PDE-Based Anisotropic Disparity-Driven Stereo Vision

Henning Zimmer[1,2], Andrés Bruhn[1], Levi Valgaerts[1], Michael Breuß[1], Joachim Weickert[1],
Bodo Rosenhahn[2], and Hans-Peter Seidel[2]

[1]Mathematical Image Analysis Group, Faculty of Mathematics and Computer Science
Building E1.1, Saarland University, 66041 Saarbrücken, Germany
Email:{zimmer,bruhn,valgaerts,breuss,weickert}@mia.uni-saarland.de
[2]Max-Planck Institute for Informatics,
Stuhlsatzenhausweg 85, 66123 Saarbrücken, Germany
Email:{rosenhahn,hpseidel}@mpi-sb.mpg.de

Abstract

Recent variational stereo approaches suffer from at least one of the following drawbacks: Either they use an *isotropic disparity-driven* smoothness term that ignores the directional information of the disparity field, or they apply *anisotropic image-driven* regularisation that suffers from oversegmentation artifacts. As a remedy, we present a novel *anisotropic disparity-driven* approach for stereo vision. It is designed as a highly adaptive anisotropic diffusion-reaction equation that incorporates a diffusion process which has been used successfully for image denoising and inpainting. Its directional adaptation allows to better control the smoothing w.r.t. the local structure of the disparity field. Experiments that compare our model to a recent isotropic variational method and a probabilistic graph cut approach demonstrate the superior quality of our approach. Moreover, a multigrid algorithm allows for moderate run times that do not depend on the disparity range.

1 Introduction

Stereo vision is an important and challenging part of computer vision research. Although first attempts go back to Marr and Poggio [17] in 1976, qualitatively good results are still hard to obtain. In the usual binocular case, one is given two images of the same scene, captured from two different views, which we denote by 'left' and 'right', respectively. In order to recover the missing depth information of the scene, one has to solve a *correspondence problem*: For each pixel in the left image one has to determine the corresponding *disparity*, i.e., the change of its position w.r.t. the right image.

There are different methods to compute the disparity, which can basically be divided into four classes: (i) *Feature-based* approaches [9], which match characteristic points in the images, e.g., corners, (ii) *area-based* approaches [20], matching pixels if patches around them exhibit a certain similarity, (iii) *phase-based* approaches [7], that use the phase information in the Fourier domain, and finally (iv) *energy-based* approaches [2, 12, 13, 14, 16, 20, 21], which find the disparity by minimising an energy functional that penalises deviations from data and smoothness assumptions. The latter class can be further divided into *probabilistic* and *variational* approaches. The first type [12, 13, 20] models images and disparity as Markov random fields and tries to find the most probable disparity, given the two images. This comes down to the minimisation of a discrete energy which is usually done by graph cuts (GC) [13], belief propagation (BP) [12] or dynamic programming (DP) [14] algorithms. These methods are quite successful as they usually impose strict smoothness assumptions, modelling a piecewise constant disparity. However, Li and Zucker [15] have shown that such approaches may have severe drawbacks if the assumption of a piecewise constant disparity is violated. This can be the case if the depth is varying smoothly, for instance in the presence of curved or slanted surfaces. Moreover, probabilistic approaches suffer from their discrete nature, since they only assign integer disparity values to the pixels.

These restrictions do not apply to the second type of energy-based methods, *variational* approaches. Here, the disparity is computed by the minimisation of a continuous energy functional which can be done by a gradient descent method. This requires

to compute the steady-state of a partial differential equation (PDE), which is of *diffusion-reaction* type. Variational approaches go back to the work of Horn and Schunck [10], where they were first successfully introduced in optical flow computations. For stereo, they were used, among others, in the work of Slesareva *et al.* [21], where the authors adapted the very accurate optical flow method of [3] to the weakly calibrated case. By exploiting the known geometry of the two views, they restrict the search for correspondences along epipolar lines. In this work we restrict ourselves to the scenario where the two images have been *rectified* beforehand and displacements only occur in horizontal direction. Thus the disparity boils down to a pixelwise scalar value. A recent variational stereo method for the rectified case is proposed in [2], which additionally incorporates segmentation ideas and occlusion handling to further improve results at disparity boundaries.

