead of concentrations grey-values spread over the image and the equilibrium is reached when all pixels have got the same grey-value. Corresponding to the conservation of mass the average grey-value will be conserved. The way grey-values spread over the image depends on the diffusion matrix D. There are three cases that can be distinguished:

- *Homogeneous diffusion*, where D is the identity matrix
- *Isotropic diffusion*, where D is the identity matrix with a scalar pre-factor depending on the local image structure
- *Anisotropic diffusion*, where D can be an arbitrary symmetric, positive definite matrix depending on the local image structure.

1. except there is a chemical reason why two special liquids do not mix

Another classification of diffusion methods in image processing can be made concerning the value of the data. Until now the following cases appear in the literature:

- *Scalar-valued data*
- *Vector-valued data*
- *Matrix-valued data*

A third possibility to classify diffusion methods is the dimension of the data set. In fact it is easy to generalize diffusion to any arbitrary dimension. The only difference exists in the number of neighbors that must be considered for each data element. 2D and 3D data sets are very common in image processing and computer vision, but there are also applications with 4D data sets.

2.1 Homogeneous Diffusion

For homogeneous diffusion the diffusion matrix D is the identity matrix. So the diffusion process for an image f can be described by the following diffusion equation:

Homogeneous diffusion

$$\partial_t u = \operatorname{div}(\nabla u) = \Delta u$$
$$u(x, 0) = f(x)$$

This equation can be solved analytically yielding the unique solution

$$u(t) = K_{\sqrt{2t}} * u(0)$$

where K_σ is the Gaussian kernel with standard deviation σ.

Thus homogeneous diffusion is an linear operation and equivalent to Gaussian smoothing. There is a direct relation between the diffusion time t and the standard deviation of the Gaussian kernel σ:

$$t = \frac{1}{2}\sigma^2.$$

Since images have a finite size, boundary conditions have to be added. It is defined that there is no flux across boundaries.

2.2 Isotropic Diffusion

Isotropic diffusion has its roots in the work of Perona and Malik [PM87]. The main idea is to reduce smoothing at discontinuities in the data. The corresponding diffusion equation reads:

Isotropic diffusion

$$\partial_t u = \operatorname{div}\left(g(|\nabla u|^2)\,\nabla u\right).$$

While for homogeneous diffusion the diffusion matrix was the identity matrix for all data elements, in the isotropic case it is multiplied by a scalar value, depending on the gradient magnitude of the data. This *diffusivity* determines how strongly the data is smoothed in the respective area and is computed via a decreasing *diffusivity function* $g(s^2)$ where s equals $|\nabla u|$. Since the diffusivity depends on the evolving data, isotropic diffusion is a nonlinear process.

2.2.1 Diffusivity functions

Perona Malik proposed two diffusivity functions

Perona-Malik I

$$g(s^2) = \frac{1}{1 + s^2/\lambda^2}$$

Perona-Malik II

$$g(s^2) = \exp\left(-\frac{s^2}{2\lambda^2}\right)$$

In areas where $|\nabla u|$ is 0, g equals 1 yielding conventional Gaussian smoothing. As $|\nabla u|$ increases, smoothing is reduced resulting in the preservation of edges. In fact strong edges become even sharper as for high $|\nabla u|$ the *flux function* $g(s^2) \cdot s$ decreases again. This phenomenon is called *backward diffusion*. The contrast parameter λ determines the threshold above which backward diffusion takes place.

Further diffusivity functions mentioned in the literature are:

Weickert [Wei98]

$$g(s^2) = \begin{cases} 1 & (s^2 = 0) \\ 1 - \exp\left(-\dfrac{3.31488}{s^8/\lambda^8}\right) & (s^2 > 0) \end{cases}$$

Charbonnier [CBFAB94]

$$g(s^2) = \frac{1}{\sqrt{1 + s^2/\lambda^2}}$$

The Charbonnier diffusivity function differs from the other functions as it does not allow backward diffusion. Such diffusivities are interesting for cases where edges should not be enhanced but only be preserved. The Weickert diffusivity on the other hand decreases faster than the functions proposed by Perona and Malik leading to more segmentation-like results.

2.2.2 Regularized diffusion

Due to its edge enhancing capability nonlinear diffusion has in its original design two major drawbacks. In the continous setting there is no general well-posedness theory for diffusivity functions allowing backward diffusion. Though Weickert and Benhamouda showed in [WB97] that spatial discretization creates well-posedness, there is still another problem. Noise is misinterpreted as edges and tried to be preserved. Therefore images which are heavily distorted by noise are not smoothed at all. This problem can be removed by smoothing $|\nabla u|$ or presmoothing u before its derivatives are computed [CLMC92]. This also creates well-posedness for the continous setting. The modified diffusion equation is:

Regularized isotropic diffusion

$$\partial_t u = \operatorname{div}\left(g(|\nabla_\sigma u|^2)\,\nabla u\right) \qquad \text{with } \nabla_\sigma u = \nabla(K_\sigma * u)\,.$$

It should be mentioned that this presmoothing is not necessary for diffusivity functions without backward diffusion, because in the case of forward diffusion there is always some minimum smoothing removing the noise.

2.2.3 Discretization and implementation aspects

For implementing nonlinear diffusion the derivatives of the continuous diffusion equations have to be discretized. This is done by finite difference approximations. In general, first order approximations are used:

Discretization in space

$$\partial_{xx} u = \frac{1}{h_1}\left[\left(\frac{g_{i+1,j}+g_{ij}}{2}\;\frac{u_{i+1,j}-u_{ij}}{h_1}\right)-\left(\frac{g_{ij}+g_{i-1,j}}{2}\;\frac{u_{ij}-u_{i-1,j}}{h_1}\right)\right]+O(h_1^2)$$

$$\partial_{yy} u = \frac{1}{h_2}\left[\left(\frac{g_{i,j+1}+g_{ij}}{2}\;\frac{u_{i,j+1}-u_{ij}}{h_2}\right)-\left(\frac{g_{ij}+g_{i,j-1}}{2}\;\frac{u_{ij}-u_{i,j-1}}{h_2}\right)\right]+O(h_2^2)$$

where h_1 and h_2 are the pixel sizes in x-direction and y-direction and g_{ij} approximates $g(|\nabla_\sigma u|^2)$ in a pixel (i,j) with $\nabla_\sigma u$ approximated by $\left(\frac{u_{i+1,j}-u_{i-1,j}}{2h_1},\;\frac{u_{i,j+1}-u_{i,j-1}}{2h_2}\right)$ and u has been smoothed by a Gaussian kernel with standard deviation σ.

This scheme for two dimensions can also be written for general dimensions. Therefore the multi-dimensional data set u is rewritten as a one-dimensional vector so that each data element is represented by a single index k:

Discretization in space, arbitrary dimensional data

$$\partial_t u = \sum_{n=1}^{M} \sum_{l \in N_n(k)} \frac{g_l+g_k}{2h_n^2}\;(u_l-u_k)$$

where M is the dimension of the data set and $N_n(k)$ represents the set of neighbors of a data element k in n-direction.

A shorter matrix-vector notation is:

Discretization in space, vector-matrix notation

$$\partial_t u = A(u)u$$

$$\text{with } a_{kl} = \begin{cases} \dfrac{g_l + g_k}{2h_n^2} & (l \in N_n(k)) \\ -\displaystyle\sum_{n=1}^{M} \sum_{l \in N_n k} \frac{g_l + g_k}{2h_n^2} & (l = k) \\ 0 & (\text{else}) \end{cases}$$

Note that A is a sparse symmetric matrix where the number of rows and columns is the number of data elements.

For discretization in time there are two possibilities:

Discretization in time, explicit scheme

$$\frac{u^{k+1} - u^k}{\tau} = A(u^k)\, u^k$$

Discretization in time, semi-implicit scheme

$$\frac{u^{k+1} - u^k}{\tau} = A(u^k)\, u^{k+1}$$

Here k is the time index and τ the time step size.

The explicit scheme leads to a convolution operation or matrix-vector multiplication:

$$u^{k+1} = (I + \tau A(u^k))u^k$$

It can be proved [Wei98] to be stable for time step sizes $\tau \leq 1 / \sum_{n=1}^{M} \frac{2}{h_n^2}$. Thus for large diffusion times many iterations have to be carried out.

The semi-implicit scheme leads to a linear system:

$$(I - \tau A(u^k))u^{k+1} = u^k$$

It is stable for arbitrary large time steps [Wei98]. Therefore only one iteration step has to be carried out to obtain the result for an arbitrary diffusion time. There exist numerous numerical methods to solve such sparse linear systems like the *Jacobi method*, *Gauss-Seidel* or the *SOR method* [You71]. Though being iterative methods again, they are often faster than the explicit scheme for large diffusion times. Moreover, there are alternative semi-implicit schemes like AOS which are roughly ten times faster than an explicit scheme [WRV98].

2.3 Anisotropic Diffusion

The main problem of isotropic diffusion is that in the presence of discontinuities it reduces diffusivity in all directions. Therefore data with lots of discontinuities is not properly smoothed anymore. Anisotropic diffusion, however, reduces diffusivity merely in parallel to the gradient but still smooths perpendicular to it. The diffusion equation is:

Anisotropic diffusion [Wei94]

$$\partial_t u = \operatorname{div}\left(D(\nabla_\sigma u \nabla_\sigma u^T)\nabla u\right)$$

The matrix $D(A) = T(g(\lambda_i))T^T$ is the diffusion matrix for $A = T(\lambda_i)T^T$ where the last-mentioned expression denotes a principal axis transformation of A with the eigenvalues λ_i as the elements of a diagonal matrix and the normalized eigenvectors as the columns of the orthogonal matrix T. The diffusivity function $g(s^2)$ can be chosen the same way as for isotropic diffusion.

2.3.1 Discretization and implementation aspects

The way to find discretizations for anisotropic diffusion is basically the same as for isotropic diffusion. However, there are some further difficulties.
Let $D = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$. Then

$$\partial_t u = \operatorname{div}(D\nabla u) = \operatorname{div}\begin{pmatrix} a\partial_x u + b\partial_y u \\ b\partial_x u + c\partial_y u \end{pmatrix} = \partial_x(a\partial_x u) + \partial_x(b\partial_y u) + \partial_y(b\partial_x u) + \partial_y(c\partial_y u).$$

In addition to the terms $\partial_x(a\partial_x u)$ and $\partial_y(c\partial_y u)$ already known from isotropic diffusion there are mixed terms. Standard approximations by central differences yield:

$$\partial_x(b\partial_y u) \approx \frac{1}{2h_1}\left(b_{i+1,j}\frac{u_{i+1,j+1} - u_{i+1,j-1}}{2h_2} - b_{i-1,j}\frac{u_{i-1,j+1} - u_{i-1,j-1}}{2h_2}\right)$$

$$\partial_y(b\partial_x u) \approx \frac{1}{2h_2}\left(b_{i,j+1}\frac{u_{i+1,j+1} - u_{i-1,j+1}}{2h_1} - b_{i,j-1}\frac{u_{i+1,j-1} - u_{i-1,j-1}}{2h_1}\right)$$

Unfortunately these expressions can create negative stencil weights violating a condition needed to guarantee stability. Another more complex discretization can guarantee stability but only for condition numbers of D below 5.8284. This severely limits anisotropy. For more details see [Wei98].

In practice the above-mentioned problem is not that important. Though stability can not be guaranteed, the algorithm used to yield stable results especially for larger diffusion times. However, there is a further problem with anisotropic diffusion. The semi-implicit scheme, known from the isotropic case to be faster, may create problems. This is because for anisotropic diffusion the numerical methods to solve the linear system cannot be proved to converge. Also pure AOS schemes cannot be applied if negative stencil weights appear. Thus the best way to implement anisotropic diffusion so far is the explicit scheme that needs many iterations for large diffusion times.

Moreover, recomputing the diffusion matrix D takes more computation time in the anisotropic case. For two-dimensional, scalar-valued data this additional effort is still comparatively small, since the principal axis transformation need not be computed explicitly (the eigenvectors are already determined by the direction of the gradient). Yet for higher dimensions, vector-valued or matrix-valued data, the explicit computation of the principal axis transformation cannot be avoided anymore.

2.3.2 A faster approximation for the explicit scheme

Addressing the last-mentioned item an approximation of the explicit scheme has been used for this work. Instead of recomputing the diffusion matrix after each iteration step, the number of recomputations is reduced in the course of time. This idea is based on the observation that most of the diffusion time is needed to smooth the data within nearly homogeneous areas, while the step by step removing of discontinuities due to the recomputation of the diffusion matrix can in a certain scope also be accelerated by choosing a higher value for the contrast parameter λ.

To get an idea of how the recomputation frequency can be reduced we regard diffusion of a signal consisting of two pixels only. For such a case analytic solutions are possible. Let u_1 and u_2 be the grey-values of those two pixels with $u_1 > u_2$. Within nearly homogeneous areas diffusivity is approximately 1. Thus the diffusion equation for such areas is

$$\partial_t u = \partial_{xx} u$$

$$\partial_t u_1 = \frac{1}{h}\left(\frac{u_2 - u_1}{h} - \frac{u_1 - u_0}{h}\right) = \frac{u_2 - u_1}{h^2}$$

$$\partial_t u_2 = \frac{1}{h}\left(\frac{u_3 - u_2}{h} - \frac{u_2 - u_1}{h}\right) = -\frac{u_2 - u_1}{h^2}$$

with pixel size h and the boundary conditions $u_3 - u_2 = u_1 - u_0 = 0$ (u_0 and u_3 are only dummy pixels).
With $w := (u_1 - u_2)$ this leads to the differential equation

$$\partial_t w = -\frac{2}{h^2} w$$

with the solution

$$w(t) = w(0) \cdot e^{-\frac{2}{h^2}t}.$$

This shows that the difference in grey-value between the two pixels decreases exponentially and can be seen as a motivation for exponentially reducing the number of recomputations in the course of time. Such a procedure causes the same result with lower computational costs, if the contrast parameter is raised. This is because each update of the diffusion matrix causes increases in diffusivity mainly in areas where the magnitude of the gradient is around the contrast parameter. Thus, less updates yield discontinuities with such a magnitude of the gradient to be preserved longer. On the other hand, raising the contrast parameter yields those discontinuities to be removed earlier what leads to nearly the same result as the original method but with less computational costs.

2.3.3 Summary

As now all diffusion techniques for scalar-valued data have been presented, the differences between homogeneous, isotropic and anisotropic diffusion are briefly summarized. Figure 1 depicts those differences. While homogeneous diffusion removes all the noise but also blurrs important image structures, isotropic diffusion leaves some noise at discontinuities. Since anisotropic diffusion can smooth along discontinuities, it can also remove the noise there, thus yielding the best results.
Regarding the computational effort homogeneous diffusion is the best choice as it can be implemented by a convolution with a Gaussian kernel. Isotropic diffusion is more expensive, though there are concepts to speed up computation using semi-implicit schemes and algorithms from numerical mathematics. The superior results of anisotropic diffusion have to be payed with higher computational costs, since fast algorithms that significantly speed up computation are not yet available.

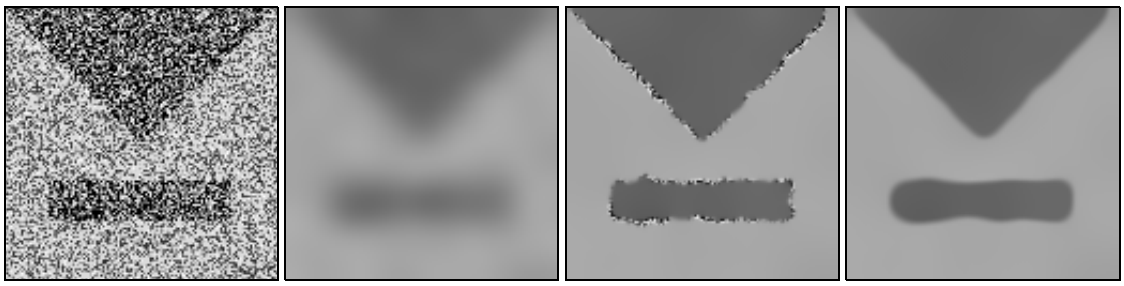

Figure 1: Different Types of Diffusion.
FROM LEFT TO RIGHT:
(a) Image heavily distorted by noise. (b) Homogeneous diffusion.
(c) Isotropic diffusion. (d) Anisotropic diffusion.

2.4 Vector-valued Diffusion

The immediate concept to adapt diffusion methods for vector-valued data is to smooth each vector channel separately. In fact for homogeneous diffusion (or Gaussian smoothing) this is an appropriate procedure. For isotropic and anisotropic diffusion, however, there arises a problem: By diffusing each channel separately, structures can develop at different locations in the channels. For the case of a color image for example this means that an edge for the blue value is being located at a different location than the edge in the channels for the red and green value.

A solution to this problem is the coupling of the channels by using a common diffusion matrix for all of them. This common diffusion matrix again is determined by using the data of all channels. This is done by summing the values of all channels before passing this sum to the diffusivity function.

Thus, the resulting diffusion equations for the isotropic and anisotropic case are:

Vector-valued isotropic diffusion [GKKJ92]

$$\partial_t u_i = \operatorname{div}\left(g\left(\sum_{k=1}^{n} |\nabla_\sigma u_k|^2 \right) \nabla u_i \right) \qquad i = 1 \ldots n$$

Vector-valued anisotropic diffusion [Wei94]

$$\partial_t u_i = \operatorname{div}\left(D\left(\sum_{k=1}^{n} (\nabla_\sigma u_k)(\nabla_\sigma u_k)^T \right) \nabla u_i \right) \qquad i = 1 \ldots n$$

Note: The values in the different channels are required to have comparable intensity ranges, otherwise the sum has to be weighted respecting the differences in the intensity range.

2.5 Matrix-valued Diffusion

Matrix-valued diffusion can be derived directly from vector-valued diffusion. Again a coupling of the channels has to be made for structures to develop at the same locations in all channels. Actually the only difference to vector-valued diffusion is the two-dimensional arrangement of the channels. This leads to the following diffusion equations:

Matrix-valued isotropic diffusion [TD01]

$$\partial_t u_{ij} = \operatorname{div}\left(g\left(\sum_{k,l=1}^{n,m} |\nabla_\sigma u_{kl}|^2 \right) \nabla u_{ij} \right) \qquad \begin{matrix} i = 1 \ldots n \\ j = 1 \ldots m \end{matrix}$$

Matrix-valued anisotropic diffusion [WB02]

$$\partial_t u_{ij} = \operatorname{div}\left(D\left(\sum_{k,l=1}^{n,m} (\nabla_\sigma u_{kl})(\nabla_\sigma u_{kl})^T \right) \nabla u_{ij} \right) \qquad \begin{matrix} i = 1 \ldots n \\ j = 1 \ldots m \end{matrix}$$

Tschumperlé and Deriche [TD01] added a reprojecting step to their technique to maintain the property of positive semidefiniteness while diffusing such matrix fields. However, in [WB02] this step is proved to be unnecessary, since the coupling of the channels with a common diffusion matrix always ensures the positive semidefiniteness property to be preserved.

3 Structure Tensor

In the last chapter a method for nonlinear smoothing of matrix-valued data was shown. This chapter will now deal with an application for this new method: the structure tensor. First the conventional linear structure tensor is reviewed, and its properties are carefully examined. This examination of the existing method allows an exact definition of what an enhanced structure tensor at least has to perform and what improvements could be possible. In the second part of this chapter such an enhanced, nonlinear structure tensor is developed and some basic tests are performed.

3.1 Linear structure tensor

The linear structure tensor was first mentioned by Förstner and Gülch [FG87]. It is a matrix field containing orientation and magnitude of structures for each pixel of an image f. The structure tensor is computed from the image gradient by the tensor product $\nabla f \nabla f^T$. Although this tensor product contains merely the same information as the gradient itself, it has the big advantage that it can be smoothed without cancellation effects for areas where gradients have opposite signs. This smoothing stabilizes the orientation information. The smoothed matrix field is called structure tensor.

Linear structure tensor

$$J_\rho = K_\rho * (\nabla f \nabla f^T)$$

3.1.1 Information content

Besides the information about orientation and magnitude of structures already present in the gradient, the structure tensor contains a further information: the coherence. This additional information has been gained by the smoothing process. Orientation, magnitude as well as coherence can all be determined from the structure tensor by a principal axis transformation $J_\rho = T(\lambda_i)T^T$, where the eigenvectors of J_ρ are the columns of T and the eigenvalues λ_i are the elements of a diagonal matrix.
The eigenvector to the largest eigenvalue then determines the orientation of the structure, while the Frobenius norm $\|J_\rho\| = \sqrt{\sum_i \lambda_i^2}$ determines its magnitude. The coherence is expressed by the condition number of J_ρ (largest eigenvalue against smallest eigenvalue) or by the measure $(\lambda_1 - \lambda_2)^2$.

The magnitude and coherence information can be used for a structure analysis. Homogeneous areas in the image yield the magnitude to be small. In areas around edges the structure tensor has a high magnitude as well as a high coherence, while corners result in a high magnitude but low coherence.

