

Nonlinear Matrix Diffusion for Optic Flow Estimation

Thomas Brox and Joachim Weickert

Faculty of Mathematics and Computer Science, Saarland University,
Building 27.1, P. O. Box 15 11 50, 66041 Saarbrücken, Germany.
{brox,weickert}@mia.uni-saarland.de

Abstract. In this paper we present a method for nonlinear diffusion of matrix-valued data. We adapt this technique to the well-known linear structure tensor in order to develop a new nonlinear structure tensor. It is then used to improve the optic flow estimation methods of Lucas and Kanade and its spatio-temporal variant of Bigün et al.. Our experiments show that the nonlinear structure tensor leads to a better preservation of discontinuities in the optic flow field.

1 Introduction

Nonlinear diffusion techniques have proved to be very useful for discontinuity-preserving denoising of scalar and vector-valued data. Apart from very recent work [10, 12], however, not many attempts have been made to design diffusion filters for matrix-valued data. One important representative of matrix-valued data fields is the *structure tensor* (*ST*), a frequently used tool for corner detection [3], texture [9] and image sequence analysis [2, 5]. The conventional formulation of the structure tensor uses Gaussian smoothing which is equivalent to linear diffusion filtering. This is well-known to blur across data discontinuities.

The goal of the present paper is to formulate a *nonlinear* structure tensor that respects discontinuities in the data. The nonlinear ST can be used in any application working with the conventional linear ST. In our paper we focus on its evaluation for estimating optic flow fields. The well-known optic flow method of Lucas and Kanade [6] or its spatio-temporal counterpart by Bigün et al. [2] use a linear ST that integrates across a neighborhood of a fixed size. The novel nonlinear ST adapts this neighborhood to the data, preserving discontinuities in the optic flow field.

Addressing these issues is the goal of the present paper. It is organized as follows. In Section 2 we present our nonlinear diffusion method for the ST. In Section 3 first the optic flow estimation method of Lucas and Kanade is briefly reviewed. We then apply the nonlinear ST and present experimental results in Section 4. The paper is concluded by a summary in Section 5.

Related work. The approach of Tschumperlé and Deriche [10] uses space-variant diffusion with a scalar-valued diffusivity. In contrast to our method with a diffusion tensor, it may be classified as isotropic. Furthermore it is not focusing

on the ST, but on diffusion tensor MRI. In the work of Weickert and Brox [12] a general method for diffusing or regularizing matrix-valued data is proposed. The present model differs from this work not only by the fact that the process is specifically adapted to the ST, but also by its application to optic flow estimation. In this sense our work is close in spirit to the interesting papers of Nagel and Gehrke [8] and Middendorf and Nagel [7]. These authors use shape-adapted Gaussians for designing structure tensors that average over regions with similar grey values. While *homogeneous* Gaussian convolution and *linear* diffusion are equivalent, it should be noted that this is no longer the case with *space-variant* Gaussian smoothing and *nonlinear* diffusion. Since scalar-valued nonlinear diffusion filters offer a sound mathematical underpinning, it appears promising to investigate also a nonlinear diffusion formulation of the structure tensor.

2 Matrix-valued Diffusion

Let us first illustrate the limitations of the conventional linear ST by an example. Figure 1a shows a synthetic test image f which is distorted by Gaussian noise with $\sigma = 30$. Figure 1b depicts the matrix product $J_0 = \nabla f \nabla f^\top$ as a colored orientation plot. The direction of the eigenvector to the largest eigenvalue of J_0 is mapped to the hue value and the largest eigenvalue to the intensity value in the HSI color model. The saturation value is set to its maximum.¹

The linear ST J_ρ can be seen in Figure 1c. It is derived from J_0 by smoothing each component by a Gaussian kernel with standard deviation ρ . This technique closes structures of a certain scale very well. It also removes the noise appropriately. On the first glance surprising for a linear technique is the preservation of orientation discontinuities. However, discontinuities in the magnitude are not preserved causing object boundaries to dislocate.