One important design aspect of variational methods is the choice of the *regulariser* modelling the smoothness assumptions. Recent variational stereo approaches either use isotropic disparity-driven regularisers [2, 21], which adapt the smoothing of the disparity map w.r.t. the magnitude of the disparity gradient, or anisotropic image-driven regularisers that try to preserve edges in accordance with the image data [1, 16]. For most cases, anisotropic processes have shown to be superior to their isotropic counterparts, as they offer a higher accuracy at image edges and thin structures. Disparity-driven methods generally have an advantage over image-driven ones that tend to give oversegmented results. However, a method that combines these two advantages, has not been proposed so far in a stereo context. To fill the void in existing smoothing strategies, this paper introduces an anisotropic disparity-driven stereo method, which takes into account directional information of the disparity field and thus allows to distinguish between smoothing along and across disparity edges. In [23], Weickert and Schnörr present a theoretical framework for the design of regularisers in the context of optical flow computation, which also includes an anisotropic flow-driven smoothness term. We will show that it is not possible to directly adopt this regularisation in the stereo case, as the resulting diffusion process remains isotropic. As a remedy, we propose a different strategy: Instead of deriving a suitable energy functional, we will directly model a highly adaptive anisotropic diffusion process within the diffusion-reaction equation. Since a corresponding energy formulation is no longer required, we can design more powerful smoothing strategies that are based on nonlinear anisotropic diffusion filters. These filters have already shown their usefulness in the context of image denoising [22] and PDE-based inpainting [8]. In particular, our method will exhibit a distinct behaviour at corners, edges and homogeneous regions.

Our paper is organised as follows: Section 2 introduces basic concepts of variational stereo. After discussing existing types of regularisers we present our new anisotropic method in Section 3. Section 4 shortly describes the solution of the arising PDE, while Section 5 shows experiments that compare our new anisotropic method with an isotropic one, as well as a GC method. Section 6 concludes the paper with a summary and gives an outlook to possible future work.

2 Variational Stereo

2.1 Basic Structure

Assume we are given the rectified image pair $f_l, f_r : \Omega \to \mathbb{R}$, denoting the left and the right view, respectively. Here $\Omega \subset \mathbb{R}^2$ is a rectangular image domain. We further assume that the images are presmoothed by a convolution with a Gaussian kernel of standard deviation σ_{pre}. The unknown horizontal disparity component $u : \Omega \to \mathbb{R}$ is found by minimising an energy functional of the form

$$E(u) = \int_\Omega \left[M(f_l, f_r, u) + \alpha\, V(\nabla u) \right] \, \mathrm{d}\mathbf{x} \,, \quad (1)$$

where $\mathbf{x} := (x, y)^\top \in \Omega$ and $\nabla := (\partial_x, \partial_y)^\top$ denotes the spatial gradient operator. The *data term* $M(f_l, f_r, u)$ models how well the disparity u matches the given data f_l and f_r. In general, this is done by imposing one or several constancy assumptions on image properties. The *smoothness term* or *regulariser* $V(\nabla u)$ enforces the disparity to be smoothly varying in space by penalising large gradients of u. Its influence on the overall energy is steered by a smoothness weight $\alpha > 0$.

We find a minimiser u of the energy functional (1) via a gradient descent method by introducing an artificial evolution parameter t. In other words, we are looking for the steady state solution of the

diffusion-reaction equation

$$u_t = \left(\partial_x V_{u_x} + \partial_y V_{u_y}\right) - \frac{1}{\alpha}\,\partial_u M\ , \quad (2)$$

for $t \to \infty$, with homogeneous Neumann boundary conditions $\partial_\mathbf{n} u = 0$ on $\partial\Omega$. Here the subscripts of u denote partial derivatives and $\mathbf{n}$ denotes the normal vector of the image boundary $\partial\Omega$. The term between brackets on the righthand side comprises the diffusion part which results from the smoothness term of the energy functional. The last term constitutes the reaction part of the equation and stems from the data term.