3.1.2 Noise removal

Although it makes often more sense to remove noise of the image f already in the image itself before computing the structure tensor, it is also removed in the structure tensor due to its smoothing step. This is demonstrated in Figure 2. Figure 2a shows a synthetic test image distorted by Gaussian noise with $\sigma = 30$. In Figure 2b the matrix product $J_0 = \nabla f \nabla f^T$ is depicted as a colored orientation plot, where the direction of the eigenvector to the largest eigenvalue is mapped to the hue value and the largest eigenvalue itself is mapped to the intensity value of the HSI color model. The saturation value is set to its maximum. Figure 2c shows the linear structure tensor. Most of the noise of Figure 2b has been removed here.

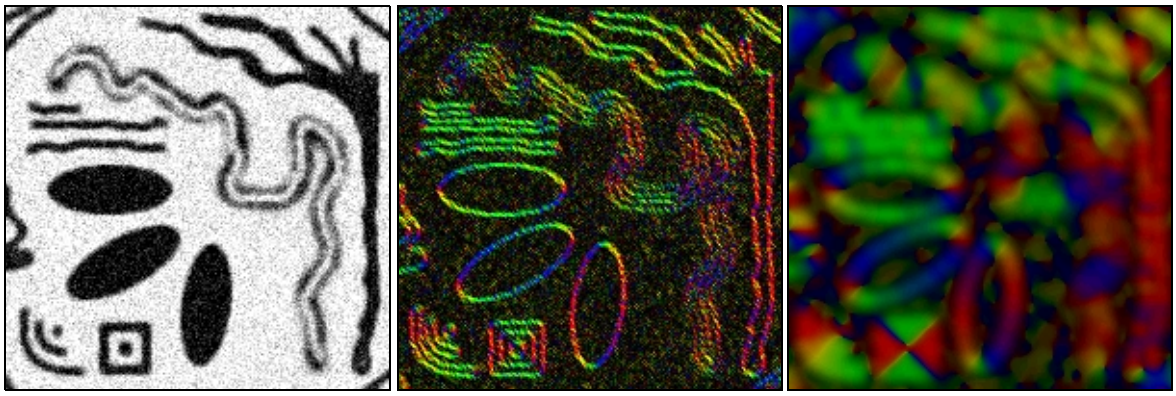

Figure 2: Synthetic test image
FROM LEFT TO RIGHT:
(a) Image f distorted by Gaussian noise with $\sigma = 30$. (b) Tensor product $\nabla f \nabla f^T$.
(c) Linear structure tensor J_ρ with $\rho = 3$.

3.1.3 Closing of structures

In many applications for the structure tensor it is desirable that structures of a certain scale are closed. This means there has to be a filling effect of orientation information from structured areas into areas without structure as far as these areas are small in respect to a certain scale. By means of the structures in the lower left of Figure 2 it can be seen that the linear structure tensor fulfills this requirement appropriately.

3.1.4 Preservation of orientation discontinuities

Figure 2c, especially the rectangular structure in the lower left, also reveals the preservation of orientation discontinuities. One would not expect such a property from a smoothing method without explicit discontinuity preservation. In fact, discontinuities in the matrix field are not preserved. The preservation of orientation discontinuities, however, is caused by its special representation as a tensor product. Two neighboring orientations with the same magnitude and a gap of 90 degree even would not influence each other at all.

Proof

Consider two perpendicular vectors $v_1 = \begin{pmatrix} x \\ y \end{pmatrix}$ and $v_2 = \begin{pmatrix} -y \\ x \end{pmatrix}$.

Their matrix products are then

$\begin{pmatrix} a_1 & b_1 \\ b_1 & c_1 \end{pmatrix}$ and $\begin{pmatrix} a_2 & b_2 \\ b_2 & c_2 \end{pmatrix}$ with

$a_1 = x^2 \qquad b_1 = xy \qquad c_1 = y^2$

$a_2 = y^2 \qquad b_2 = -xy \qquad c_2 = x^2$

Thus the weighted average of both matrix fields with weight $\alpha > \frac{1}{2}$ is

$\begin{pmatrix} a & b \\ b & c \end{pmatrix}$ with

$a = \alpha x^2 + (1-\alpha)y^2$

$b = (2\alpha - 1)xy$

$c = \alpha y^2 + (1-\alpha)x^2$

The larger eigenvalue λ of this matrix is

$$\lambda = \frac{a+c}{2} + \sqrt{\left(\frac{a+c}{2}\right)^2 + b^2 - ac} = \alpha(x^2 + y^2)$$

The eigenvector to the largest eigenvalue then is $c \cdot \begin{pmatrix} u \\ 1 \end{pmatrix}$ with $c \in I\!R$ and

$$u = \frac{b}{\lambda - a} = \frac{(2\alpha - 1)xy}{\alpha(x^2 + y^2) - \alpha x^2 - (1-\alpha)y^2} = \frac{(2\alpha - 1)xy}{(2\alpha - 1)y^2} = \frac{x}{y}$$

Thus the resulting vector has the same orientation as v_1.

Figure 3 shows the equilibration velocity $\varphi^1 - \varphi^0$, where φ^0 and φ^1 are the difference in orientation before and after the smoothing respectively, in dependence of φ^0. For small differences the velocity is small, since there is almost nothing to equilibrate. It rises until a maximum is reached for differences of about 60 degree. Then the curve falls again, and for 90 degree there is no equilibration at all.

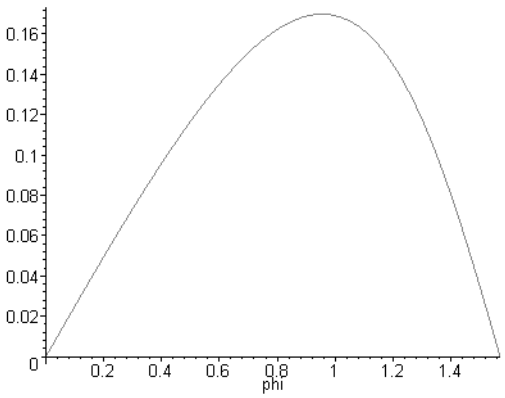

Figure 3: Equilibration velocity of two orientations with an orientation difference of *phi* (rad).

However, this holds only for the equilibration between two data elements with the same magnitude not influenced by further neighbors. In fact, many data elements with the same orientation can cause a single data element to adapt its orientation. Otherwise noise could not be removed. Also a data element with a low magnitude will easily adapt to a neighbor with a higher magnitude. In these cases the order of the eigenvalues changes, causing the eigenvector to the largest eigenvalue to turn by 90 degree.

Nevertheless, there remains a discontinuity preserving property enabling two regions to keep different orientations without blurring effects during the smoothing process.

3.1.5 Dislocation and blurring effects

While orientation discontinuities are preserved rather well, discontinuities in the magnitude are removed. This blurring effect is typical for Gaussian smoothing and can also be observed in Figure 2. Edges disappear with increasing ρ and the remaining edges dislocate. In fact this effect is the most important drawback of the linear structure tensor and motivation for the construction of a nonlinear structure tensor.

3.2 Nonlinear structure tensor

A nonlinear structure tensor with discontinuity preserving capabilities addressing the blurring effect of the conventional linear structure tensor can be built by using matrix-valued nonlinear diffusion instead of Gaussian smoothing (or homogeneous diffusion). However, it has to be ensured that the removal of one drawback will not harm all the positive properties of the linear structure tensor. So it is conceivable that the diffusion method must be adapted specifically to its application on the structure tensor. In the previous section of this chapter all important properties of the linear structure tensor were discussed in detail. They will now serve as a kind of duty book for the design of an appropriate nonlinear structure tensor.

3.2.1 What a nonlinear structure tensor has to accomplish

- **Noise removal.** Although image noise can also be removed in the image itself before computing the structure tensor, a method containing a smoothing step should be able to handle noise.
- **Closing of structures.** Small areas without any structure should be filled with the structure information of bordering structured areas. A single parameter (like ρ in the linear case) should determine how small areas must be to be affected by this filling effect yielding a scale space property.
- **Discontinuity preservation.** Orientation discontinuities as well as discontinuities in the magnitude should be preserved. In the best case the same parameter as mentioned above would also determine whether a discontinuity is worth preserving or whether it should be removed. This would lead to a scale space which respects closing of structures as well as discontinuity preservation.
- **Few robust parameters.** There should be as few parameters as possible. In the best case there would be a single scale parameter like for the linear structure tensor. Further parameters should be at least robust against moderate variations so they can be fixed for a whole set of input data.

3.2.2 A nonlinear structure tensor using the standard diffusion technique

A straight-forward approach to a nonlinear structure tensor is to apply the matrix-valued diffusion methods of the previous chapter. Starting with the matrix field $\nabla f \nabla f^T$ of an image f, nonlinear diffusion with diffusion time t yields the nonlinear structure tensor J_t. Figure 4 depicts the results using isotropic and anisotropic diffusion respectively.

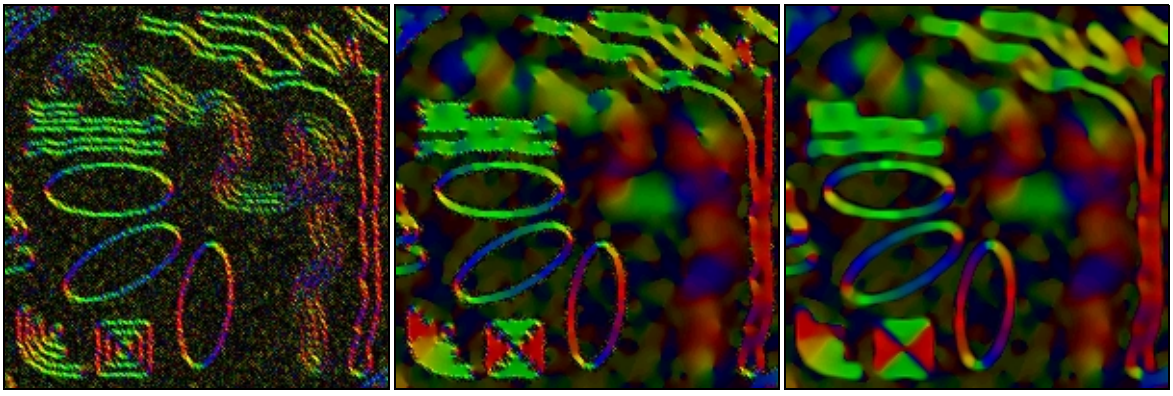

Figure 4: Nonlinear structure tensor using the standard diffusion technique.
FROM LEFT TO RIGHT:
(a) Matrix field $J_0 = \nabla f \nabla f^T$. (b) J_t obtained by isotropic diffusion and $t = 6$.
(c) J_t obtained by anisotropic diffusion and $t = 6$.

For both cases the Weickert diffusivity function with $\lambda = 300$ and
$\sigma = 2.24$ was used.

It is clearly visible that the anisotropic version of the nonlinear structure tensor outperforms the isotropic one. The isotropic version shows the typical effect of noise not removed at discontinuities. This is obvious because at discontinuities isotropic diffusion reduces diffusivity in all directions.
However, it can also be observed that there is another drawback of isotropic diffusion when being applied to the structure tensor. The closing of structures is far more problematic than in the anisotropic case. Especially at orientation discontinuities there is no appropriate closing of structures. In fact there is no scale space property in this respect because the closing is irregular. Near orientation discontinuities unstructured areas must be smaller than elsewhere in the image to be closed.
Both the isotropic and anisotropic version preserve orientation discontinuities what is not surprising as even homogeneous diffusion has this property. Discontinuities in the magnitude are mainly preserved, yet the diagonal structure with lower contrast dislocates in both versions. While for the anisotropic case there is still hope that it is possible to fix this by choosing a smaller diffusion time or a lower contrast parameter λ, there is no chance to find parameters for the isotropic case that both close the remaining structures and preserve the discontinuity of the low contrast structure.
Thus summarizing the first three items of the above-mentioned requirements anisotropic diffusion can more or less fulfill all of them while for isotropic diffusion there has to be made a choice between closing of structures and discontinuity preservation. For applications like optic flow estimation this is unacceptable.

We now focus on the anisotropic version of the nonlinear structure tensor as it performed quite well for the first three items. In order to optimize the discontinuity preservation capabilities and to verify the performance in respect to the requirement of few robust parameters the parameter settings are regarded in detail.

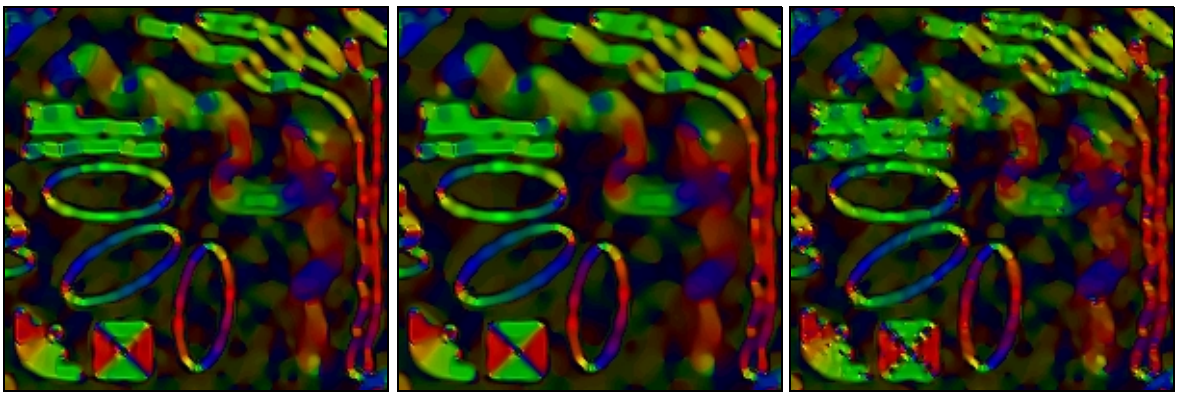

Figure 5: Parameter variations.
FROM LEFT TO RIGHT:
(a) $t = 6, \lambda = 150, \sigma = 2.24$ (b) $t = 12, \lambda = 150, \sigma = 2.24$
(c) $t = 6, \lambda = 300, \sigma = 2.12$

In Figure 5a the contrast parameter λ has been reduced in order to improve discontinuity preservation for the low contrast structure. Although an improvement could be achieved, there is still some blurring effect left. Nevertheless, the closing of structures is already affected. Moreover, noise removal is also reduced in some areas.
Figure 5b shows that a higher diffusion time can not completely make up for the drawbacks induced by the lower contrast parameter. This indicates that there is no perfect parameter setting for the test image that closes all structures appropriately without any blurring effects at the structure with lower contrast.
Figure 5c reveals the parameter settings also to be quite sensitive to variations. A slight change in the setting of the presmoothing parameter σ leads to a considerably worse result. The contrast parameter λ is much more robust. However, it still has to be adapted to the image and can not be fixed for a whole set of input data. So the requirement of only a few robust parameters can not be fulfilled using standard matrix-valued diffusion techniques. Although, besides these issues, the technique works quite well, it is therefore reasonable to try some adaptations in order to remove the remaining drawbacks.

3.2.3 A nonlinear structure tensor using an adapted diffusion technique

As we concentrate here merely on the structure tensor, it is possible to adapt the diffusion technique specifically to this sort of matrix field. One major problem that arises using the standard technique is the double preservation of orientation discontinuities. While orientation discontinuities are already preserved for homogeneous diffusion due to the special representation of orientations as a tensor product, they are preserved a second time due to the nonlinear diffusion technique which handles orientation discontinuities the same way as discontinuities in the magnitude. This results in an irregular closing of structures and noise reduction, what becomes most obvious for the isotropic version of the nonlinear structure tensor but also affects the anisotropic case (see for example Figure 5c). In fact, also the sensitivity to parameter variations arises from this double preservation of orientation discontinuities, since the parameter values are either too high for areas with orientation discontinuities or too low for areas without them.

Therefore the best way to improve the results of the nonlinear structure tensor is to avoid this double preservation of orientation discontinuities. This can be achieved by steering the diffusivity not by the complete matrix field, which also contains the orientation information, but only by its magnitude. The diffusion equations for the isotropic and anisotropic case then are:

Isotropic diffusion

$$\partial_t u = \operatorname{div}\left(g\left(\left| \nabla_\sigma \left(\sqrt[4]{\sum_{k,l} u_{kl}^2} \right) \right|^2 \right) \right)$$

Anisotropic diffusion

$$\partial_t u = \operatorname{div}\left(D\left(\left[\nabla_\sigma \left(\sqrt[4]{\sum_{k,l} u_{kl}^2} \right) \right] \left[\nabla_\sigma \left(\sqrt[4]{\sum_{k,l} u_{kl}^2} \right) \right]^T \right) \right)$$

In these equations the fourth root seems to be quite arbitrary. Though there is a good motivation for it. For diffusion time $t = 0$ the structure tensor of an image f is

$$J_0 = \begin{pmatrix} f_x^2 & f_x f_y \\ f_x f_y & f_y^2 \end{pmatrix}$$

with subscripts denoting partial derivatives. Thus the corresponding expression of the diffusion equation is

$$\sqrt[4]{\sum_{k,l} u_{kl}^2} = \sqrt[4]{f_x^4 + 2f_x^2 f_y^2 + f_y^4} = \sqrt[4]{(f_x^2 + f_y^2)^2} = \sqrt{f_x^2 + f_y^2} = |\nabla f|$$

what can be interpreted as the image gradient driving the diffusion. However, this is only exactly the case for $t = 0$. The diffusivity is adapted to the new structure tensor after each time step. So it is still a nonlinear diffusion process. Note that also the positive semidefiniteness of the matrix field is still preserved with this technique, since there is still a coupling of all matrix components [WB02].

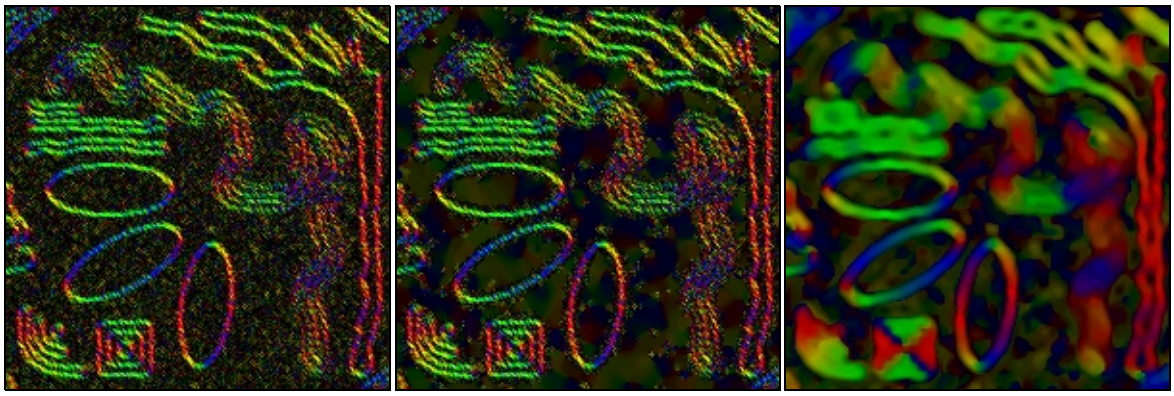

Figure 6: Adapted nonlinear structure tensor.
FROM LEFT TO RIGHT:
(a) Matrix product $J_0 = \nabla f \nabla f^T$. (b) Nonlinear structure tensor J_t with isotropic diffusion.
(c) Nonlinear structure tensor J_t with anisotropic diffusion.

Figure 6 depicts the results achieved with the new nonlinear structure tensor. Obviously the isotropic technique does not produce satisfactory results as there is almost no smoothing. An optimization of the parameter setting did also not yield any improvements. On the other hand, the anisotropic version performs rather well. The reason is obviously the fact that anisotropic diffusion is able to smooth along edges while isotropic diffusion is not. The fact that we use an edge detector (the gradient magnitude) to drive the diffusion process therefore explains the bad performance of the isotropic technique. However, it is still surprising that the anisotropic version also closes structures very well though it should not be able to smooth across discontinuities. This is because discrete anisotropic diffusion mostly cannot adapt exactly to the direction of a discontinuity due to the pixel grid. Thus anisotropic diffusion always smooths slightly across discontinuities. Exceptions are synthetic images where all edges perfectly fit to the pixel raster. An example is shown in Figure 7. However, in real image data such structures do not occur.

Figure 7: Synthetic image

Nevertheless, Figure 6c looks somehow distorted in some regions. However, some smoothing of the diffusivity has revealed to be suitable to avoid these unpleasant effects. Already some slight smoothing is sufficient to remove them. Throughout this work a Gaussian kernel with a standard deviation of $\sqrt{2}$ has been used. In Figure 8 the resulting nonlinear structure tensor is depicted together with its linear counterpart.