This problem can be addressed by replacing the convolution with a Gaussian kernel by a discontinuity preserving diffusion method. However, all capabilities of the linear ST should remain. Therefore, we keep the diffusivity at its maximum except at locations where discontinuities in the magnitude exist. This is done by regarding J_0 as initial matrix field that is evolved under the diffusion equation

$$\partial_t u_{ij} = \operatorname{div} \left(D \left(\left(\nabla_\sigma \sqrt{\sum_{k,l} u_{kl}^2} \right) \left(\nabla_\sigma \sqrt{\sum_{k,l} u_{kl}^2} \right)^\top \right) \nabla u_{ij} \right) \quad \forall i, j \quad (1)$$

where the evolving matrix field $u_{ij}(x, t)$ uses $J_0(x)$ as initial condition for $t = 0$. The matrix $D(A) = T(g(\lambda_i))T^\top$ is the diffusion tensor for $A = T(\lambda_i)T^\top$ where the last-mentioned expression denotes a principal axis transformation of A with the eigenvalues λ_i as the elements of a diagonal matrix (λ_i) and the normalized eigenvectors as the columns of the orthogonal matrix T .

The diffusivity $g(s^2)$ is a decreasing function such as $g(s^2) = 1 - e^{-\frac{3.31488\lambda^8}{s^8}}$ with a contrast parameter λ . By ∇_σ we denote the nabla operator where Gaussian

¹ A color version of this paper will be provided in the internet.

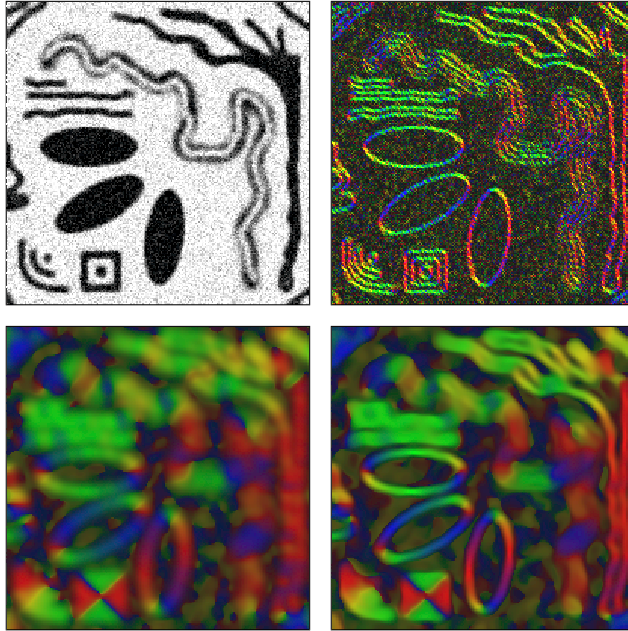


Fig. 1. (a) TOP LEFT: Synthetic image with Gaussian noise. (b) TOP RIGHT: $J_0 = \nabla f \nabla f^\top$. (c) BOTTOM LEFT: Linear structure tensor J_ρ with $\rho = 3$. (d) BOTTOM RIGHT: Nonlinear structure tensor J_t with $t = 12.5$.

derivatives with standard deviation σ are used. For more detailed information about diffusion equations in general we refer to [11].

On the first glance the fourth root in Eq. 1 seems to be quite arbitrary, but there is a good motivation for it: For diffusion time $t = 0$ the structure tensor is

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

where subscripts denote partial derivatives. In this case we have

$$\sqrt[4]{\sum_{k,l} u_{kl}^2} = \sqrt[4]{f_x^4 + 2f_x^2 f_y^2 + f_y^4} = \sqrt[4]{(f_x^2 + f_y^2)^2} = \sqrt{f_x^2 + f_y^2} = |\nabla f|. \quad (3)$$

This leads to the interpretation that the image gradient ∇f drives the diffusion. Precisely speaking, this is only exactly the case for $t = 0$: The diffusivity is adapted to the new structure tensor after each time step, therefore resulting in a nonlinear diffusion process.