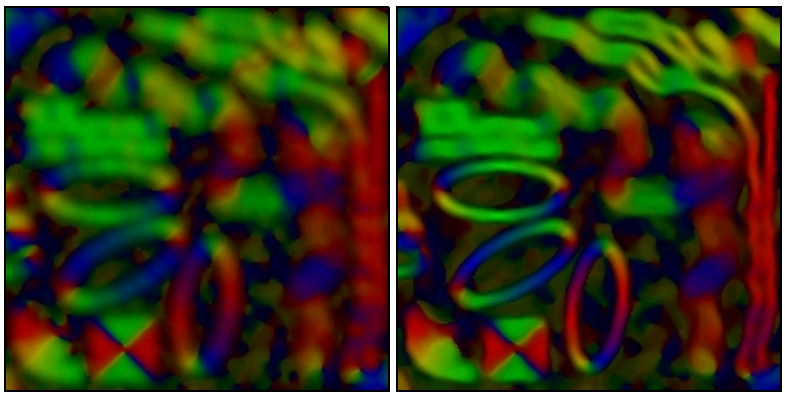

Figure 8: Linear and nonlinear structure tensor
LEFT: Linear structure tensor. RIGHT: Nonlinear structure tensor

Now that a version of the nonlinear structure tensor using an adapted diffusion technique has been developed, it is time to review whether it can fulfill the requirements listed in section 3.2.1.

- **Noise removal.** In Figure 8 it is visible that the nonlinear structure tensor can handle noise the same way as the linear original.
- **Closing of structures.** It can also be seen that structures of a certain scale are closed appropriately. Moreover, in Figure 11 it becomes obvious that the scale can be steered by a single parameter, namely the diffusion time *t*.
- **Discontinuity preservation.** In contrast to the linear structure tensor, which only preserves orientation discontinuities, the nonlinear structure tensor also preserves discontinuities in the magnitude. This was actually the motivation for the construction of a nonlinear structure tensor and it can be seen that this goal could be achieved.
- **Few robust parameters.** The main reason why the diffusion technique was adapted specifically to the structure tensor was the requirement of parameters to be more robust. Since the linear structure tensor has only one single robust parameter (the standard deviation of the Gaussian kernel), a similar outcome for the nonlinear structure tensor is necessary to be useful for the same set of applications. Figure 9 and Figure 10 reveal the additional parameters of the nonlinear structure tensor to be so robust that they can be fixed and be seen as constants. All experiments for this work (also those of the following chapters) have been performed with the same diffusivity function (Weickert diffusivity), the same contrast parameter $\lambda = 0.1$ and the same presmoothing parameter $\sigma = 1.58$. Moreover, Figure 11 depicts the scale space property of the nonlinear structure tensor with the scale determined by the last remaining parameter. The same way as the linear structure tensor can be steered merely by the parameter ρ the diffusion time *t* determines the result of the nonlinear structure tensor.

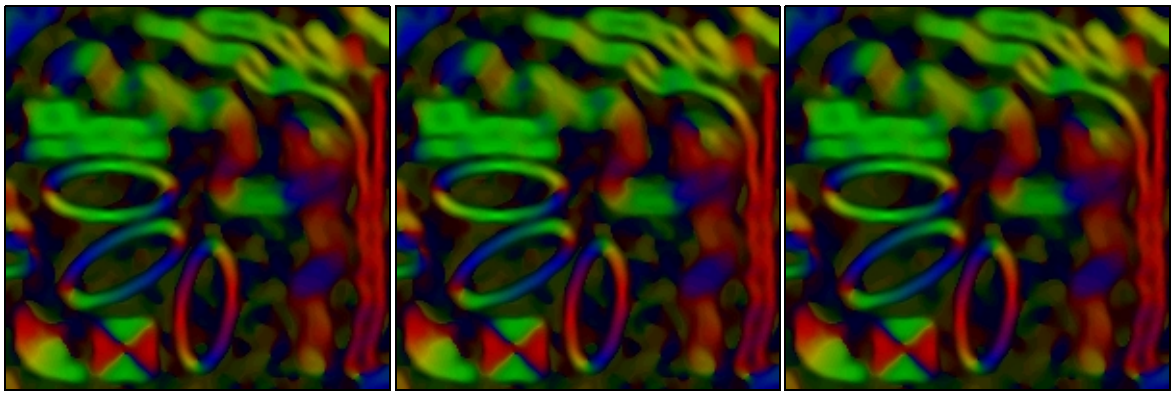

Figure 9: Variation of the contrast parameter λ.
FROM LEFT TO RIGHT:
(a) $\lambda = 0.01$. (b) $\lambda = 0.1$. (c) $\lambda = 1$.
All other parameters were kept fixed at $t = 15$ and $\sigma = 1.58$.

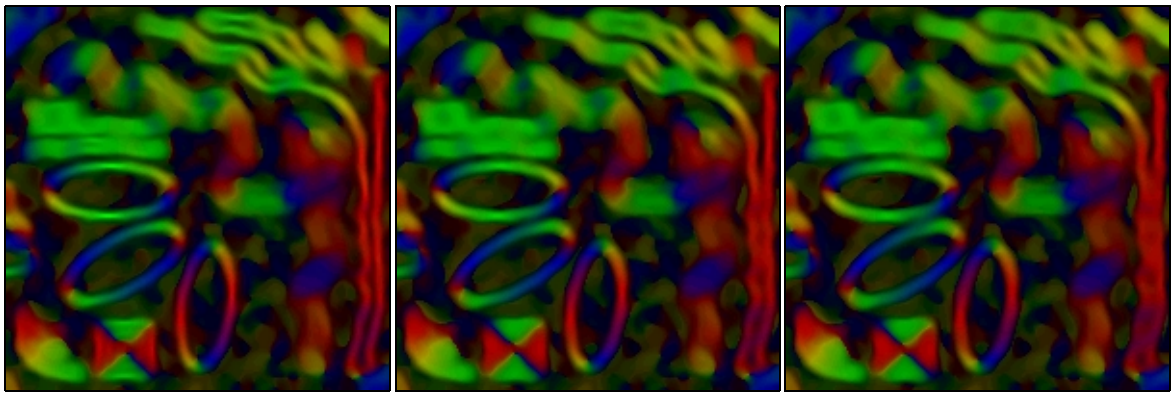

Figure 10: Variation of the presmoothing parameter σ.
FROM LEFT TO RIGHT:
(a) $\sigma = 1$. (b) $\sigma = 1.58$. (c) $\sigma = 2$.
All other parameters were kept fixed at $t = 15$ and $\lambda = 0.1$.

3.2.4 Summary

Summarizing this chapter it was possible to design a new nonlinear structure tensor preserving all useful properties of the conventional linear structure tensor including its convenient usage with only one decisive parameter. Moreover, as a further property, the nonlinear structure tensor preserves discontinuities in the magnitude. Some basic tests were performed in this chapter, yet only testing within the environment of a real application would show the effects of the new technique. This is done in the next chapters.

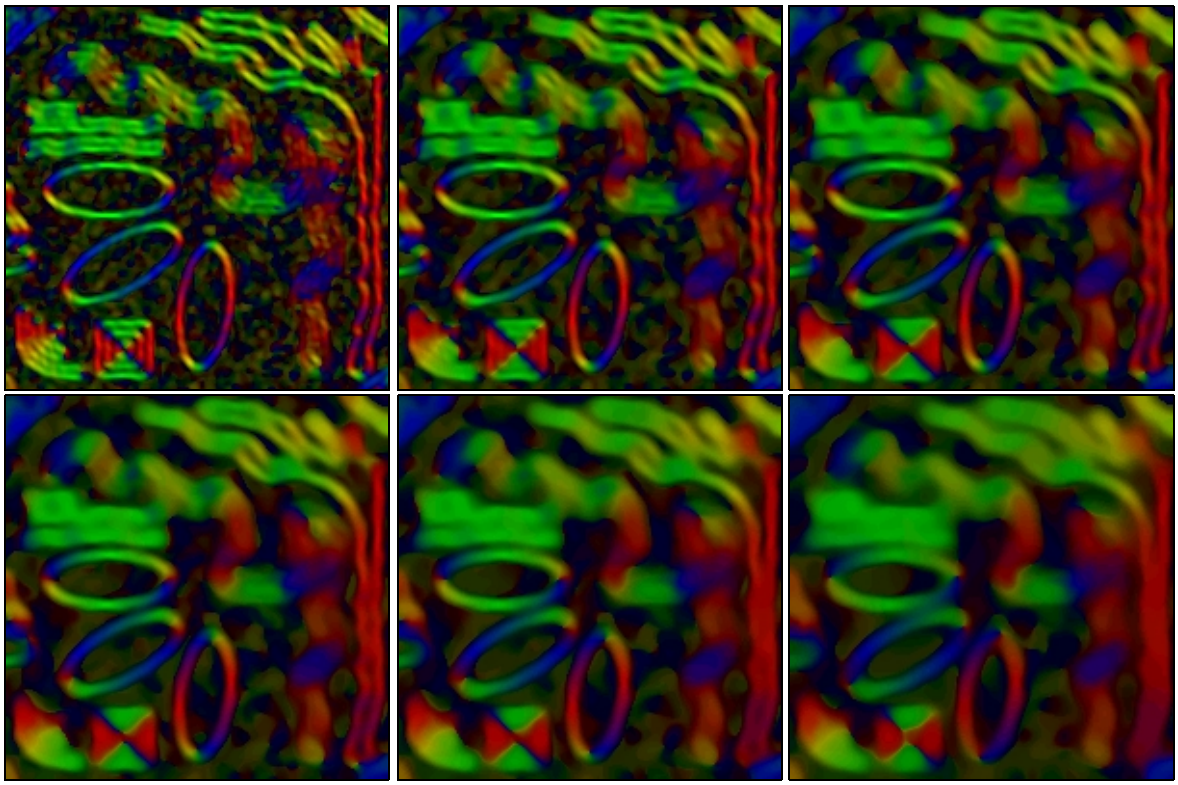

Figure 11: Variation of the diffusion time t.
UPPER ROW FROM LEFT TO RIGHT:
(a) $t = 2$. (b) $t = 5$. (c) $t = 10$.
LOWER ROW FROM LEFT TO RIGHT:
(d) $t = 15$. (e) $t = 25$. (f) $t = 50$.
All other parameters were kept fixed at $\lambda = 0.1$ and $\sigma = 1.58$.

4 Optic Flow Estimation

In the previous chapter a nonlinear structure tensor has been developed. In order to verify its performance in real world applications optic flow estimation as a popular application field of computer vision will be considered in this chapter. There is one classic differential optic flow estimation technique using a structure tensor. Is was introduced 20 years ago by Lucas and Kanade [LK81,Luc84]. A spatio-temporal version of this technique was proposed by Bigün et al. [BG88,BGW91]. In the first part of this chapter both methods will be reviewed in order to apply and test the nonlinear structure tensor later.

Recently, Weickert et. al [WBS01] used the structure tensor for an extension of the Horn and Schunck method [HS81]. They also demonstrated that the Lucas-Kanade/Bigün as well as the Horn-Schunck approach are both special cases of their new CLG technique. In this chapter the nonlinear structure tensor will also be applied to the CLG method, and therefore the techniques of Horn-Schunck and CLG will be reviewed in the first part as well.

Based on the discussion about the role of smoothing in the field of differential optic flow estimation techniques [WBS01] as well as a framework for convex regularizers [WS01a] a general differential optic flow estimation technique will be presented at the end of this chapter. This general method unifies the Lucas-Kanade/Bigün approach, along with its adaptation using the nonlinear structure tensor, and the Horn-Schunck approach, along with its discontinuity preserving adaptations. The result can also be interpreted as a discontinuity preserving version of the CLG method.

First some notions concerning optic flow estimation are introduced. Consider an image sequence $f(x, y, z)$ with z standing for the time axis and let $(u, v)^T(x, y, z)$ be the searched optic flow field. Corresponding image objects at times z and $z + 1$ are assumed to have the same grey-value. This results in the original *optic flow constraint*:

$$f(x + u, y + v, z + 1) = f(x, y, z)$$

For small displacements this can be linearized by a first order Taylor expansion leading to the linearized optic flow constraint:

$$f_x u + f_y v + f_z = 0$$

where subscripts denote partial derivatives.

As this is only one equation for two flow components, the optic flow is not uniquely determined by this constraint. Only the displacement parallel to the image gradient, the so-called normal flow, can be computed. This is called the *aperture problem*. An additional assumption has to be made in order to obtain a unique solution for the optic flow field.

While most optic flow estimation techniques have the linearized or at least the original optic flow constraint in common, they differ in the way how they circumvent the aperture problem.

4.1 Review of existing techniques

In this section, which is based on [WBS01], [Bru01] and [WS01a], the existing techniques used for this work are reviewed. All these techniques belong to the class of differential optic flow estimation methods and can be covered by the general method introduced later in section 4.3.

4.1.1 Lucas-Kanade and Bigün

First the methods of Lucas-Kanade [LK81,Luc84] and Bigün et al. [BG88,BGW91] are considered. Their idea to cope with the aperture problem is to assume the optic flow vector to be constant in a neighborhood B_ρ of size ρ. Since the aperture problem is circumvented by using local assumptions, Lucas-Kanade and Bigün are so-called *local* flow estimation techniques.

According to Lucas-Kanade, for each point (x_0, y_0) the optic flow can be estimated by a least square fit minimizing the local energy function

$$E_{LK}(u, v) = \frac{1}{2} \int_{B_\rho(x_0, y_0)} (f_x u + f_y v + f_z)^2 dx dy.$$

Instead of the sharp window B_ρ often a convolution with a Gaussian kernel K_ρ is used yielding

Lucas-Kanade

$$E_{LK}(u, v) = K_\rho * (f_x u + f_y v + f_z)^2.$$

A minimum (u, v) of E satisfies $\partial_u E = 0$ and $\partial_v E = 0$, leading to the linear system

$$\begin{pmatrix} K_\rho * f_x^2 & K_\rho * (f_x f_y) \\ K_\rho * (f_x f_y) & K_\rho * f_y^2 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -K_\rho * (f_x f_z) \\ -K_\rho * (f_y f_z) \end{pmatrix}.$$

Provided the system matrix is not singular this linear system can be solved. Such singular matrices appear in regions where the image gradient vanishes or the aperture problem remains present, because the direction of the gradient is all the same within the neighborhood. Using sufficiently broad Gaussian filters avoid such situations to appear in practice. However, the smaller eigenvalue of the system matrix can still become close to zero resulting in an uncertain estimation of the optic flow in such regions. Therefore the smaller eigenvalue of the system matrix is often used as a confidence measure. For points where the value becomes too small the estimated optic flow is rejected yielding non-dense flow fields.

Although the methods of Lucas-Kanade and Bigün make the same assumptions to circumvent the aperture problem, they differ in the way how the least square fit is performed. Bigün formulates the minimization problem as follows:

Bigün

$$E_{LK}(u, v) = w^T J_\rho w$$

with $w^T = (u, v, 1)$ and J_ρ denoting the linear structure tensor $K_\rho * (\nabla f \nabla f^T)$. It is easy to see that both formulations are equivalent.
Bigün minimizes $E(u, v)$ by performing a principal axis transformation of the structure tensor. The eigenvector w to the smallest eigenvalue then determines the optic flow vector (the third component of w is normalized to 1). Again there is a confidence measure, which allows to identify uncertain estimations in order to sparsify the flow field. If the aperture problem is still relevant for some points, the second smallest eigenvalue becomes close to zero and can therefore serve as confidence measure. Additionally, the Bigün approach allows to identify uncertain estimations due to noise or discontinuities. In such cases the smallest eigenvalue is not close to zero. This additional information can also be used to sparsify the flow field.

Although the energy function is equivalent in both cases, the optic flow estimated by the two versions of Lucas-Kanade and Bigün turns out to be different. Until now, there are only speculations to explain this phenomenon. A possible explanation is the following: The linear system proposed by Lucas and Kanade is a least square approach. In the case of the aperture problem it returns the vector with the smallest magnitude from the subspace given by the two eigenvectors which both have eigenvalues close to zero. The method of Bigün, however, returns an arbitrary vector from this subspace. This explanation also goes together with experimental results, where the Lucas-Kanade method yields a flow field, the magnitude of which is often to small in regions where the aperture problem is still relevant, while the results obtained by the method of Bigün et al. show some bad estimates with very high magnitudes in such areas. For this work the linear system proposed by Lucas and Kanade was used as its results are more robust.

However, Bigün et al. made a further extension to the Lucas-Kanade approach. While Lucas and Kanade assumed the optic flow only to be constant in a spatial neighborhood, Bigün et al. extended the neighborhood to the temporal domain. As a result the structure tensor is also smoothed along the time axis. However, there are no further adaptations necessary. Therefore the spatio-temporal extension can also be used together with the linear system of Lucas and Kanade. In general a spatio-temporal approach leads to better estimation results.

4.1.2 Horn-Schunck

While Lucas-Kanade and Bigün made local assumptions to circumvent the aperture problem, the method proposed by Horn and Schunck [HS81] is a *global* approach. The optic flow constraint is embedded in the global energy functional

Horn-Schunck

$$E_{HS}(u, v) = \int_{\Omega} ((f_x u + f_y v + f_z)^2 + \alpha(|\nabla u|^2 + |\nabla v|^2))dxdy$$

that has to be minimized in order to determine the optic flow components $u(x, y)$ and $v(x, y)$. Note that these components are now functions in *x* and *y,* since they are not constant as in the local approach of Lucas-Kanade.

The energy functional consists of a *data term*

$$\int_{\Omega} (f_x u + f_y v + f_z)^2 dxdy$$

reflecting the optic flow constraint and a *smoothness term*

$$\int_{\Omega} (|\nabla u|^2 + |\nabla v|^2) dxdy$$

assuming the flow field to be smooth by penalizing high gradient magnitudes. This assumption again tackles the aperture problem.

The regularization parameter α weighs the importance of the two assumptions against each other. While small α favour the grey-value constance, large α lead to a smoother flow field. Note that for areas where few image structure is available the data term becomes small yielding an automatic strengthening of the smoothness term. Therefore the flow information is transferred from areas with a clear image structure to areas where the image is poorly structured.

Since the flow components are not constant like in the Lucas-Kanade case, their computation is a bit more complicated and leads to differential equations. A theorem from the calculus of variations [CH53,Els61] states that the minimizing functions of the convex energy functional

$$E(u, v) = \int_{\Omega} F(x, y, u, v, u_x, u_y, v_x, v_y)dxdy$$

are necessarily satisfying the Euler-Lagrange equations

$$\partial_x F_{u_x} + \partial_y F_{u_y} - F_u = 0$$
$$\partial_x F_{v_x} + \partial_y F_{v_y} - F_v = 0$$

with the boundary conditions $\partial_n u = 0$ and $\partial_n v = 0$.

Thus the Euler-Lagrange equations for the Horn-Schunck functional are:

$$\Delta u - \frac{1}{\alpha}(f_x^2 u + f_x f_y v + f_x f_z) = 0$$

$$\Delta v - \frac{1}{\alpha}(f_y^2 v + f_x f_y u + f_y f_z) = 0$$

with reflecting boundary conditions and where Δ denotes the Laplace operator $\Delta := \partial_{xx} + \partial_{yy}$. These diffusion equations have a unique solution [Sch91]. Using finite difference approximations they yield the sparse linear system

$$\sum_{j \in N(i)} (u_i - u_j) - \frac{f_x}{\alpha}(f_x u_i + f_y v_i + f_z) = 0 \qquad i = 1 \ldots M$$

$$\sum_{j \in N(i)} (u_i - u_j) - \frac{f_y}{\alpha}(f_x u_i + f_y v_i + f_z) = 0 \qquad i = 1 \ldots M$$

where $N(i)$ denotes the set of neighbors of pixel i, u_i and v_i are the optic flow components in pixel i, and M is the total number of pixels.
As already known from Chapter 2, such a linear system can be solved iteratively by numerical methods like the *Jacobi* method, the *Gauss-Seidel* method or the *SOR* method. They can be shown to converge for arbitrary initial values.

Another possibility is to use the differential equations to obtain a diffusion-reaction system given by

$$\partial_t u = \Delta u - \frac{1}{\alpha}(f_x^2 u + f_x f_y v + f_x f_z)$$

$$\partial_t v = \Delta v - \frac{1}{\alpha}(f_y^2 v + f_x f_y u + f_y f_z)$$

The functions u and v can then be computed by a vector-valued diffusion approach with the explicit scheme

$$u^{k+1} = \frac{u^k + \tau A(u^k, v^k)u^k - \frac{\tau}{\alpha} f_x (f_y v^k + f_z)}{1 + \frac{\tau}{\alpha} f_x^2}$$

$$v^{k+1} = \frac{v^k + \tau A(u^k, v^k)v^k - \frac{\tau}{\alpha} f_y (f_x u^k + f_z)}{1 + \frac{\tau}{\alpha} f_y^2}$$

where k denotes the time index, τ denotes the time step size and A is the diffusion matrix known from Chapter 2 (see especially section 2.2.3). The solution for u and v is obtained for $t \to \infty$. The advantage of such a diffusion approach is the direct applicability of discontinuity preserving diffusion methods. This is interesting in the case of discontinuity preserving variants of the Horn-Schunck method that will be discussed next. The disadvantage of the diffusion approach with the explicit scheme is its slow convergence. The numerical methods solving the linear system converge much faster. Especially the SOR method is a good choice.

Thus for the Horn-Schunck approach, the smoothness term of which comes down to simple homogeneous diffusion, solving the linear system should be favoured. An alternative that has not been tested for this work is given by implicit diffusion schemes also using fast converging methods.

4.1.3 Extensions of the Horn-Schunck approach

The Horn-Schunck method has been extended in two ways. One extension is its formulation in the spatio-temporal domain. The other extension is the usage of discontinuity preserving smoothness terms. While spatio-temporal extensions have not been so popular, the literature knows a lot of discontinuity preserving versions of the Horn-Schunck approach [AELS99, ADK99, BA91, Coh93, HB93, KTB96, Nag83, Nes93, PGPO94, Sch94, SH89, WS01a]. In the following we will concentrate on the framework described in [WS01a] that considers both discontinuity preserving as well as spatio-temporal regularizers.

In [WS01a] Weickert and Schnörr distinguish the regularizers in *image-driven* and *flow-driven* as well as *isotropic* and *anisotropic*. The latter classification is in compliance with the classification of diffusion methods made in Chapter 2. While image-driven methods use the gradient of the image sequence to determine discontinuities, a flow-driven approach determines discontinuities by means of the estimated optic flow field. Since discontinuities in the image sequence do not necessarily induce discontinuities in the optic flow field, especially for strongly textured image sequences, image-driven methods often yield oversegmentations. Therefore flow-driven methods in general perform better than image-driven methods. In this work only the flow-driven approach has been considered.

Nonlinear isotropic regularizers

With an isotropic, flow-driven regularizer the global energy functional looks like

Horn-Schunck with a nonlinear isotropic regularizer

$$E_{HSI}(u, v) = \int_{\Omega} ((f_x u + f_y v + f_z)^2 + \alpha \Psi(|\nabla u|^2 + |\nabla v|^2))\,dx\,dy$$

where $\Psi(s^2)$ is a differentiable and increasing penalizer function that is strictly convex in s. With $\Psi(s^2) = s^2$ this comes down to the conventional Horn-Schunck approach. Weickert and Schnörr described a penalizer function based on Charbonnier [CBFAB94] which is

$$\Psi(s^2) = \varepsilon s^2 + (1-\varepsilon)\lambda^2 \sqrt{1 + \frac{s^2}{\lambda^2}} \qquad 0 < \varepsilon < 1, \quad \lambda > 0.$$

The parameter ε is only required for proving well-posedness and can be fixed to some very small value close to zero. The other parameter λ has already been introduced in Chapter 2 in the term of diffusivity functions. It serves as a contrast parameter determining whether a discontinuity is strong enough to be preserved.

The Euler-Lagrange equations for the above-mentioned energy functional are

$$\text{div}\,(\Psi'(|\nabla u|^2 + |\nabla v|^2)\nabla u) - \frac{1}{\alpha}(f_x^2 u + f_x f_y v + f_x f_z) = 0$$

$$\text{div}\,(\Psi'(|\nabla u|^2 + |\nabla v|^2)\nabla v) - \frac{1}{\alpha}(f_y^2 v + f_x f_y u + f_y f_z) = 0$$

with the diffusivity Ψ' denoting the derivative of Ψ with respect to its argument.
For the above-mentioned penalizer function it is

$$\Psi'(s^2) = \varepsilon + \frac{1-\varepsilon}{\sqrt{1 + s^2/\lambda^2}}.$$

Diffusivity functions have already appeared in Chapter 2. In fact, as already mentioned in the last subsection, the minimum of the energy functional can also be found by a vector-valued diffusion process. While for the conventional Horn-Schunck method such an approach results in homogeneous diffusion, in this case it comes down to isotropic diffusion. However, for the isotropic case it is still possible to solve the linear system using fast converging iterative methods like SOR. As already mentioned above it is recommendable to use them instead of the explicit diffusion scheme, because in experiments they revealed to converge more than 20 times faster.

Nonlinear anisotropic regularizers

Having in mind the classification of diffusion methods mentioned in Chapter 2, one could also think of an anisotropic regularizer. In fact, Weickert and Schnörr introduced in [WS01a] an anisotropic, flow-driven regularizer yielding the energy functional

Horn-Schunck with a nonlinear anisotropic regularizer

$$E_{HSA}(u, v) = \int_\Omega ((f_x u + f_y v + f_z)^2 + \alpha \,\text{tr}\, \Psi(\nabla u \nabla u^T + \nabla v \nabla v^T))\,dx\,dy$$

where $\text{tr}(A)$ denotes the sum of the diagonal elements of a matrix A.
The energy functional can again be minimized in compliance with the Euler-Lagrange equations leading to

$$\text{div}\,(\Psi'(\nabla u \nabla u^T + \nabla v \nabla v^T)\nabla u) - \frac{1}{\alpha}(f_x^2 u + f_x f_y v + f_x f_z) = 0$$

$$\text{div}\,(\Psi'(\nabla u \nabla u^T + \nabla v \nabla v^T)\nabla v) - \frac{1}{\alpha}(f_y^2 v + f_x f_y u + f_y f_z) = 0$$

or the diffusion-reaction system

$$\partial_t u = \text{div}\,(\Psi'(\nabla u \nabla u^T + \nabla v \nabla v^T)\nabla u) - \frac{1}{\alpha}(f_x^2 u + f_x f_y v + f_x f_z)$$

$$\partial_t v = \text{div}\,(\Psi'(\nabla u \nabla u^T + \nabla v \nabla v^T)\nabla v) - \frac{1}{\alpha}(f_y^2 v + f_x f_y u + f_y f_z)$$

Using the diffusion-reaction system, a solution can again be computed by means of a vector-valued diffusion approach, now applying anisotropic diffusion. The linear system obtained by the Euler-Lagrange equations, however, may cause problems, since the iterative methods to solve the linear system do not necessarily converge. A possibility to speed up the computation used for this work is to compute an initial estimation of the flow field with the help of isotro-

pic diffusion and SOR. Afterwards the explicit scheme can be used to get an anisotropic solution. As both methods have very similar results, now there are only few iterations necessary to obtain a solution very close to the minimum.

Spatio-temporal regularizers

Considering the spatio-temporal extension of the conventional Horn-Schunck method introduced by Elad and Feuer [EF98], the flow components u and v become functions in x, y and z. Besides, the Nabla operator has to be understood as a spatio-temporal operator $\nabla_3 = (\partial_x, \partial_y, \partial_z)$. The resulting energy functional is

$$E_{HS3D}(u, v) = \int_{\Omega} ((f_x u + f_y v + f_z)^2 + \alpha(|\nabla_3 u|^2 + |\nabla_3 v|^2))\,dx\,dy\,dz$$

where Ω now includes the time axis. In the Euler-Lagrange equations this leads to

$$\Delta_3 u - \frac{1}{\alpha}(f_x^2 u + f_x f_y v + f_x f_z) = 0$$

$$\Delta_3 v - \frac{1}{\alpha}(f_y^2 v + f_x f_y u + f_y f_z) = 0$$

with the spatio-temporal Laplace operator $\Delta_3 = \partial_{xx} + \partial_{yy} + \partial_{zz}$. In the discretization leading to the linear system this only yields for each pixel two additional neighbors along the time axis. If seeking the minimum by the diffusion-reaction system the extension is just as simple, since the two-dimensional diffusion process simply becomes three-dimensional. The major drawback of the spatio-temporal approach is though the increased amount of memory necessary to hold the complete image sequence as well as the flow components.

In [WS01a] Weickert and Schnörr also described the spatio-temporal versions of their isotropic and anisotropic regularizers. They work the same way as the spatio-temporal version of the basic Horn-Schunck method. A more detailed treatment of especially the spatio-temporal extension and its advantages against the spatial version can be found in [WS01b]. In general the additional regularizing along the time axis yields better results than the pure spatial processing.

4.1.4 CLG

In [WBS01] Weickert et al. introduced CLG (Combined Local-Global) as a combination of the *local* Lucas-Kanade and the *global* Horn-Schunck approach. It can also be seen as an extension of the Horn-Schunck approach using the structure tensor known from the Lucas-Kanade and Bigün method. Rewriting the global energy functional of Horn-Schunck

$$E_{HS}(u, v) = \int_{\Omega} ((f_x u + f_y v + f_z)^2 + \alpha(|\nabla u|^2 + |\nabla v|^2))dxdy$$

the following way

$$E_{HS}(u, v) = \int_{\Omega} (w^T J_0 w + \alpha(|\nabla u|^2 + |\nabla v|^2))dxdy$$

with $w = (u, v, 1)$ and $J_0 = \nabla f \nabla f^T$, the similarities to the quadratic form of the energy function of the Lucas-Kanade method introduced by Bigün et al.

$$E_{LK}(u, v) = w^T J_\rho w$$

with J_ρ denoting the structure tensor, becomes evident. For the CLG method J_0 from the Horn-Schunck method is simply replaced by the structure tensor J_ρ:

CLG

$$E_{CLG}(u, v) = \int_{\Omega} (w^T J_\rho w + \alpha(|\nabla u|^2 + |\nabla v|^2))dxdy.$$

Since only pre-computed data is changed, the process in order to minimize the global energy functional stays the same. The only thing that has to be done is to smooth the components of J_0 before minimizing the energy functional. By extending the structure tensor as well as the smoothness term to the spatio-temporal domain, as shown for Lucas-Kanade and Horn-Schunck, a spatio-temporal version of CLG is obtained.

Although the idea of CLG is quite simple it constitutes an important progress, as both the Lucas-Kanade method and Horn-Schunck method can be seen as special cases of CLG with certain parameter settings. For the Lucas-Kanade technique the parameter α is zero, while for the Horn-Schunck approach the same holds for the parameter ρ. This combination of two existing techniques leads later in this chapter to an even more general formulation where also discontinuity preserving variants can be integrated into the concept.

4.1.5 Characteristics of the mentioned techniques

At the end of this section the characteristics of the mentioned techniques are being discussed. This should only be seen as a brief summary. A more detailed discussion about the performance of the presented techniques including the extensions developed in this work can be found in the next chapter where the results of several experiments with the different methods are presented.

On the previous pages two basic optic flow estimation techniques have been described: first the local method of Lucas-Kanade and Bigün and then the global method of Horn-Schunck. Finally a technique was mentioned that combined both approaches, the CLG method. One rea-

son why such a combined method appeared was the fact that local and global methods have different benefits and drawbacks.

The Lucas-Kanade technique works fine for sequences where mainly translational motion is present. On the other hand it has problems to cope with divergent motion. This is mainly because this local method assumes the flow field to be constant within a local area, and this assumption is violated for divergent motion. The same problem appears with flow discontinuities. Here again the flow field can not be assumed to be constant. However, the Lucas-Kanade method has a useful confidence measure to sort out where the flow estimation is rather reliable and where not. Moreover, it is quite robust against noise and compared to other techniques an estimation can be computed very quickly. Also the technique is more or less simple to implement.

The global approach of Horn and Schunck behaves differently. First it can handle both translational and divergent motion, as the flow field is assumed to be smooth but not constant. Of course, this still creates problems at flow discontinuities, yet for this reason several discontinuity preserving extensions of the Horn-Schunck technique have been developed. Another benefit of a global method is that a dense flow field can be guaranteed. On the other hand, a major drawback of the Horn-Schunck method is its sensitivity to image noise. Maybe the most important problem with a global approach, however, is the implementation aspect. Besides the fact that more computation time is necessary to find a solution, it is not straight-forward to implement a global approach in a correct way. This might be the reason why global methods like Horn-Schunck in the literature often yield suboptimal results.

It is not surprising that a combined method like CLG can not circumvent the last-mentioned problem. However, CLG can combine some other benefits of local and global methods. So the noise sensitivity of Horn-Schunck is not a problem anymore. The same holds for divergent motion that caused problems in local approaches. Furthermore, it can be observed that even in sequences with hardly any noise CLG is sometimes superior to the pure Horn-Schunck technique [Bru01]. Flow discontinuities, however, are not yet addressed by the CLG method. Both the structure tensor as well as the smoothness term are based on linear homogeneous diffusion. This will be changed at the end of this chapter.

4.2 Optic flow estimation with the nonlinear structure tensor

Reminding the intention why considering optic flow estimation we will now come back to the nonlinear structure tensor developed in the last chapter. From all optic flow estimation techniques the Lucas-Kanade method seems to be the best choice to verify the performance of the nonlinear structure tensor in comparison to the linear original. The reason is that in Lucas-Kanade all assumptions made to estimate the flow field are already taken into account by the structure tensor. The only additional processing step is the solution of the linear system which is numerical standard work. So a possible superiority of the nonlinear structure tensor against the linear structure tensor must become obvious when being applied to the method of Lucas-Kanade.

4.2.1 Lucas-Kanade and the nonlinear structure tensor

For the original Lucas-Kanade technique the linear system was

$$\begin{pmatrix} J_{\rho 00} & J_{\rho 01} \\ J_{\rho 01} & J_{\rho 11} \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -J_{\rho 02} \\ -J_{\rho 12} \end{pmatrix}$$

where the $J_{\rho ij}$ are the components of the linear structure tensor J_ρ.
Applying the nonlinear structure tensor instead, the same linear system is obtained

$$\begin{pmatrix} J_{t00} & J_{t01} \\ J_{t01} & J_{t11} \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -J_{t02} \\ -J_{t12} \end{pmatrix}$$

with J_{tij} now denoting the components of the nonlinear structure tensor J_t. As a result a better discontinuity preservation can be expected. This will be tested in the next chapter.

Of course, this new version of the Lucas-Kanade technique can also be extended to the spatio-temporal domain. In this case the diffusion process used to obtain the nonlinear structure tensor must be three-dimensional. As the extension of diffusion methods to arbitrary dimensional data is quite simple, this implies no problem.

4.2.2 CLG and the nonlinear structure tensor

Since also the CLG technique as a combination of the Lucas-Kanade method and Horn-Schunck method uses a structure tensor, the nonlinear structure tensor can also be applied here. In this case the differences between the both versions of the structure tensor may become less obvious than with the purely local approach, because the smoothness term of the CLG method could conceal most of the benefits gained with the nonlinear structure tensor. Nevertheless, it seems interesting whether such an extension could yield better results.

The global energy functional of the original CLG method was

$$E_{CLG}(u, v) = \int_\Omega (w^T J_\rho w + \alpha(|\nabla u|^2 + |\nabla v|^2))dxdy$$

with J_ρ being the linear structure tensor. Replacing it by the nonlinear structure tensor J_t yields

$$E_{CLGA}(u, v) = \int_\Omega (w^T J_t w + \alpha(|\nabla u|^2 + |\nabla v|^2))dxdy.$$

Again this spatial version can be extended to the spatio-temporal domain by using a spatio-temporal structure tensor and a spatio-temporal regularizer.

4.3 A general differential optic flow estimation technique

When Weickert et al. presented their CLG method in [WBS01] they also considered the meaning of smoothing steps with regard to optic flow estimation. They distinguished three smoothing steps for their CLG technique:

1. *Noise scale*. There is a presmoothing of the image sequence in order to remove noise or aliasing effects.
2. *Integration scale*. There is a smoothing step in order to integrate across an area of certain size where a constant flow field is expected. This leads to the structure tensor.
3. *Regularization*. The smoothness term in the global approach can also be considered as a smoothing step that smooths the resulting flow field.

For all three smoothing steps they used homogeneous diffusion (or equivalent methods). On the other hand Weickert and Schnörr introduced in [WS01a] a framework that unifies several global optic flow estimation techniques by regarding them as one method using different diffusion techniques for the regularization. When we focus on flow-driven regularizers there appear three kinds of diffusion:

1. *Homogeneous diffusion*
2. *Isotropic diffusion*
3. *Anisotropic diffusion*

The same possibilities exist for the scalar-valued data at the noise scale. Thus there is only the integration scale missing to complete the framework. With the help of matrix-valued diffusion and the nonlinear structure tensor, however, nonlinear diffusion is available for the integration scale, too.

These reflections yield a general differential optic flow estimation technique, which can be formulated by the following energy functional:

Dicontinuity preserving CLG

$$E(u, v) = \int_{\Omega} (w^T J_t w + \alpha\Theta(\nabla u, \nabla v))dxdy$$

where $w = (u, v, 1)$ and J_t is a structure tensor using an arbitrary diffusion technique with diffusion time t to smooth the initial matrix product $J_0 = \nabla f_\sigma \nabla f_\sigma^T$. Here f_σ denotes the image sequence smoothed by an arbitrary diffusion technique with diffusion time σ. $\Theta(\nabla u, \nabla v)$ is in the isotropic case (including the homogeneous case) $\Psi(|\nabla u|^2 + |\nabla v|^2)$ and in the anisotropic case $\mathrm{tr}\,\Psi(\nabla u \nabla u^T + \nabla v \nabla v^T)$.

For this general technique there are four relevant parameters:

1. The diffusion time σ for the presmoothing of the image sequence
2. The diffusion time t for the structure tensor
3. The regularization parameter α
4. The contrast parameter λ in the case of a non-quadratic regularizer

With α set to zero, one ends up with the conventional or discontinuity preserving Lucas-Kanade/Bigün method depending on the kind of diffusion used for the structure tensor. On the other hand, a diffusion time t of zero for the structure tensor yields a Horn-Schunck based method, the discontinuity preserving capabilities of which depending on the regularizer. All other cases are mixtures between those purely local or global methods and are based on the CLG approach that is also covered by this framework using only linear diffusion techniques. A discussion about which parameter settings and combinations of diffusion methods are useful will follow in the next chapter after some test results have been presented.

In the review of the optic flow estimation techniques used for this framework it could be seen that all techniques can be extended to the spatio-temporal domain. Therefore there also exists a spatio-temporal version of the general method. All smoothing steps have simply to be performed in the spatio-temporal domain.

5 Tests

In this chapter some tests with the previously described optic flow estimation techniques are performed. These tests have three intensions. First, the performance of the nonlinear structure tensor in comparison to its linear counterpart should be determined. Second, the results of the CLG method with discontinuity preserving extensions presented in the last chapter are compared to current results from the literature. Finally, the tests should show the baviour of different combinations of diffusion methods and give a hint on what combinations are appropriate for certain applications.

5.1 Test environment

For testing all estimation techniques mentioned in the last chapter the general optic flow estimation technique was implemented. For each of the three smoothing steps one of the following diffusion methods could be selected: *none*, *homogeneous*, *isotropic*, *anisotropic*.

The first two smoothing steps were computed using scalar-valued diffusion and the linear or nonlinear structure tensor respectively. The nonlinear structure tensor based on isotropic diffusion was ignored in the tests, as it was already shown in Chapter 3 to produce inappropriate results.
As far as there was no regularization term selected (yielding a pure local method) the optic flow was estimated by solving the linear system according to Lucas-Kanade. Otherwise the SOR method was used to solve the linear system resulting from the Euler-Lagrange equations. Here, 150 iterations with an overrelaxation factor ω of 1.95 were performed in the case of a quadratic regularizer. For a non-quadratic, isotropic regularizer the extrapolation factor was reduced to 1.9 in order to avoid stability problems for small contrast parameters. The number of iterations consequently had to be increased to 500 iterations. In the case of an anisotropic regularizer first the SOR method was executed with the respective isotropic regularizer and the same parameters. Afterwards 500 iterations with the explicit scheme for anisotropic diffusion were performed to get the final result. For details about the SOR method see [Bru01], [Yo71] or [PTVF92].

The results were obtained on a standard PC with an Intel Pentium III 450 CPU and Windows 98. The depicted flow fields are colored orientation plots with the direction of the flow field mapped on the hue value and its magnitude mapped on the intensity value of the HSI color

model. The saturation value has always been set to its maximum. With this illustration method details of the flow field are better visible than with the generally used vector plots, as it need not to be subsampled. This is especially important to verify the performance of the methods in respect to their discontinuity preservation properties.

5.1.1 Sequences

In the following the sequences used for testing are described. They are all well-known sequences from the literature and can be obtained from the internet. While one sequence allows only a subjective assessment, for the others the ground truth flow field is available. This allows a quantitative comparison of the performance of different techniques. The sequences have been selected in order to ensure a wide variety of challenging features. In the following description those features are named for each sequence. Due to the great number of different techniques that had to be tested small sequences were preferred.

Hamburg Taxi sequence

Figure 12: Hamburg Taxi sequence, frame 9.

Size:	256, 190, 20
Correct Flow field:	Not available Flow estimated between frame 9 and 10
Origin:	`ftp://csd.uwo.ca/pub/vision` Created at the university of Hamburg
Description:	The Hamburg Taxi sequence is a real world sequence with no correct flow field available. It is of a rather poor quality with a lot of noise. There are four moving objects: A taxi going around the corner, a car in the lower left moving to the right, a van in the lower right moving to the left and a pedestrian in the upper left moving along the pavement. The motion is mainly translational. Besides the low quality and the minimum movement of the pedestrian also the large areas without any motion and the resulting discontinuities are interesting features of this sequence.