The nonlinear ST obtained by Equation 1 is depicted in Figure 1d. The result is exactly what we expect from a nonlinear ST: While object boundaries are no longer dislocated, all positive properties of the linear ST remain valid. Noise is

removed, structures of a certain scale are closed and orientation discontinuities are preserved. Although there are some additional parameters for the nonlinear ST, they are not really a problem. The diffusion time t simply replaces the scale parameter ρ of the linear ST. The other parameters, namely the diffusivity function $g(s^2)$, its contrast parameter λ as well as the presmoothing parameter σ , are very robust against variations and can be fixed, still yielding good results for a whole set of input data.

3 Optic Flow Estimation

Before we test the performance of the new nonlinear structure tensor by applying it to optic flow estimation, the classic estimation method of Lucas and Kanade [6] using the linear ST is briefly reviewed.

Assuming that image structures do not alter their grey values during their movement can be expressed by the optic flow constraint

$$f_x u + f_y v + f_z = 0. \quad (4)$$

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 (*aperture problem*). A second assumption has to be made. Lucas and Kanade proposed to assume the optic flow vector to be constant within some neighborhood B_ρ of size ρ . The optic flow in some point (x_0, y_0) can then be estimated by the minimizer of the local energy function

$$E(u, v) = \frac{1}{2} \int_{B_\rho(x_0, y_0)} (f_x u + f_y v + f_z)^2 dx dy. \quad (5)$$

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

$$\begin{pmatrix} \int_{B_\rho} f_x^2 dx dy & \int_{B_\rho} f_x f_y dx dy \\ \int_{B_\rho} f_x f_y dx dy & \int_{B_\rho} f_y^2 dx dy \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -\int_{B_\rho} f_x f_z dx dy \\ -\int_{B_\rho} f_y f_z dx dy \end{pmatrix}. \quad (6)$$

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

$$\begin{pmatrix} K_\rho * f_x^2 & K_\rho * f_x f_y \\ K_\rho * f_x f_y & K_\rho * f_y^2 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -K_\rho * f_x f_z \\ -K_\rho * f_y f_z \end{pmatrix} \quad (7)$$

where the entries of the linear system are five of the components of the linear ST J_ρ . The linear system can be solved provided the system matrix is not singular. Such singular matrices appear in regions where the image gradient vanishes. They also appear in regions where the aperture problem remains present, leading to the smaller eigenvalue of the system matrix being close to 0. In this case one may only compute the so-called normal flow (the optic flow component parallel

to the image gradient). Using sufficiently broad Gaussian filters, however, will greatly reduce such singular situations. In these cases, one may obtain results with densities close to 100 %. In order to improve the quality of the estimated optic flow field, the density can be reduced by using the smaller eigenvalue of the system matrix as confidence measure [1],

By replacing all spatial integrations in Equations 5–7 by spatio-temporal integrations, one ends up with a method that is equivalent to the structure tensor approach of Bigün et al. [2]. The spatio-temporal approach in general yields better results than the spatial one.

The Gaussian convolution in the structure tensor methods of Lucas–Kanade and Bigün is well known to be equivalent to linear diffusion filtering. With our knowledge from Section 2 we may now introduce corresponding *nonlinear* versions of both techniques by replacing the linear ST in equation 7 by the nonlinear one. As mentioned above only the diffusion time is a critical parameter. All results presented in the next section have been achieved with the diffusivity function $g(s^2) = 1 - e^{-\frac{3.31488\lambda^8}{s^8}}$, a contrast parameter $\lambda = 0.1$ and a regularization parameter $\sigma = 1.5$. Only for the noise experiments in Table 2, we adapted σ to the noise.

4 Results

A good image sequence to demonstrate the discontinuity preserving property of the new technique is the Hamburg taxi sequence. In this scene there are four moving objects: a taxi turning around the corner, a car moving to the right, a van moving to the left and a pedestrian in the upper left². Figure 2 shows that using the linear structure tensor in the Lucas–Kanade method causes the flow fields of moving objects to dislocate, whereas the nonlinear structure tensor ensures that object boundaries are preserved in a better way. Like in Barron et al. [1] the sequence was presmoothed along the time axis by a Gaussian kernel with $\sigma = 1$.