Street sequence

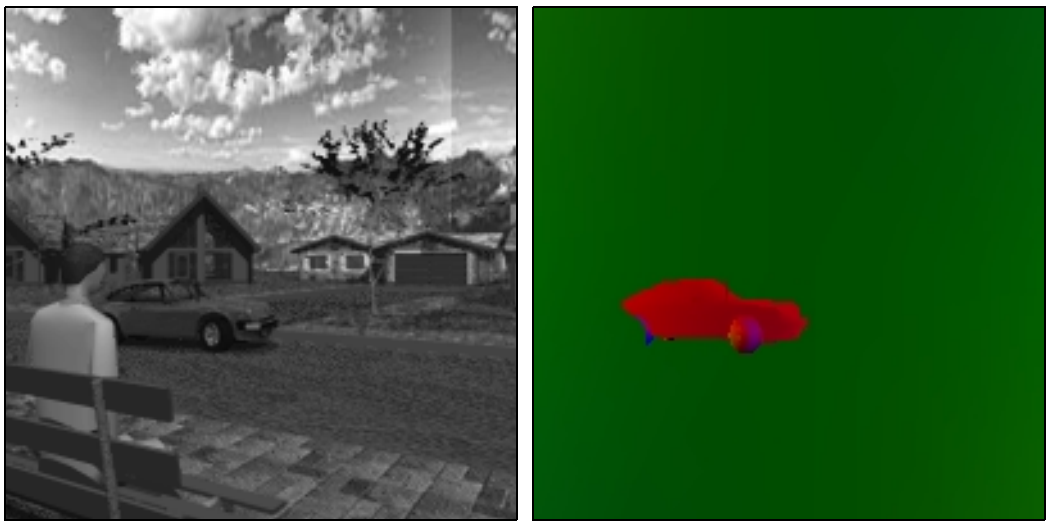

Figure 13: Street sequence.
LEFT: Frame 10.
RIGHT: Correct flow field between frame 10 and 11.

Size:	200, 200, 20
Correct Flow field:	Available between frame 10 and 11
Origin:	`www.cs.otago.ac.nz/research/vision` Created by Galvin et al. [GMN+98]
Description:	The Street sequence is a synthetic sequence with the correct flow field available. There is a car moving from the left to the right. Additionally the camera pans to the right. The motion is completely translational except the wheels of the car which is rotational motion. Challenging features of this sequence are aliasing effects as well as discontinuities and the motion of the wheels.

Office sequence

Figure 14: Office sequence.
LEFT: Frame 10.
RIGHT: Correct flow field between frame 10 and 11.

Size:	200, 200, 20
Correct Flow field:	Available between frame 10 and 11
Origin:	`www.cs.otago.ac.nz/research/vision` Created by Galvin et al. [GMN+98]
Description:	The Office sequence is a synthetic sequence with the correct flow field available. There is no movement apart from the camera moving towards the center of the image. The motion is completely divergent. Challenging features are aliasing effects as well as the window, because outside the window there is hardly any motion. There are further discontinuities at the chair and the desk.

Yosemite sequence with cloudy sky

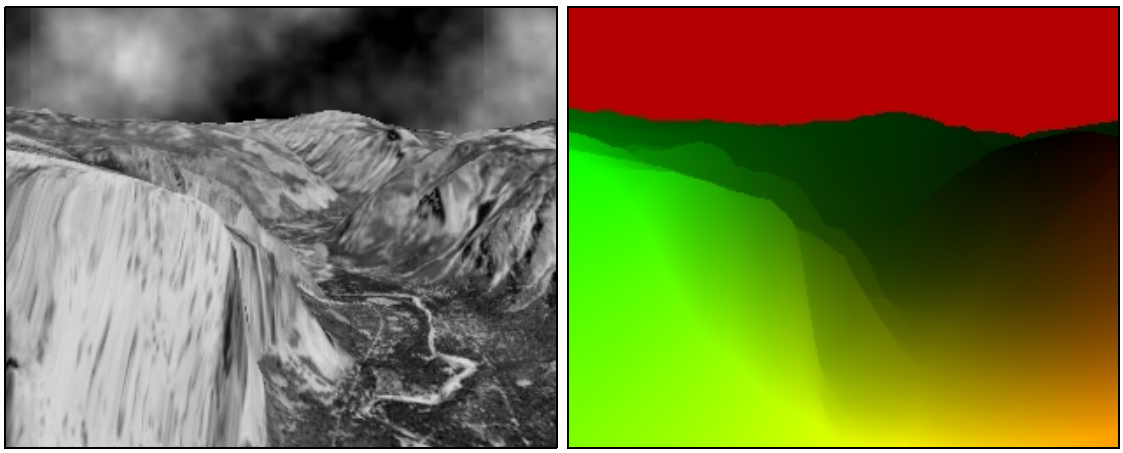

Figure 15: Yosemite sequence with cloudy sky.
LEFT: Frame 8.
RIGHT: Correct flow field between frame 8 and 9.

Size:	316, 252, 15
Correct Flow field:	Available between frame 8 and 9
Origin:	`ftp://ftp.csd.uwo.ca/pub/vision` Created by Lynn Quam at SRI
Description:	The Yosemite sequence with cloudy sky is a synthetic sequence with the correct flow field available. The clouds move to the right and the camera moves towards the horizon. There is both translational and divergent motion in this sequence. This mixture of motion types is one interesting aspect. Other challenging features are large displacements up to 5 pixels/frame in the lower left, discontinuities at the horizon, and the mountains as well as the clouds changing their grey-value.
Note:	The y-components of the correct flow field have opposite sign.

5.1.2 Quality measure

For the sequences with the correct flow field available the *average angular error* (AAE) introduced by Barron et al. [BFB94] was used as a quantitative quality measure. This measure is very popular and therefore enables a more or less fair comparison between the results of the tested techniques and those from the literature.

With (u_c, v_c) being the correct flow field and (u_e, v_e) being the estimated flow the angular error e_i for some pixel i is

$$e_i = \arccos\left(\frac{u_{ci}u_{ei} + v_{ci}v_{ei} + 1}{\sqrt{(u_{ci}^2 + v_{ci}^2 + 1)(u_{ei}^2 + v_{ei}^2 + 1)}}\right).$$

The AAE then is

$$aae = \frac{1}{N}\sum_{i=1}^{N} e_i$$

with N being the number of pixels.

Note that this measure not only penalizes errors in the orientation but also errors in the magnitude. However, it is limited to a maximum of 180°, so some large errors do not have much influence on the AAE. For this reason the standard deviation of the angular error serves as an additional measure pointing out large errors for some pixels. It is defined as

$$\sigma_e = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (e_i - aae)^2}.$$

5.1.3 Parameter optimization

There appear parameters for each of the three smoothing steps. For the first step, the noise scale, these are:

- *Diffusion time* t_{pre}
- *Anisotropy factor* $t_{\text{spatial}}/t_{\text{temporal}}$ (appears only for spatio-temporal smoothing)
- *Contrast parameter* λ_{pre} (appears only in the case of isotropic or anisotropic diffusion)

All those parameters were only optimized once for each sequence.

For the second smoothing step, the structure tensor, the parameters are:

- *Diffusion time* t_{ST}
- *Anisotropy factor* $t_{\text{spatial}}/t_{\text{temporal}}$ (appears only with spatio-temporal techniques)
- *Contrast parameter* λ_{ST} (appears only within the nonlinear structure tensor)
- *Presmoothing parameter* σ_{ST} (appears only within the nonlinear structure tensor)

While the diffusion time was optimized for each sequence and each technique, the other parameters were kept fixed at $t_{\text{spatial}}/t_{\text{temporal}} = 3$, $\lambda_{ST} = 0.1$ and $\sigma_{ST} = 1.58$.

Finally, for the regularization term of the global energy functional there are:

- *Regularization parameter* α
- *Contrast parameter* λ (appears only in the case of a non-quadratic regularizer)

Both parameters were optimized for each sequence and each technique.

So there are three parameters which had to be optimized for each sequence and technique. The main problem is, that those parameters depend on each other. In order to find the optimal parameter setting, combinations of a set of values for each of the three parameters had to be tested. These sets were constructed by first choosing a wide range within which the optimal parameter setting was expected. This range was sampled with a sampling rate that still yielded an acceptable number of test runs. The best combination of parameters was then used to restrict the ranges enabling an increase of the sampling rates. This procedure was performed iteratively until the AAE changed only by 1% or, in the case of a non-existent ground truth flow field, differences in the result were only marginal. For a first idea of what the best parameter setting would look like the work of Bruhn [Bru01] has been very helpful.

5.2 Testing the nonlinear structure tensor

In this section the nonlinear structure tensor is tested with the help of the Lucas-Kanade method. The results of the Lucas-Kanade method using the nonlinear structure tensor are juxtaposed to the original Lucas-Kanade method using the conventional linear structure tensor.

5.2.1 Hamburg Taxi sequence - Preservation of discontinuities in the magnitude

The Hamburg Taxi sequence seems to be a good choice to test the new discontinuity preserving property of the nonlinear structure tensor as there are four moving objects in front of a non-moving background. This yields lots of discontinuities in the magnitude of the structure tensor. Unfortunately, there is no correct flow field available for this sequence, so only a visual assessment of the results is possible. However, since there is no synthetic sequence yet available in the literature with moving objects in front of a non-moving background, the sequence was chosen anyway. Moreover, it is a good test for a method to apply it to a real image sequence.

Figure 16 shows the optic flow fields estimated by the spatial Lucas-Kanade method using the linear structure tensor and the nonlinear structure tensor respectively. Due to artefacts in temporal direction caused by the camera hardware at that time some presmoothing along the time axis with homogeneous diffusion and diffusion time 0.5 has been performed.

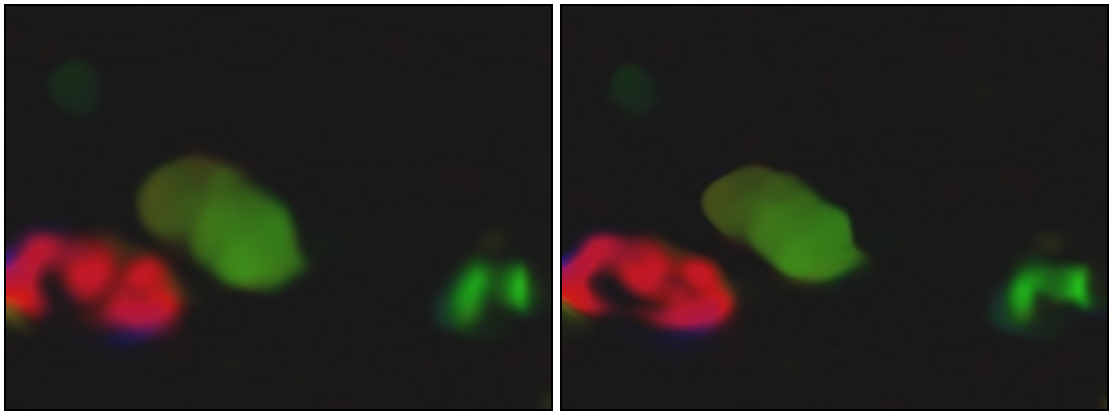

Figure 16: Optic flow field of the Hamburg Taxi sequence
estimated with the spatial Lucas-Kanade method.
LEFT: Linear structure tensor at $t = 10$.
RIGHT: Nonlinear structure tensor at $t = 15$.

Comparing the results achieved with the two different structure tensors the superiority of the nonlinear structure tensor becomes obvious. Discontinuities in the magnitude are much better preserved than in the linear case, and therefore the moving objects do not dislocate that much. This proves that the nonlinear structure tensor can fulfill the expectations and replace the conventional linear structure tensor in a real application in order to improve results.

5.2.2 Street sequence - Preservation of orientation discontinuities

In another test the Street sequence was used to verify the performance of the nonlinear structure tensor in respect to orientation discontinuities. Although the linear structure tensor is also able to preserve orientation discontinuities, as mentioned in chapter 3, it should be tested whether there are still some advantages with the nonlinear structure tensor in the case of optic flow estimation. Moreover, the Street sequence offers the correct flow field enabling the use of a quantitative measure like the average angular error to notice also small differences in quality.
For the spatial Lucas-Kanade method the Street sequence was presmoothed with homogeneous diffusion and a diffusion time of 0.25. For the spatio-temporal variant there has also been applied a spatio-temporal presmoothing with the same diffusion time and an anisotropy factor of 1.

Figure 17 shows the result with the spatial Lucas-Kanade method. There is not that much difference between the linear and the nonlinear structure tensor visible than in the case of the Hamburg Taxi sequence. This is because orientation discontinuities are preserved in both cases. The problem of an area around the discontinuity where the flow is underestimated also remains in both cases. However, Table 1 reveals that the method using the nonlinear structure tensor still performs better than the original Lucas-Kanade method as the average angular error is reduced by 7%. The same holds for the spatio-temporal versions depicted in Figure 18, though the difference here is only 3%.

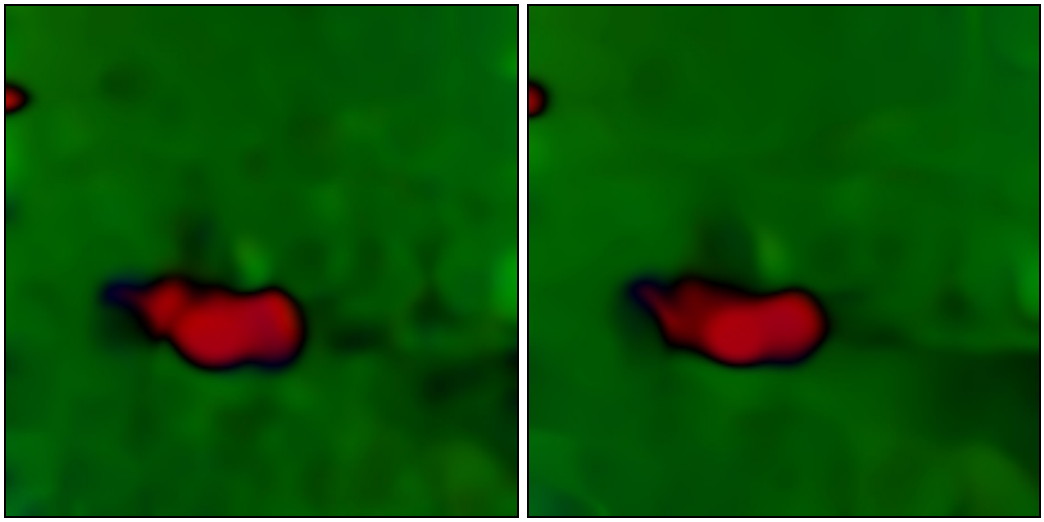

Figure 17: Optic flow of the Street sequence
estimated with the *spatial* Lucas-Kanade method.
LEFT: Linear structure tensor.
RIGHT: Nonlinear structure tensor.

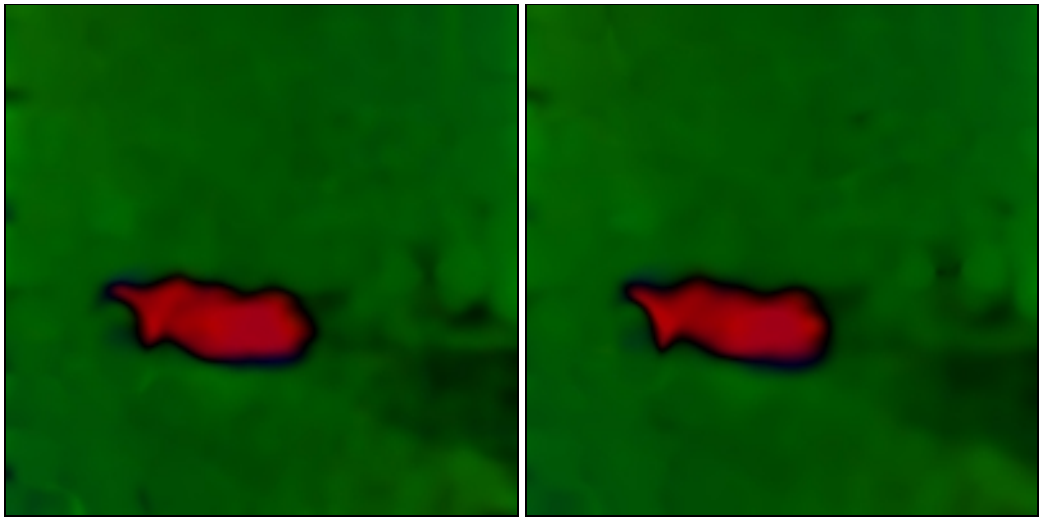

Figure 18: Optic flow of the Street sequence
estimated with the *spatio-temporal* Lucas-Kanade method.
LEFT: Linear structure tensor.
RIGHT: Nonlinear structure tensor.

Table 1: AAE of the Street sequence

method	diffusion time	AAE	standard deviation
Lucas-Kanade 2D linear	10	6.29°	± 13.55°
Lucas-Kanade 2D nonlinear	34	5.88°	± 13.18°
Lucas-Kanade 3D linear	5	5.28°	± 13.41°
Lucas-Kanade 3D nonlinear	7	5.14°	± 13.03°

5.2.3 Yosemite sequence - Overall performance

Finally, some tests with the Yosemite sequence have been performed. This sequence consists of a combination of both translational and divergent motion so it is not optimal for the Lucas-Kanade technique. However, it has a lot of challenging features as well as the correct flow field available and so makes a good sequence for a conclusion of this test series. The sequence was presmoothed spatially by homogeneous diffusion and a diffusion time of 0.75. For the spatio-temporal version an additional presmoothing along the time axis with a diffusion time of 0.25 has been performed.
Figure 19 and Figure 20 show the optimal results of the spatial and spatio-temporal Lucas-Kanade method respectively. The average angular errors are listed in Table 2. There the nonlinear structure tensor can be seen to result in a 9% better AAE in the spatial case. For the spatio-temporal case the lead again reduces to 3% but is still relevant.

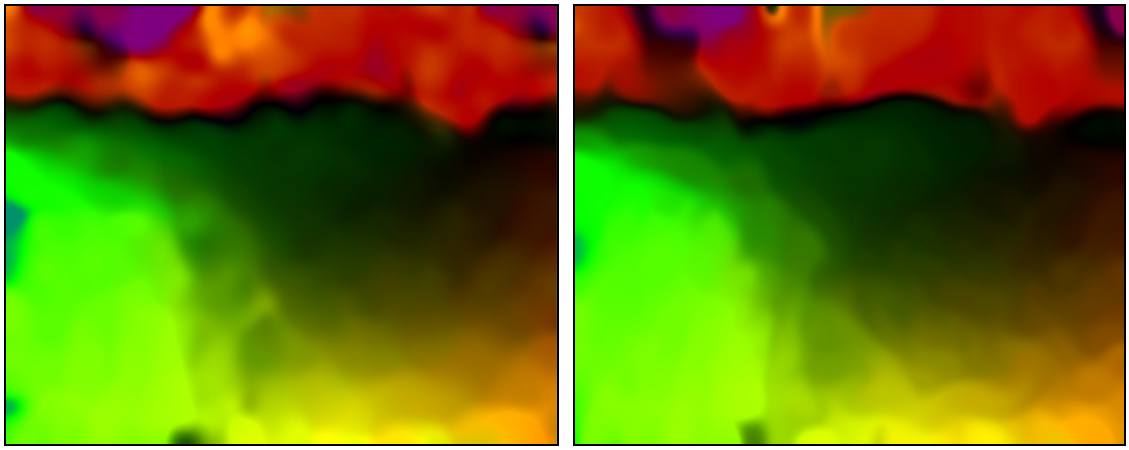

Figure 19: Optic flow of the Yosemite sequence
estimated with the *spatial* Lucas-Kanade method.
LEFT: Linear structure tensor.
RIGHT: Nonlinear structure tensor.

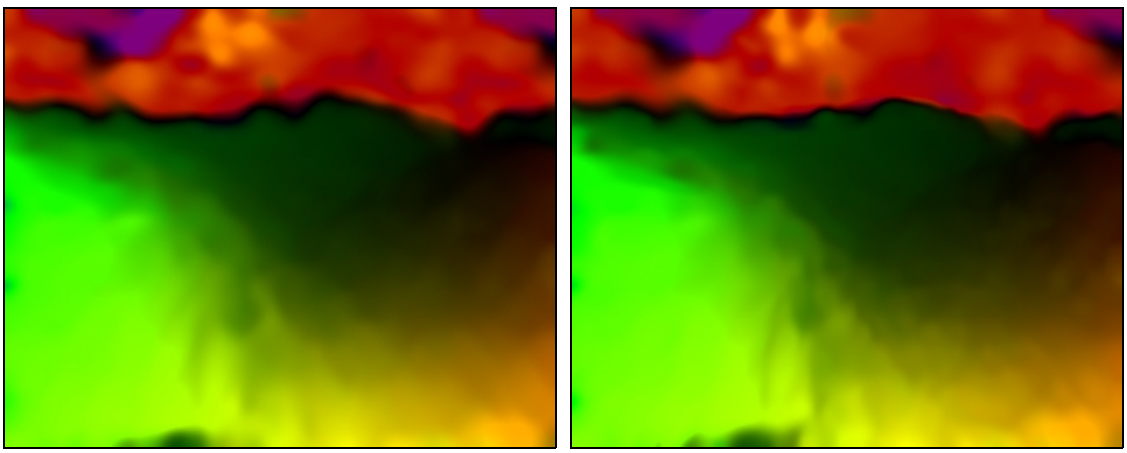

Figure 20: Optic flow of the Yosemite sequence
estimated with the *spatio-temporal* Lucas-Kanade method.
LEFT: Linear structure tensor.
RIGHT: Nonlinear structure tensor.