In another experiment we used a synthetic street sequence³. For this sequence created by Galvin et al. [4] the ground truth flow field is available. This enables the computation of the average angular error between the estimated flow and the ground truth flow field as a quantitative measure [1]. Table 1 shows the angular errors of some algorithms from the literature as well as the linear and nonlinear ST methods. A direct comparison between the angular errors of the method using the linear ST and nonlinear ST respectively, quantifies the improvement achieved with our new technique. It should be noted that the discontinuity locations constitute only a small subset of the entire image. This explains an effect that can be observed for all discontinuity preserving optic flow methods: Visually significant improvements at edges can only lead to moderate improvements for a global measure such as the average angular error.

² The sequence is available from <ftp://csd.uwo.ca> under the directory `pub/vision`

³ The sequence can be obtained from www.cs.otago.ac.nz/research/vision

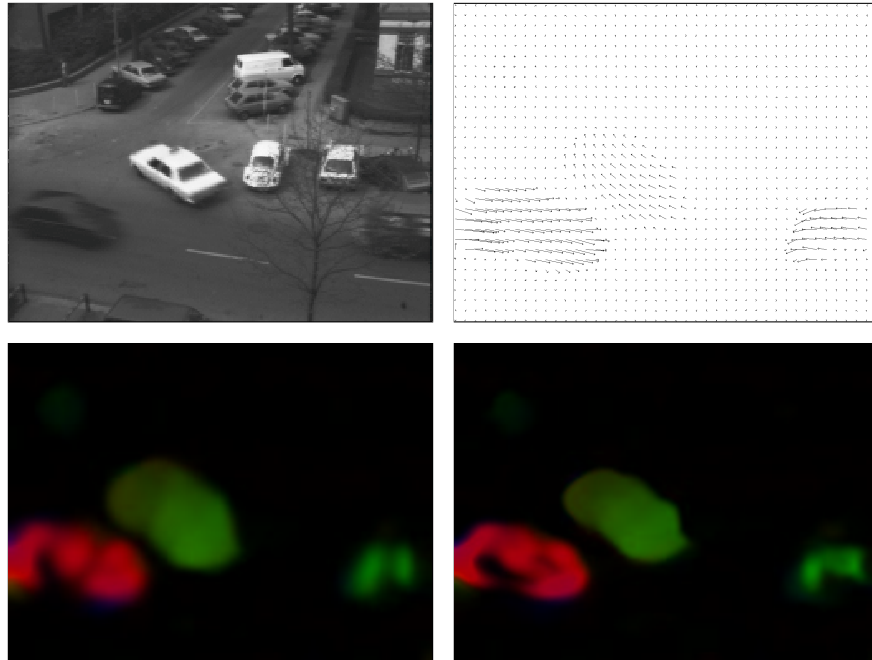


Fig. 2. TOP LEFT: Hamburg Taxi Sequence (Frame 9). TOP RIGHT: Optic flow field with the nonlinear ST. BOTTOM LEFT: Flow magnitude with the linear ST. BOTTOM RIGHT: Flow magnitude with the nonlinear ST.

Since the classic Lucas and Kanade technique using the linear ST is known to be very robust against noise, we degraded the Street sequence by Gaussian noise to test whether the nonlinear ST yields any drawbacks in this respect. Table 2 demonstrates that our technique is superior to the original one even in the presence of severe noise.

5 Conclusions

In this paper it was shown how new diffusion methods for matrix-valued data can be used to construct a nonlinear structure tensor. We presented a special diffusion method which keeps all the benefits of the conventional linear ST, but tackles its problem of object delocalizations. Afterwards it was shown that the nonlinear ST could serve to improve the classic optic flow estimation techniques of Lucas and Kanade and of Bigün et al.. Our experiments revealed the superiority of the nonlinear ST to the conventional linear one. Moreover, also in comparison to other estimation techniques our method performed competitive. This supports the expectation that the nonlinear ST can also improve the results of other methods where a linear ST is used.

Table 1. Street sequence. Comparison between the best results from the literature and our results. AAE = average angular error.