Table 2: AAE of the Yosemite sequence

method	diffusion time	AAE	standard deviation
Lucas-Kanade 2D linear	21	8.94°	± 13.14°
Lucas-Kanade 2D nonlinear	135	8.22°	± 12.59°
Lucas-Kanade 3D linear	13	7.69°	± 12.58°
Lucas-Kanade 3D nonlinear	18	7.49°	± 12.53°
Lucas-Kanade 2D linear, 55% density	15	6.43°	± 13.34°
Lucas-Kanade 2D nonlinear, 55% density	40	6.06°	± 13.06°

It has also been tested whether the quality of the confidence measure mentioned for the Lucas-Kanade method in the last chapter suffers from the nonlinear structure tensor. The last two entries in Table 2 reveal that the method using the nonlinear structure tensor is still superior to the original method for non-dense flow fields. Of course, the difference in quality becomes smaller because the critical areas where the nonlinear structure tensor performs better than the linear one, especially the discontinuity at the horizon, are no longer part of the measurement.

Table 3: AAE of the Yosemite sequence at different noise levels.
σ_n denotes the standard deviation of the noise.
Spatial version, 100% density.

σ_n	Linear		Nonlinear	
	t_{ST}	AAE	t_{ST}	AAE
0	21	8.94°	135	8.22°
5	30	9.61°	164	8.93°
10	54	11.42°	229	11.19°
20	133	15.38°	543	16.56°
40	180	23.07°	717	24.59°

Since the classic Lucas-Kanade method using the linear structure tensor is known to be quite robust against noise, the Yosemite sequence has been degraded by Gaussian noise to test whether the nonlinear structure tensor yields any drawbacks in this respect. For this test the presmoothing has not been changed compared to the sequence without noise in order to observe the noise robustness of the structure tensor. No confidence measure has been applied yielding a flow field density of 100%. Table 3 demonstrates the Lucas-Kanade method using the nonlinear structure tensor to be still robust against noise and that it can keep its lead for rather high noise levels. For very severe noise, however, it looses its superiority against the linear structure tensor and even performs worse. This is because for such high noise levels texture information and noise can not be distinguished anymore. While the nonlinear structure tensor misinterprets some noise as additional structure, its linear counterpart is more conservative, as it does not try to find information in the data but simply removes both noise and structure. Table 3 reveals that for high noise levels such a conservative behaviour gets beneficial. For usual noise levels, however, this strategy neglects useful information. In a test at the end of this chapter it will also be demonstrated that severe noise can be handled quite well by an appropriate presmoothing of the sequence.

5.3 Results with the general optic flow estimation technique

After the usefulness of the nonlinear structure tensor has been verified in the last section, now the performance of the general technique is regarded. There are many combinations of diffusion methods for the three stages. It is expected that the most complex combination, where for all stages anisotropic diffusion is used, will supply the best results. However, such a combination would also induce the highest computational costs. Thus the more simple combinations have also been tested.

However, there is another reason why also the more simple combinations have been tested. A series of different combinations nicely reveals the benefits of replacing a simple diffusion method like homogeneous diffusion by a more complex one. In particular, it becomes obvious of how large the gain of quality is. Moreover, it can be seen which challenges of a sequence can be addressed by a certain increase in complexity. This promotes the understanding of what a certain diffusion method really does at the corresponding smoothing stage and what kind of problems still appear with the general method because of incorrect assumptions or other reasons.

For the tests in this section only sequences with the correct flow field available have been used. This facilitates the comparison of the different combinations of diffusion methods, especially since some improvements might be so minimal that only a quantitative measure can make them visible.

5.3.1 Street sequence

We begin with the synthetic Street sequence. Like in the tests before homogeneous diffusion with a diffusion time of 0.25 has been used for presmoothing. Other diffusion methods for this first smoothing step will only be discussed for the Yosemite sequence. There also some noise was added to the sequence. As will be seen then it makes no sense to use a more complex diffusion method for presmoothing as long as the sequence is not degraded by noise.

All results are listed in Table 5 by a descending average angular error. There it becomes visible that for the Street sequence the classic Horn-Schunck method with a linear structure tensor supplies the worst result. Even the simpler linear Lucas-Kanade method performs better. The original CLG technique can improve the result of Lucas-Kanade only marginally. This is because the Street sequence with its merely translational motion fits perfectly well to the assumption of a locally constant flow field. On the other hand, the smoothness assumption of the Horn-Schunck approach is violated due to the discontinuity of the car. Regarding the parameter settings of the original CLG reveals that in fact it works mainly as local method in this case, as the regularization parameter is rather small. So the smoothing of the structure tensor is very useful for this sequence. However, the standard deviation can also be observed to rise in the case of a smoothed structure tensor. This indicates some essential errors made by smoothing the structure tensor. Using the nonlinear structure tensor improves the result. However, the standard deviation, even though it gets better in this case, is still high. On the other hand, it can be seen that it is much smaller for the Horn-Schunck method despite its poor

AAE. A CLG approach can benefit from both the better AAE of the local method and the lower standard deviation of the global method.
Using a discontinuity preserving regularizer improves the AAE as well as the standard deviation significantly. This is because the main problem of the Horn-Schunck method for this sequence is addressed in this case. The additional usage of the nonlinear structure tensor can further improve the result a little. The linear structure tensor, however, has not a positive effect anymore. This again proves the superiority of the nonlinear structure tensor. Some further improvement is obtained by replacing the isotropic regularizer by an anisotropic one. Also interesting, of course, is the step from the spatial to the spatio-temporal domain. Here some further improvements of almost 10% can be obtained.
Summarizing the results for this sequence almost every additional step towards a higher complexity also yields better results.

Table 4: AAE for the Street sequence

method (ST, regularizer, dimension)	t_{ST}	α	λ	AAE	standard deviation
None, Linear, 2D	-	500	-	6.50°	± 11.32°
Linear, None, 2D	10	-	-	6.29°	± 13.55°
Linear, Linear, 2D	7	50	-	6.27°	± 12.89°
Nonlinear, None, 2D	34	-	-	5.88°	± 13.18°
Nonlinear, Linear, 2D	45	5	-	5.67°	± 12.23°
None, Isotropic, 2D	-	23000	0.0025	3.82°	± 8.59°
Linear, Isotropic, 2D	0	23000	0.0025	3.82°	± 8.59°
Nonlinear, Isotropic, 2D	1	22000	0.0025	3.73°	± 8.82°
Nonlinear, Anisotropic, 2D	1	21000	0.0025	3.65°	± 8.83°
None, Linear, 3D	-	60	-	6.02°	± 11.97°
Linear, None, 3D	5	-	-	5.28°	± 13.41°
Nonlinear, None, 3D	7	-	-	5.14°	± 13.03°
None, Isotropic, 3D	-	9000	0.0025	3.52°	± 9.33°
Nonlinear, Isotropic, 3D	0	9000	0.0025	3.52°	± 9.33°
Nonlinear, Anisotropic, 3D	0	9000	0.0025	3.31°	± 9.40°

The estimated flow fields are depicted in Figure 21 and Figure 23 for the spatial and spatio-temporal methods respectively. Figure 22 and Figure 24 show the angular errors for each pixel. Here the problems of the respective methods become obvious. It is visible that the aliasing effects in the trees cause many methods to estimate a too large flow there. This need not be a drawback in every application. A human observer also detects some increased motion in this area. However, most applications want these aliasing effects to be ignored. This is best done by smoothing the structure tensor.
However, it gets also visible that, on the other hand, the smoothness term is very important for a more reliable estimation of the flow. Whenever the local part of the technique dominates the estimation process there are large errors in the frontal area of the car. This also explains the higher standard deviations for those methods.

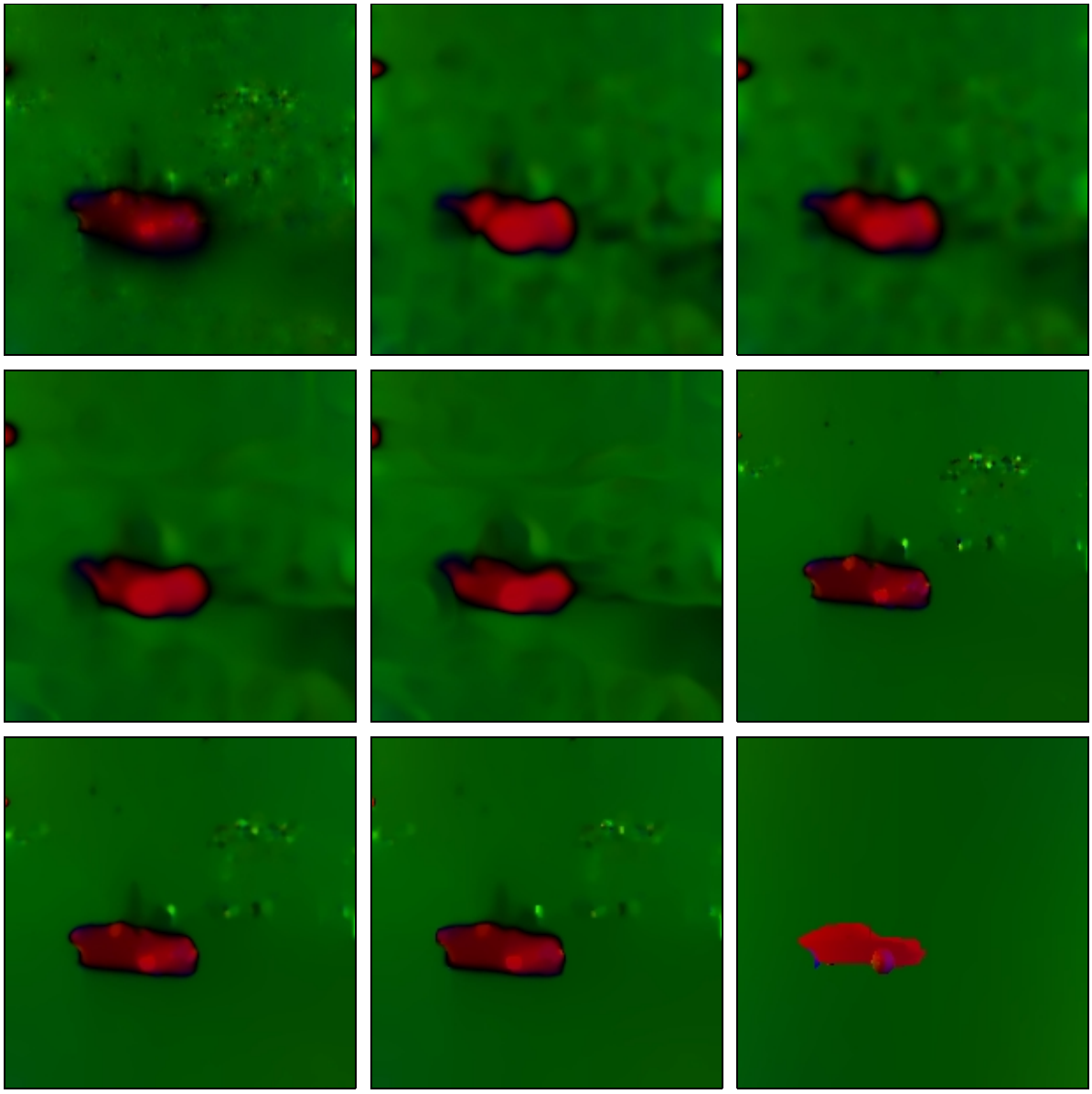

Figure 21: Estimated flow fields of the Street sequence, spatial versions.
From Left to Right, Top to Bottom:
(a) None, Linear. (b) Linear, None. (c) Linear, Linear.
(d) Nonlinear, None. (e) Nonlinear, Linear. (f) None, Isotropic.
(g) Nonlinear, Isotropic. (h) Nonlinear, Anisotropic. (i) Correct flow field.

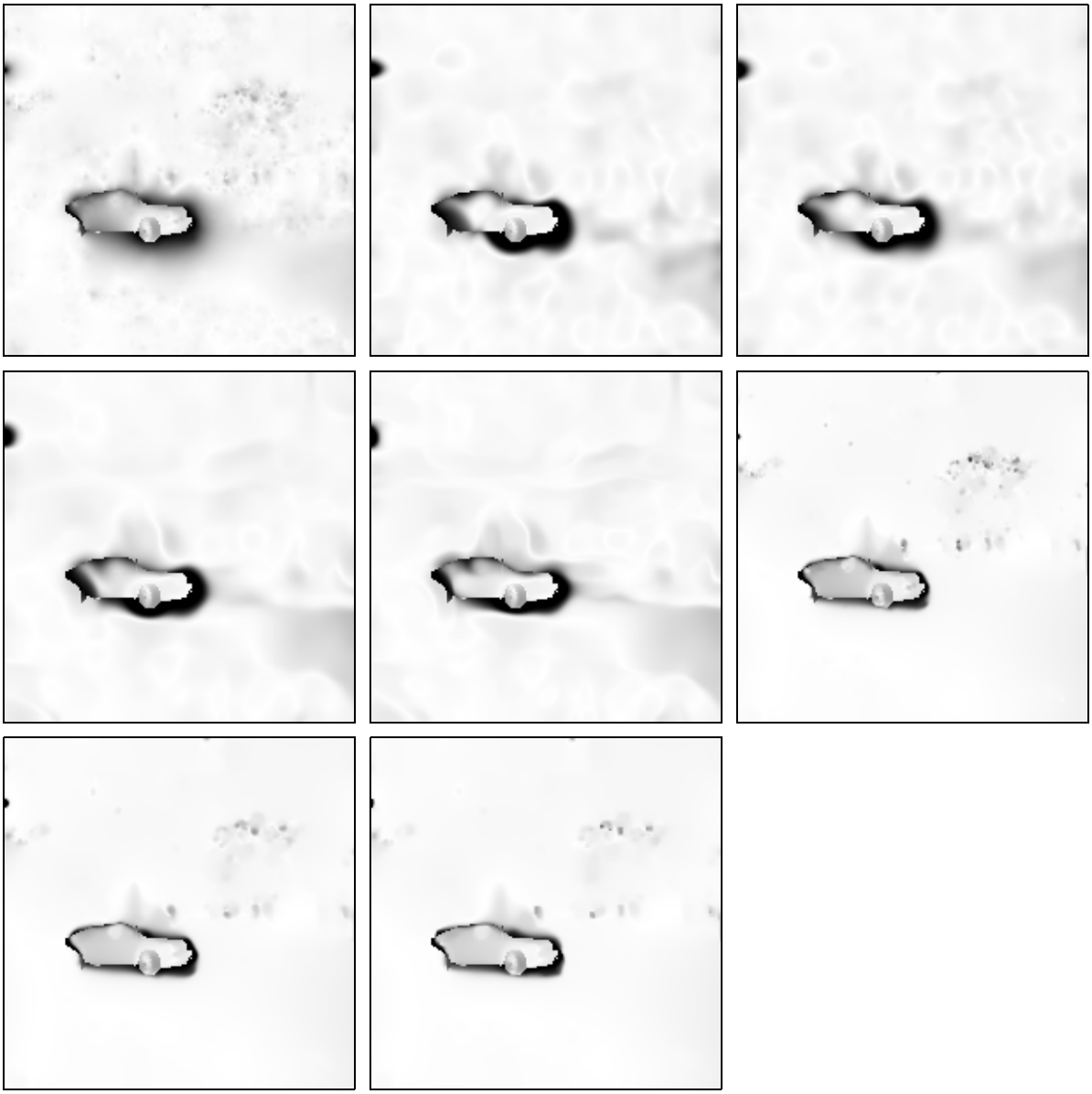

Figure 22: Street sequence. Angular error of the estimated flow fields, spatial versions.
FROM LEFT TO RIGHT, TOP TO BOTTOM:
(a) None, Linear. (b) Linear, None. (c) Linear, Linear.
(d) Nonlinear, None. (e) Nonlinear, Linear. (f) None, Isotropic.
(g) Nonlinear, Isotropic. (h) Nonlinear, Anisotropic.

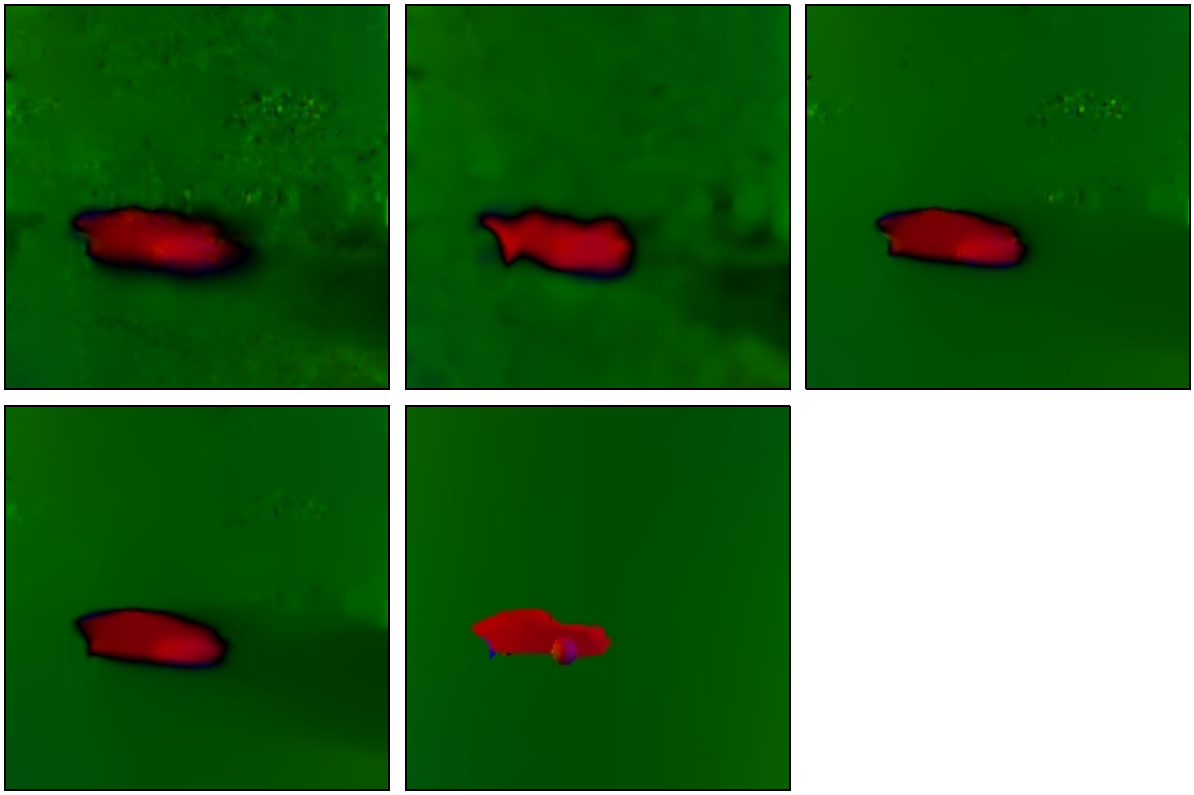

Figure 23: Estimated flow fields of the Street sequence, spatio-temporal versions.
FROM LEFT TO RIGHT, TOP TO BOTTOM:
(a) None, Linear. (b) Nonlinear, None. (c) None, Isotropic.
(d) Nonlinear, Anisotropic. (e) Correct flow field.

In Table 5 some of the results are compared to results from the literature. For the Street sequence there where results available from Galvin et al. [GMN+98] as well as Weickert and Schnörr [WS01b]. Comparing the results reveals the importance of implementation aspects, since some of the methods in the literature should be equivalent to some of ours, but supply worse results. So the result of our implementation of the Horn-Schunck method has a better AAE with a significantly higher density than that mentioned in [GMN+98]. The differences between the results of Weickert and Schnörr and ours (None, Isotropic, 2D/3D) can be explained by the fact that they used the explicit diffusion scheme instead of the SOR method. It seems that the explicit scheme requires more iterations to reach its steady state.

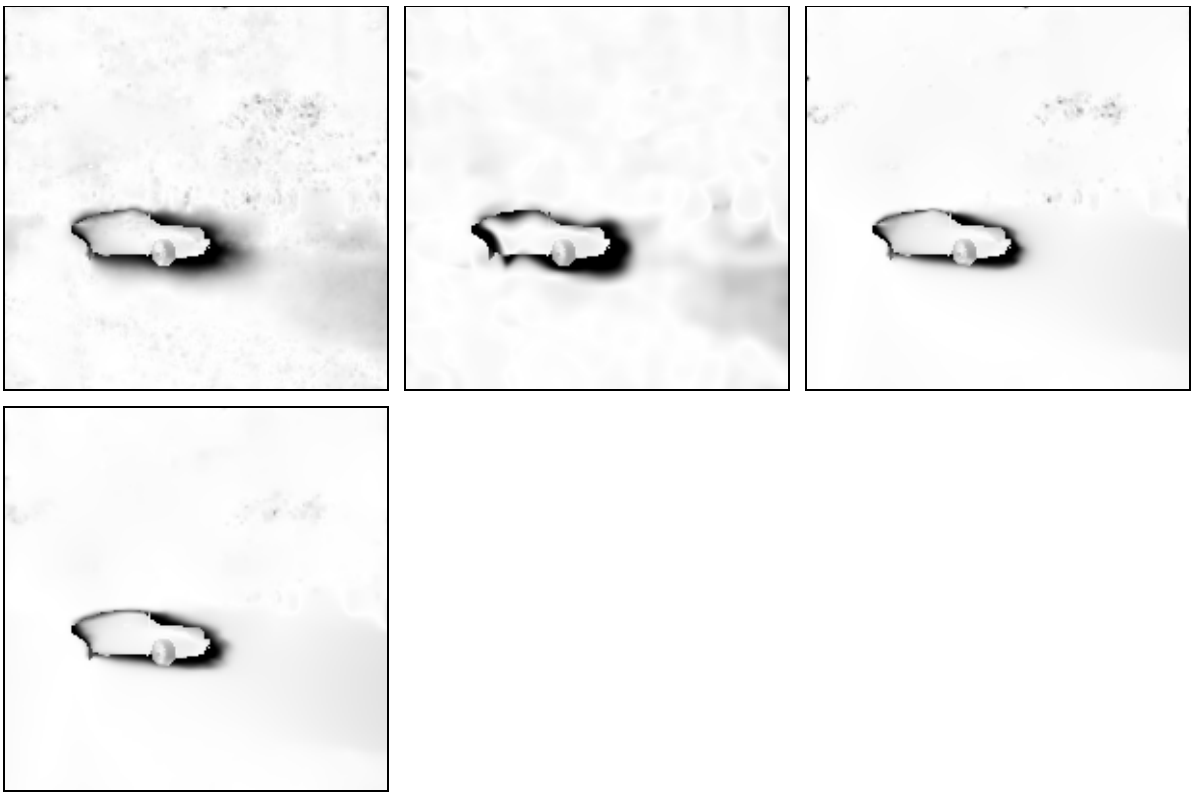

Figure 24: Angular error of the estimated flow fields, spatio-temporal versions.
FROM LEFT TO RIGHT, TOP TO BOTTOM:
(a) None, Linear. (b) Nonlinear, None. (c) None, Isotropic.
(d) Nonlinear, Anisotropic.

Table 5: Street sequence. Our results compared to some in the literature.

method	AAE	density
Camus [GMN+98]	13.69°	100%
Proesman et al. [GMN+98]	7.41°	100%
Weickert-Schnörr 2D [WS01b]	6.62°	100%
None, Linear, 2D (Horn-Schunck)	**6.50°**	**100%**
Weickert-Schnörr 3D [WS01b]	4.85°	100%
None, Isotropic, 2D (Weickert-Schnörr 2D)	**3.82°**	**100%**
Nonlinear, Anisotropic, 2D	**3.65°**	**100%**
None, Isotropic, 3D (Weickert-Schnörr 3D)	**3.52°**	**100%**
Nonlinear, Anisotropic, 3D	**3.31°**	**100%**
Anandan [GMN+98]	10.58°	54%
Horn-Schunck [GMN+98]	6.62°	46%
Singh [GMN+98]	6.18°	78%
Linear, None, 2D (Lucas-Kanade)	**4.82°**	**52%**
Nonlinear, None, 2D	**4.51°**	**53%**
Linear, None, 3D (Bigün)	**3.49°**	**58%**
Nonlinear, None, 3D	**3.30°**	**57%**

5.3.2 Office sequence

The next test sequence is the Office sequence. It consists of totally divergent motion with some discontinuities due to near and distant objects. For presmoothing the same settings as for the Street sequence have been used: Homogeneous diffusion with a diffusion time of 0.25. For the spatio-temporal methods this presmoothing has also be done along the time axis.

Table 6 shows the results for the different methods. As expected, the purely local methods give rather bad results, because their assumption of constant flow is violated for divergent motion. Moreover, for the Lucas-Kanade method with the nonlinear structure tensor there is an area where the flow can not be estimated at all (see Figure 25b), as the linear system gets singular due to missing gradients in this region.
The global methods perform much better. However, the local philosophy can improve those results. For the conventional CLG technique with a linear structure tensor this improvement is rather small. With the nonlinear structure tensor, however, the result gets 7% better compared to the classic Horn-Schunck method.
Non-quadratic, discontinuity preserving regularizers, on the other hand, have no effect on the results. This is surprising especially because there are lots of discontinuities in this sequence. However, those discontinuities are very weak. A much smaller contrast parameter would have to be used in order to preserve them, yet this has negative effects because of the aliasing in the sequence. The positive effect of preserved discontinuities is therefore compensated by the negative effects due to the aliasing and in fact leads, because of the large contrast parameter in the optimal parameter setting, to a quadratic regularizer.
A spatio-temporal approach, however, improves the results considerably. Gains up to 30% can be achieved here.

Table 6: AAE for the Office sequence

method (ST, regularizer, dimension)	t_{ST}	α	λ	AAE	standard deviation
Linear, None, 2D	134	-	-	5.72°	± 5.17°
Nonlinear, None, 2D	328	-	-	5.19°	± 5.00°
None, Linear, 2D	-	2500	-	4.37°	± 4.29°
None, Isotropic, 2D	-	2500	3	4.37°	± 4.29°
Linear, Linear, 2D	6	2800	-	4.34°	± 4.40°
Linear, Isotropic, 2D	6	2800	3	4.34°	± 4.39°
Nonlinear, Linear, 2D	50	1800	-	4.08°	± 4.19°
Nonlinear, Isotropic, 2D	50	1800	3	4.08°	± 4.19°
None, Linear, 3D	-	350	-	3.37°	± 3.82°
None, Isotropic, 3D	-	350	3	3.37°	± 3.82°
Nonlinear, Linear, 3D	1	450	-	3.34°	± 3.96°

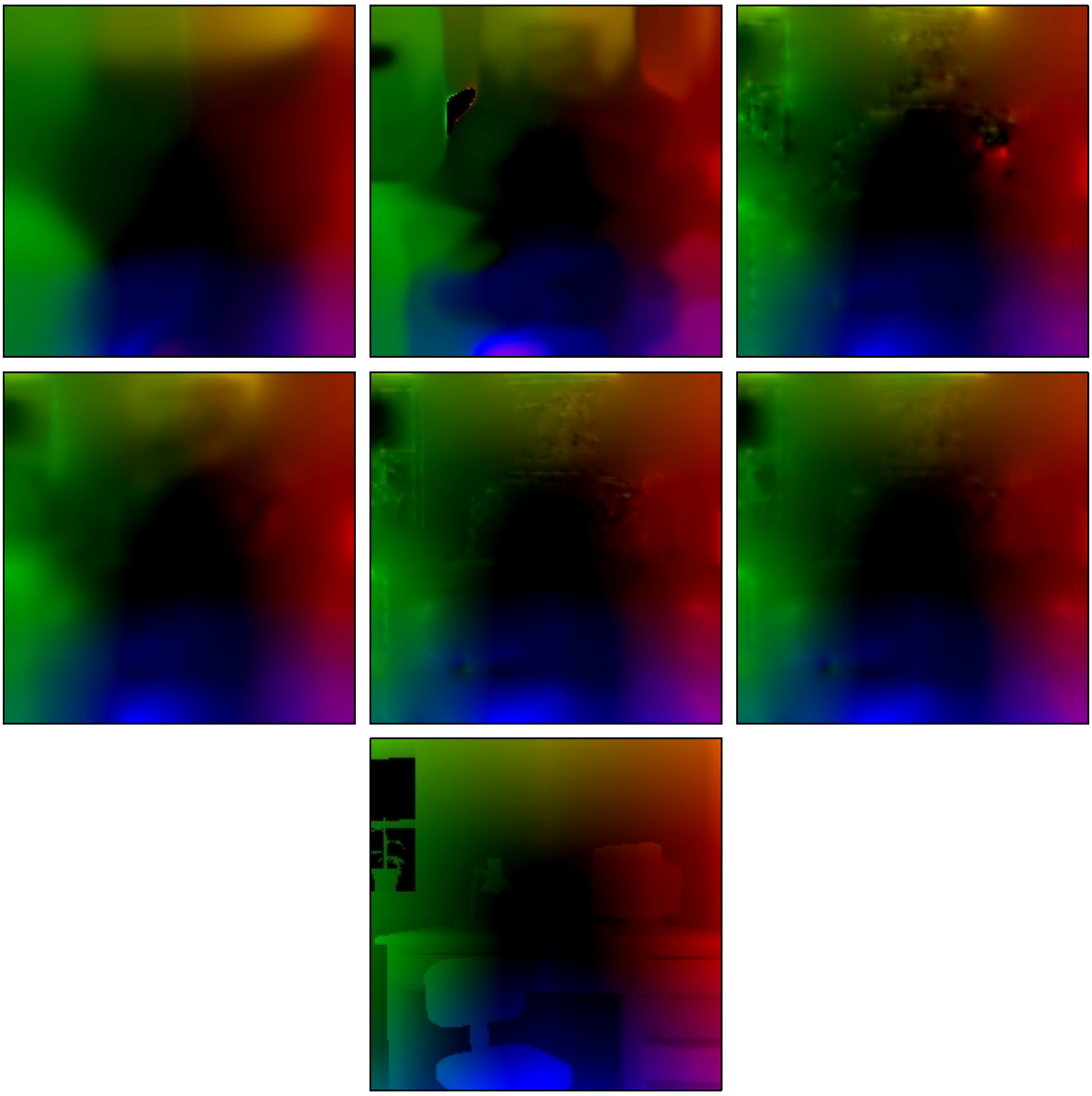

Figure 25: Estimated flow fields of the Office sequence.
FROM LEFT TO RIGHT, TOP TO BOTTOM:
(a) Linear, None, 2D. (b) Nonlinear, None, 2D. (c) None, Linear, 2D.
(d) Nonlinear, Linear, 2D. (e) None, Linear, 3D. (f) Nonlinear, Linear, 3D.
(g) Correct flow field.

In Figure 25 the estimated flow fields are depicted. They show that the nonlinear structure tensor can deal with the aliasing effects to some degree but most of the work has to be done by the global part of the method. Spatio-temporal versions yield more accurate results. This is beneficial especially for this sequence with its fine structures within the smoothly changing divergent flow field.
Unfortunately, there are no accurate numbers for the Office sequence available in the literature. Therefore a comparison like for the other sequences is not possible here.

5.3.3 Yosemite sequence

The last test sequence in this series is again the popular Yosemite sequence. It consists of mostly divergent motion in the canyon area and translational motion in the clouds region. However, for the clouds region the assumption of grey-value constancy is not fulfilled. Furthermore, in the lower left there are large displacements of up to 5 pixels/frame. This can be a problem for differential estimation methods. The motion discontinuity at the horizon is also a feature that causes problems, especially because the motion of the clouds alone is already difficult to estimate. Therefore the Yosemite sequence with cloudy sky is one of the most challenging test sequences in the literature and in general yields the worst average angular errors from all sequences.

As Table 7 reveals, this is also true for our methods. The results have all been obtained with the presmoothed sequence using homogeneous diffusion and a diffusion time of 0.75 in spatial direction. For the spatio-temporal methods an additional smoothing along the time axis with a diffusion time of 0.25 has been performed.

Table 7: AAE for the Yosemite sequence without noise and different combinations of diffusion methods.

method (ST, regularizer, dimension)	t_{ST}	α	λ	AAE	standard deviation
Linear, None, 2D	21	-	-	8.94°	± 13.14°
Nonlinear, None, 2D	135	-	-	8.22°	± 12.59°
None, Linear, 2D	-	450	-	7.20°	± 9.01°
Linear, Linear, 2D	0	450	-	7.20°	± 9.01°
Nonlinear, Linear, 2D	14	360	-	7.07°	± 9.48°
None, Isotropic, 2D	-	2000	0.02	6.39°	± 8.30°
Linear, Isotropic, 2D	0	2000	0.02	6.39°	± 8.30°
Nonlinear, Isotropic, 2D	0	2000	0.02	6.39°	± 8.30°
None, Anisotropic, 2D	-	2000	0.02	6.33°	± 8.28°
Nonlinear, Anisotropic, 2D	0	2000	0.02	6.33°	± 8.28°
Linear, None, 3D	13	-	-	7.69°	± 12.58°
Nonlinear, None, 3D	18	-	-	7.49°	± 12.53°
None, Linear, 3D	-	190	-	6.31°	± 8.87°
None, Anisotropic, 3D	-	2100	0.01	6.11°	± 8.85°
None, Isotropic, 3D	-	2500	0.01	5.76°	± 8.77°
Nonlinear, Isotropic, 3D	0	2500	0.01	5.76°	± 8.77°

Of course, the purely local methods perform worst, since their assumption of locally constant flow is violated for divergent motion. Furthermore, Figure 26 reveals them also to have more problems with the discontinuity at the horizon than the global methods. Although the nonlinear structure tensor can handle this problem better than the linear one, the estimation results for the canyon still reach too far into the sky.

Regarding Table 7 reveals the smoothing of the structure tensor in a global approach to yield only marginal effects, and merely in the case of the nonlinear structure tensor combined with a quadratic regularizer there is any improvement at all. With a non-quadratic regularizer or the linear structure tensor the smoothing step is useless for this sequence. The non-quadratic regularizer, on the other hand, causes improvements of approximately 10% compared to the quadratic regularizer. The same holds for a spatio-temporal approach when compared to the corresponding spatial method. Replacing the isotropic regularizer by an anisotropic one, however, has only little effects.
With the spatio-temporal approach the anisotropic regularizer performs even worse. However, it must be noted that, due to its large computation time, for the optimization of the anisotropic case the optimum parameters from the isotropic approach were taken and only slightly changed. Perhaps a better result could have been achieved with some completely different parameter setting.

Because of its popularity there are quite many results for this sequence available in the literature. In Table 8 only those results with 100% density have been chosen in order to guarantee a fair comparison. It can be seen that the tested methods perform rather well compared to those in the literature. Only Alvarez et al. as well as Mémin and Pérez using multiscale approaches perform better.

Table 8: Yosemite sequence with cloudy sky. Our results compared to those in the literature.

method	AAE
Anandan [BFB94]	13.36°
Singh [BFB94]	10.44°
Nagel [BFB94]	10.22°
Horn-Schunck [BFB94]	9.78°
Uras et al. [BFB94]	8.94°
CLG 2D [WBS01]	7.14°
None, Anisotropic, 2D	**6.33°**
CLG 3D [WBS01]	6.18°
None, Isotropic, 3D	**5.76°**
Alvarez et al. [AELS99]	5.53°
Mémin-Pérez [MP98]	5.38°

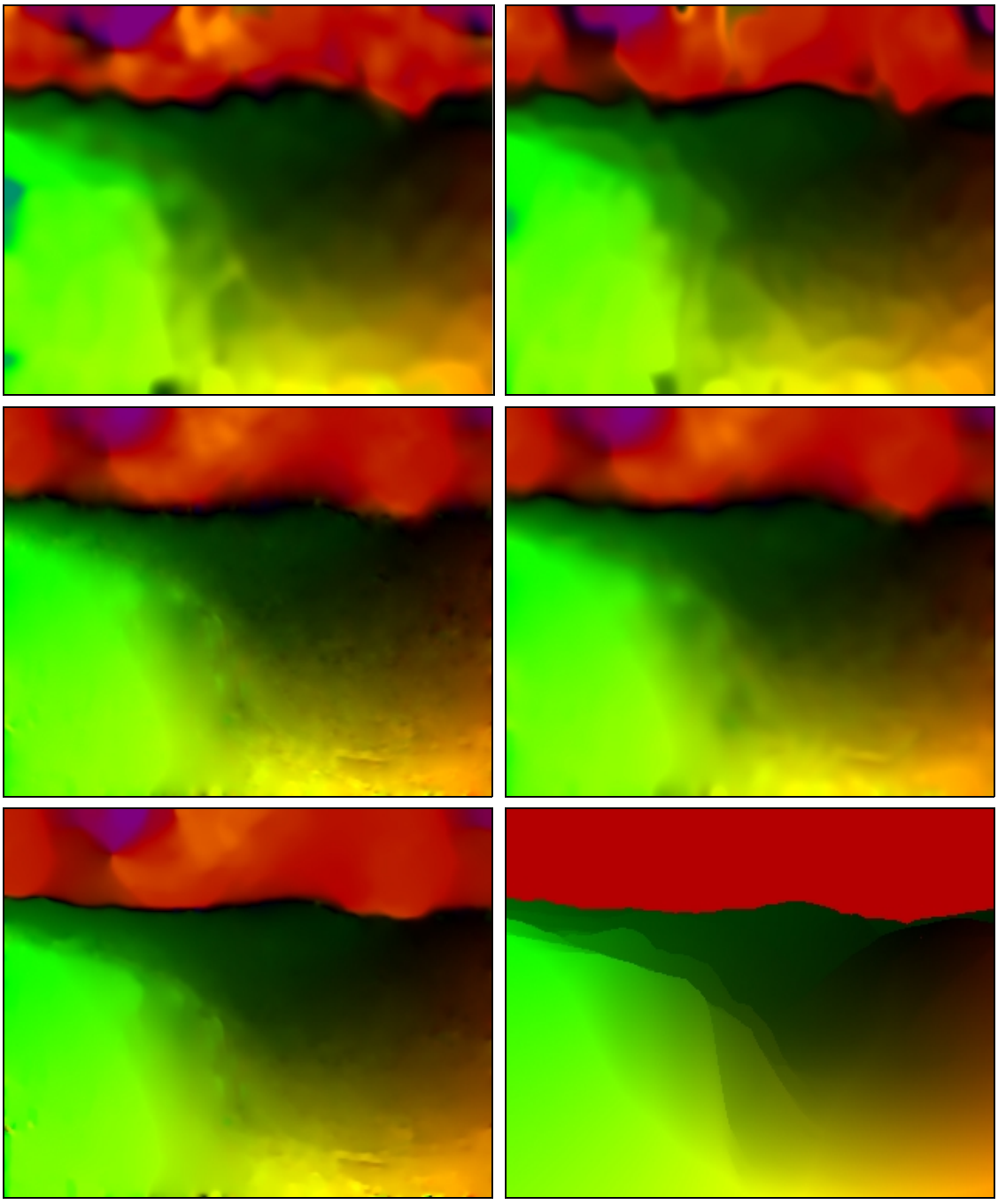

Figure 26: Estimated flow fields of the Yosemite sequence, spatial versions.
FROM LEFT TO RIGHT, TOP TO BOTTOM:
(a) Linear, None. (b) Nonlinear, None.
(c) None, Linear. (d) Nonlinear, Linear.
(e) None, Anisotropic. (f) Correct flow field.