Technique	AAE	Density
Camus [4]	13.69°	100%
Proesman et al. [4]	7.41°	100%
Weickert-Schnörr 2D [13]	6.62°	100%
Lucas-Kanade (2D) Linear	6.29°	100%
Lucas-Kanade (2D) Nonlinear	5.88°	100%
Bigün (3D) Linear	5.28°	100%
Bigün (3D) Nonlinear	5.14°	100%
Weickert-Schnörr 3D [13]	4.85°	100%
Uras et al. [4]	6.93°	54%
Horn-Schunck [4]	6.62°	46%
Singh [4]	6.18°	78%
Lucas-Kanade (2D) Linear	4.82°	52%
Lucas-Kanade (2D) Nonlinear	4.51°	53%
Bigün (3D) Linear	3.49°	58%
Bigün (3D) Nonlinear	3.30°	57%

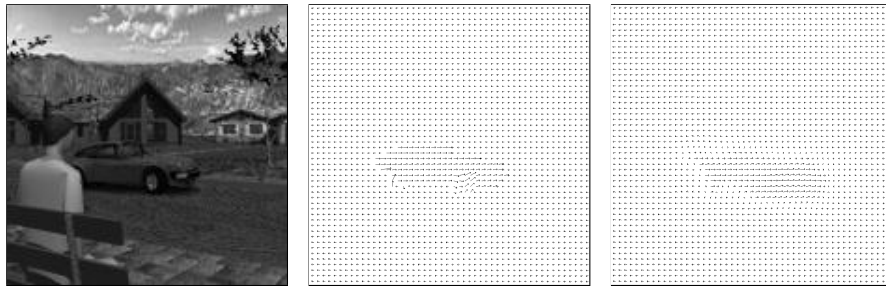


Fig. 3. LEFT: Detail from Street sequence (Frame 10, 150×150 pixels). CENTER: Ground truth flow field. RIGHT: Flow field with the nonlinear ST (3D).

Table 2. Street Sequence. Average angular errors for different noise levels (2D version, 100% density). σ_n denotes the standard deviation of the Gaussian noise.

σ_n	linear ST	nonlinear ST
0	6.29°	5.88°
5	7.63°	7.33°
10	10.93°	10.36°
20	15.87°	14.79°

Acknowledgements

Our research on matrix-valued smoothing methods is partly funded by the projects WE 2602/1-1 and SO 363/9-1 of the *Deutsche Forschungsgemeinschaft (DFG)*. This is gratefully acknowledged.

References

1. J. L. Barron, D. J. Fleet, and S. S. Beauchemin. Performance of optical flow techniques. *International Journal of Computer Vision*, 12(1):43–77, Feb. 1994.
2. J. Bigün, G. H. Granlund, and J. Wiklund. Multidimensional orientation estimation with applications to texture analysis and optical flow. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(8):775–790, Aug. 1991.
3. 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 Photogrammetric Data*, pages 281–305, Interlaken, Switzerland, June 1987.
4. 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*, Southampton, England, Sept. 1998.
5. B. Jähne. *Spatio-Temporal Image Processing*, volume 751 of *Lecture Notes in Computer Science*. Springer, Berlin, 1993.
6. 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, Aug. 1981.
7. M. Middendorf and H.-H. Nagel. Estimation and interpretation of discontinuities in optical flow fields. In *Proc. Eighth International Conference on Computer Vision*, volume 1, pages 178–183, Vancouver, Canada, July 1995. IEEE Computer Society Press.
8. 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.
9. A. R. Rao and B. G. Schunck. Computing oriented texture fields. *CVGIP: Graphical Models and Image Processing*, 53:157–185, 1991.
10. 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, pages 948–953, Kauai, HI, Dec. 2001. IEEE Computer Society Press.
11. J. Weickert. *Anisotropic Diffusion in Image Processing*. Teubner, Stuttgart, 1998.
12. J. Weickert and T. Brox. Diffusion and regularization of vector- and matrix-valued images. Technical Report 58, Department of Mathematics, Saarland University, Saarbrücken, Germany, Mar. 2002.
13. 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.