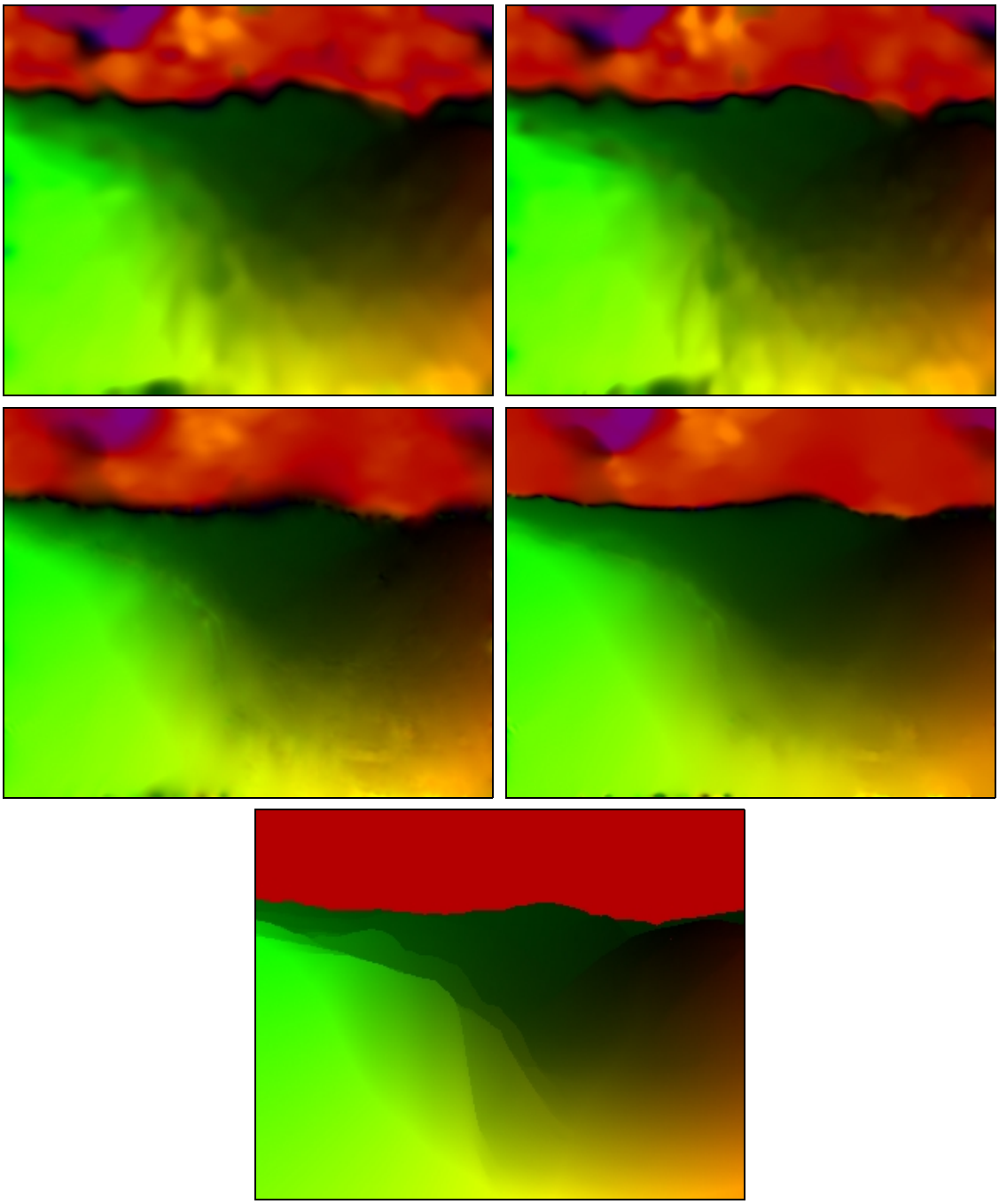

Figure 27: Estimated flow fields of the Yosemite sequence, spatio-temporal versions.
FROM LEFT TO RIGHT, TOP TO BOTTOM:
(a) Linear, None. (b) Nonlinear, None.
(c) None, Linear. (d) None, Isotropic.
(e) Correct flow field.

5.3.4 Noise and presmoothing methods

While up to now the test sequences were without noise, in the following tests the Yosemite sequence was degraded by different levels of Gaussian noise in order to test different diffusion methods for the presmoothing step. So far only homogeneous diffusion has been applied here. In the following it will be checked whether other diffusion methods perform better for certain noise levels.

In order to verify the performance of the different presmoothing methods the conventional Horn-Schunck technique with no smoothing of the structure tensor and a quadratic regularizer was applied to the presmoothed sequence, since this technique is known to be very sensitive to noise. A presmoothing that performs best for this technique can also be assumed to be nearly optimal for the other estimation methods.
For the nonlinear diffusion methods the Charbonnier diffusivity was chosen. There is no backward diffusion with this diffusivity function, so there is edge preservation but no edge enhancement. Edge enhancement would have negative effects for the optic flow estimation methods, because it propagates areas without structure information.

Table 9: Results for the Yosemite sequence and different noise levels with the conventional Horn-Schunck technique and different presmoothing methods.
σ_n denotes the Gaussian noise added to the sequence.

method	σ_n	t_{pre}	λ_{pre}	α	AAE	standard deviation
Homogeneous	0	0.75	-	450	7.20°	± 9.01°
Isotropic		0.75	64	500	7.23°	± 9.06°
Anisotropic		0.75	64	500	7.24°	± 9.06°
Homogeneous	5	1.5	-	700	8.27°	± 9.57°
Isotropic		1.5	16	700	7.97°	± 10.18°
Anisotropic		5	1	700	7.87°	± 9.33°
Homogeneous	10	2.25	-	1000	10.10°	± 9.95°
Isotropic		2	16	1000	9.02°	± 9.69°
Anisotropic		3	4	1000	8.91°	± 9.59°
Homogeneous	40	5.25	-	2000	19.07°	± 12.84°
Isotropic		4.5	16	2250	15.50°	± 11.74°
Anisotropic		5.5	8	2000	15.54°	± 11.80°

The results show that for a sequence without noise a nonlinear diffusion method yields no improvements. In fact, the nonlinear diffusion methods work like the linear method, as the contrast parameter is chosen very high for the optimum result. When noise is added to the sequence things become different. It can be seen that even for rather low noise levels the nonlinear methods perform better, and for the heavily distorted sequence they are clearly superior. Mostly the anisotropic version performs slighly better than the isotropic one.

As the structure tensor has also some capabilities to remove noise, it was tested with the noisy sequence, too. For presmoothing the optimum settings from Table 9 were taken.

Table 10: Yosemite sequence at different noise levels. Comparison between different smoothing methods for the structure tensor. For all methods the optimum presmoothing parameters from Table 9 were taken. σ_n denotes the Gaussian noise added to the sequence.

Smoothing method for the ST	σ_n	t_{ST}	α	AAE	standard deviation
None (Horn-Schunck)	0	-	450	7.20°	± 9.01°
Linear (CLG)		0	450	7.20°	± 9.01°
Nonlinear		14	360	7.07°	± 9.48°
None (Horn-Schunck)	5	-	700	7.87°	± 9.33°
Linear (CLG)		0	700	7.87°	± 9.33°
Nonlinear		40	400	7.71°	± 10.61°
None (Horn-Schunck)	10	-	1000	8.91°	± 9.59°
Linear (CLG)		0	1000	8.91°	± 9.59°
Nonlinear		56	600	8.65°	± 10.90°
None (Horn-Schunck)	40	-	2250	15.50°	± 11.74°
Linear (CLG)		150	1400	13.75°	± 13.24°
Nonlinear		550	1200	13.92°	± 13.21°

For low noise levels the linear structure tensor obviously does not yield any improvements. Only for rather high noise levels the CLG technique is considerably better than the conventional Horn-Schunck method. Then it outperforms even the technique with the nonlinear structure tensor that is in principle superior for lower noise levels. Such kind of behaviour could already be observed when the nonlinear structure tensor was tested with the Lucas-Kanade method. However, the difference between the linear and nonlinear structure tensor now decreased to only 1%.

5.4 Summary

At the end of this chapter, having the results from the test runs in mind, some summarizing statements can be made. First of all, the nonlinear structure tensor turned out to be useful not only for the Lucas-Kanade method but also for the general technique. It always performed better than the linear structure tensor, except in the case of sequences that were heavily distorted by noise. However, such high levels of noise are very rare in practise.

From a theoretical point of view the nonlinear structure tensor is also important in order to complete the framework of the general optic flow estimation technique. In most cases the application of the nonlinear structure tensor yielded better results. Of course it could not always improve results, as especially for divergent motion the assumptions that justify the application of a smoothed structure tensor are not valid. In such cases the diffusion time of the structure tensor, also the nonlinear one, is better set to zero.

Although the general method can not always take advantage of all smoothing steps, in general the most complex combination of diffusion methods using anisotropic diffusion for all stages is the best one, at least from a scientific point of view. In practical applications, of course, it would have to be decided whether a slightly improved result is worth the additional effort. This holds especially for the usage of an anisotropic regularizer. In cases where computational costs play an important role the CLG technique using the nonlinear structure tensor and a quadratic regularizer probably offers the best compromise between quality and efficiency. For noisy sequences it is further advisable to use a nonlinear diffusion method for presmoothing.

6 Conclusions

6.1 Summary

In the previous chapters two novel methods were introduced. First, it was demonstrated that recent concepts for matrix-valued diffusion can be used to construct a new nonlinear structure tensor. As the straight-forward approach was shown to have some drawbacks, a modified diffusion process was introduced. This diffusion process was designed to be closer to homogeneous diffusion in order to keep the benefits of the conventional linear structure tensor. The result was demonstrated to meet the requirements for a nonlinear structure tensor. In tests with optic flow estimation techniques the nonlinear structure tensor then also proved to be superior to the linear structure tensor in a real application.

Secondly, the CLG technique from Weickert et al. [WBS01] based on linear smoothing processes was extended to a discontinuity preserving, nonlinear version. The resulting technique unifies several methods in the field of differential optic flow estimation such as Horn-Schunck, Lucas-Kanade and their discontinuity preserving variants. Furthermore, all these methods are put in a framework that consists of three smoothing steps and different combinations of diffusion techniques. In this scope the novel nonlinear structure tensor was important in order to complete the framework. In a test series the performance of various of such combinations was verified. It could be observed that the usage of more advanced nonlinear diffusion processes improved the results. The best combinations were shown to be competitive or even superior to state-of-the-art optic flow estimation techniques described in the literature. By comparing the results of actually similar methods it also turned out that implementation aspects are very important to make use of the full potential of a technique.

6.2 Further Work

The results achieved with the nonlinear structure tensor nourish expectations that, besides optic flow estimation, it can also be applied in other methods where the linear structure tensor is used. Especially in the field of texture analysis the nonlinear structure tensor will likely benefit from its discontinuity preservation property. However, there are further applications such as edge and corner detection where improvements in comparison to the linear structure tensor are also possible. However, it must be mentioned that the nonlinear structure tensor has higher computational costs than the linear one. Therefore efficient algorithms for anisotropic diffusion are also an essential topic for future research.

For the presented general optic flow estimation technique there are still several improvements conceivable. So it is possible to try a method based on robust statistics [BA96, YD99] instead of the structure tensor. In the sense of robust statistics the nonlinear structure tensor can only avoid outliers caused by motion discontinuities. Other violations against the assumption of local constant flow are not treated. However, it must be mentioned that the nonlinear structure tensor does more than just ignore outliers. It rather changes the assumption of local constant flow itself to the assumption of piecewise constant flow. Nevertheless, a direct and fair comparison between results of the Lucas-Kanade method with the nonlinear structure tensor and a method based on robust statistics for various image sequences would be quite interesting.

The test results suggest the structure tensor and the regularizer not to be orthogonal to each other, i.e. the regularizer often smooths in directions where smoothing would no longer be necessary due to the structure tensor. Therefore results could possibly be improved by adjusting the regularizer to the structure tensor. The regularizer would then be “structure-tensor-driven” instead of flow-driven. This could also speed up computation compared to flow-driven approaches, as the underlying diffusion process would become linear again.

Another possible improvement that comes into mind when regarding the cloudy sky of the Yosemite sequence is a non-quadratic data term in order to permit some violations against the assumption of grey-value constancy. In such a case some changes in the grey-value, as it happens for the clouds in the Yosemite sequence, would probably no longer lead to daring estimations of the flow field in those areas. Using a non-linearized optic flow constraint like in [AWS00] would furthermore improve the estimation of large displacements like those in the lower left corner of the Yosemite sequence.

Considering the applicability of the mentioned optic flow estimation techniques it becomes obvious that they are not usable for real time applications yet. Especially the nonlinear methods, which supply the best results, yield computation times up to an hour on contemporary hardware. Therefore the development of efficient algorithms is also an important item in future research. The same way suitable hardware that allows to make use of the high parallelism in the methods, together with the suitable parallel algorithms, is a very interesting issue in this scope.

Besides this aspect of high computation times, the applicability of the mentioned techniques also suffers from the parameters that have to be optimized for each case of application. In Chapter 5 the optimal parameter settings can be seen to differ considerably for two different image sequences, like for example the Yosemite sequence and the Street sequence. This limits the application on cases with very sharply restricted environment conditions. In this respect, methods with less or more robust parameters are very useful. It was shown in the case of the nonlinear structure tensor that it is possible to get rid of several free parameters. The same way it might be possible to develop a diffusion method for the regularizers without a contrast parameter. Another strategy is to develop techniques that find a nearly optimal parameter setting automatically by analyzing the input data. First approaches in this direction are statistical methods that estimate the noise in the data in order to choose a good parameter for the noise scale. Similar approaches for the remaining smoothing steps would certainly be more complicated.

Bibliography

[ADK99] G. Aubert, R. Deriche, and P. Kornprobst. Computing optical flow via variational techniques. *SIAM Journal on Applied Mathematics*, 60(1):156-182, 1999.

[AELS99] L. Alvarez, J. Esclarín, M. Lefébure, and J. Sánchez. A PDE model for computing the optical flow. In *Proc. XVI Congreso de Ecuaciones Diferenciales y Aplicaciones*, pages 1349-1356, Las Palmas de Gran Canaria, Spain, September 1999.

[AWS00] L. Alvarez, J. Weickert, and J. Sánchez. Reliable estimation of dense optical flow fields with large displacements. *International Journal of Computer Vision*, 39(1):41-56, August 2000.

[BA91] M. J. Black and P. Anandan. Robust dynamic motion estimation over time. In *Proc. 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 292-302, Maui, HI, Hune 1991. IEEE Computer Society Press.

[BA96] M. J. Black and P. Anandan. The robust estimation of multiple motions: parametric and piecewise smooth slow fields. *Computer Vision and Image Understanding*, 63(1):75-104, January 1996.

[BFB94] J. L. Barron, D. J. Fleet, and S. S. Beauchemin. Performance of optical flow techniques. *International Journal of Computer Vision*, 12(1):43-77, February 1994.

[BG88] J. Bigün and G. H. Granlund. Optical flow based on the inertia matrix in the frequency domain, *Proc. SSAB Symposium on Picture Processing*, Lund, Sweden, March 1988.

[BGW91] J. Bigün, G. H. Granlund, and J. Wiklund. Multidimensional orientation estimation with applications to texture analysis and optical flow. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 13(8):775-790, August 1991.

[Bru01] A. Bruhn. *Regularization in Motion Estimation*. Diploma thesis, Computer Vision, Graphics, and Pattern Recognition Group, Institute of Computer Engineering, University of Mannheim, Germany, 2001.

[CBFAB94] P. Charbonnier, L. Blanc-Féraud, G. Aubert, and M. Barlaud. Two deterministic half-quadratic regularization algorithms for computed imaging. In *Proc. 1994 IEEE International Conference on Image Processing*, volume 2, pages 168-172, Austin, TX, November 1994. IEEE Computer Society Press.

[CH53] R. Courant and D. Hilbert. *Methods of Mathematical Physics*, volume 1. Interscience, New York, 1953.

[CLMC92] F. Catté, P.-L. Lions, J.-M. Morel, and T.Coll. Image selective smoothing and edge detection by nonlinear diffusion, *SIAM Journal on Numerical Analysis*, 32:1895-1909, 1992.

[Coh93] I. Cohen. Nonlinear variational method for optical flow computation. In *Proc. Eighth Scandinavian Conference on Image Analysis*, volume 1, pages 523-530, Tromsø, Norway, May 1993.

[EF98] M. Elad and A. Feuer. Recursive optical flow estimation - adaptive filtering approach. *Journal of Visual Communication and Image Representation*, 9(2):119-138, June 1998.

[Els61] L. E. Elsgolc. *Calculus of Variations*. Pergamon, Oxford, 1961.

[FG87] W. Förstner and E. Gülch. A fast operator for detection and precise location of distinct points, corners and centres of circular features, In *Proc. ISPRS Intercommission Conference on Fast Processing of Photogrammatric Data*, pages 281-305, Interlaken, Switzerland, June 1987.

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

[GMN+98] B. Galvin, B. McCane, K. Novins, D. Mason, and S. Mills. Recovering motion fields: an analysis of eight optical flow algorithms. In *Proc. 1998 British Machine Vision Conference*, Southhampton, England, September 1998.

[HB93] F. Heitz and P. Bouthemy. Multimodal estimation of discontinuous optical flow using Markov random fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(12):1217-1232, December 1993.

[HS81] B. Horn and B Schunck. Determining optical flow. *Artificial Intelligence*, 17:185-203, 1981.

[Iij59] T. Iijima. Basic theory of pattern observation. In *Papers of Technical Group on Automata and Automatic Control*. IECE, Japan, December 1959. In Japanese.

[KTB96] A. Kumar, A. R. Tannenbaum, and G. J. Balas. Optic flow: a curve evolution approach. *IEEE Transactions on Image Processing*, 5(4):598-610, April 1996.

[LK81] B. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *Proc. Seventh International Joint Conference on Artificial Intelligence*, pages 674-679, Vancouver, Canada, August 1981.

[Luc84] B. Lucas. *Generalized Image Matching by the Method of Differences*, PhD thesis, School of Computer Science, Carnegie-Mellon University, Pittsburgh, PA, 1984.

[MP98] E. Mémin and P. Pérez. Dense estimation and object-based segmentation of the optical flow with robust techniques. *IEEE Transactions on Image Processing*, 7(5):703-719, May 1998.

[Nag83] H.-H. Nagel. Constraints for the estimation of displacement vector fields from image sequences. In *Proc. Eighth International Joint Conference on Artificial Intelligence*, colume 2, pages 945-951, Karlsruhe, Germany, August 1983.

[Nes93] P. Nesi. Variational approach to optical flow estimation managing discontinuities. *Image and Vision Computing*, 11(7):419-439, September 1993.

[NG98] H.-H. Nagel and A. Gehrke. Spatiotemporally adaptive estimation and segmentation of OF-fields. In H. Burkhardt and B. Neumann, editors, *Computer Vision - ECCV '98*, volume 1407 of *Lecture Notes in Computer Science*, pages 86-102. Springer, Berlin, 1998.

[PM87] P. Perona and J. Malik. Scale space and edge detection using anisotropic diffusion. In *Proc. IEEE Computer Society Workshop on Computer Vision*, pages 16-22, Miami Beach, FL, November 1998. IEEE Computer Society Press.

[PTVF92] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in C*. Cambridge University Press, Cambridge, UK, second edition, 1992.

[PVPO94] M. Poesmans, L. Van Gool, E. Pauwels, and A. Oosterlinck. Determination of optical flow and its discontinuities using non-linear diffusion. In J.-O. Eklundh, editor, *Computer Vision - ECCV '94*, volume 801 of *Lecture Notes in Computer Science*, pages 295-304. Springer, Berlin, 1994.

[Sch91] C. Schnörr. Determining optical flow for irregular domains by minimizing quadratic functionals of a certain class. *International Journal of Computer Vision*, 6(1):25-38, April 1991.

[Sch94] C. Schnörr. Bewegungssegmentation von Bildfolgen durch die Minimierung konvexer nicht-quadratischer Funktionale. In W. Kropatsch and H. Bischof, editors, *Mustererkennung 1994*, pages 178-185. Springer, Berlin, 1994.

[SH89] D. Shulman and J. Hervé. Regularization of discontinuous flow fields. In *Proc. Workshop on Visual Motion*, pages 81-90, Irvine, CA, March 1989. IEEE Computer Society Press.

[TD01] D. Tschumperlé and R. Deriche. Diffusion tensor regularization with constraints preservation, In *Proc. 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1, 948-953, Kauai, HI, December 2001. IEEE Computer Society Press.

[WB97] J. Weickert and B. Benhamouda. A semidiscrete nonlinear scale-space theory and its relation to the Perona-Malik paradox. In F.Solina, W. G. Kropatsch, R. Klette, and R. Bajcsy, editors, *Advances in Computer Vision*, pages 1-10. Springer, Wien, 1997.

[WB02] J. Weickert and T. Brox. Diffusion and regularization of vector- and matrix-valued images. Technical Report 58, Department of Mathematics, Saarland University, Saarbrücken, Germany, March 2002.

[WBS01] J. Weickert, A. Bruhn, and C. Schnörr. Lucas/Kanade meets Horn/Schunck: Combining local and global optic flow methods. Technical Report 17/2001, Computer Science Series, University of Mannheim, Germany, September 2001.

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

[Wei98] J. Weickert. *Anisotropic Diffusion in Image Processing*, Teubner, Stuttgart, 1998.

[Wei01] J. Weickert. Applications of partial differential equations in image processing. Lecture notes, Faculty of Mathematics and Computer Science, University of Mannheim, Germany, 2001.

[WRV98] J. Weickert, B. M. ter Romeny, and M. A. Viergever. Efficient and reliable schemes for nonlinear diffusion filtering. *IEEE Transactions on Image Processing*, 7(3):398-410, March 1998.

[WS01a] J. Weickert and C. Schnörr. A theoretical framework for convex regularizers in PDE-based computations of image motion. *International Journal of Computer Vision*, 45(3):245-264, December 2001.

[WS01b] J. Weickert and C. Schnörr. Variational optic flow computation with a spatio-temporal smoothness constraint. *Journal of Mathematical Imaging and Vision*, 14(3):245-255, May 2001.

[YD99] Y. Yacoob and L. S. Davis. Temporal multi-scale models for flow and acceleration. *International Journal of Computer Vision*, 32(2):1-17, September 1999.

[You71] D. M. Young. *Iterative Solution of Large Linear Systems.* Academic Press, New York, 1971.