For the choice of the data term of our method we will follow the approach in [21] and use a combination of the brightness and the gradient constancy assumption:

$$M(f_l, f_r, u) = \Psi_M \Big(|f_r(\mathbf{x} + \mathbf{u}) - f_l(\mathbf{x})|^2$$
$$+ \gamma\,|\nabla f_r(\mathbf{x} + \mathbf{u}) - \nabla f_l(\mathbf{x})|^2 \Big)\ . \quad (3)$$

In the above expression $\mathbf{u} := (u, 0)^\top$, and $\Psi_M(s^2)$ is a differentiable and increasing function that is convex in s. The brightness constancy constraint models the classical assumption that the grey value of a pixel does not change during its displacement [10]. The gradient constancy assumption on the other hand renders the approach more robust under varying illumination conditions, a common problem in real-world images. Its contribution to the overall data term is steered by a parameter $\gamma > 0$. Note that we refrain from linearising the data term to allow for a correct estimation of large disparities. As a robust penaliser function we choose $\Psi_M(s^2) := \sqrt{s^2 + \varepsilon^2}$, where $\varepsilon > 0$ is a small regularisation parameter. This results in a modified L^1 penalisation, which helps us cope with outliers caused by image noise or occlusions. The contribution of the data term (3) to equation (2) will be denoted by $m(f_l, f_r, u) := \partial_u M$ and can be written as follows:

$$m(f_l, f_r, u) = \Psi_M' \left(f_z^2 + \gamma\,(f_{xz}^2 + f_{yz}^2)\right)$$
$$\cdot (f_x f_z + \gamma\,(f_{xx}f_{xz} + f_{xy}f_{yz}))\ . \quad (4)$$

In this equation we made, in accordance with [21], use of the following abbreviations:

$$f_* \quad := \quad \partial_* f_r(\mathbf{x} + \mathbf{u})\ , \quad (5)$$
$$f_z \quad := \quad f_r(\mathbf{x} + \mathbf{u}) - f_l(\mathbf{x})\ , \quad (6)$$
$$f_{*z} \quad := \quad \partial_* f_r(\mathbf{x} + \mathbf{u}) - \partial_* f_l(\mathbf{x})\ , \quad (7)$$

where the variable z is used to emphasise the use of temporal differences in contrast to temporal derivatives.

2.2 Regularisation

We will now give a short overview of existing spatial regularisers for rectified variational stereo. We will follow the taxonomy of Weickert and Schnörr [23], which gives a systematic classification of convex smoothness terms for optical flow computation. Based on their connection with multichannel diffusion filtering, this classification encompasses data-driven and flow-driven as well as isotropic and anisotropic regularisers.

I. Isotropic image-driven regularisation.
This type of regularisation inhibits smoothing of the disparity field at image edges. A recent work in this area was published by Kim and Sohn [11].

II. Anisotropic image-driven regularisation.
This class of regularisers mainly became popular through the works of Mansouri *et al.* [16] and Alvarez *et al.* [1]. The smoothness term makes use of a *diffusion tensor* $D(\nabla f_l, \nabla f_r) \in \mathbb{R}^{2\times2}$ which, compared to isotropic processes, can include additional directional information. This gives rise to more degrees of freedom in the adaptation of the smoothing process to the underlying image structure. The biggest drawback of image driven regularisation lies in the fact that not every image edge necessarily matches a disparity edge. Especially in the presence of textures the resulting disparity field can suffer from oversegmentation.

III. Isotropic disparity-driven regularisation.
A remedy for oversegmented solutions can come from the use of disparity-driven regularisers, which inhibit smoothing at edges of the evolving disparity u. Indeed, most recent successful variational approaches [2, 21] use a regulariser of this type. The smoothness term takes on the form $V_{\text{ID}}(\nabla u) = \Psi_V\left(|\nabla u|^2\right)$ for a non-quadratic penaliser $\Psi_V(s^2)$ which is convex in s. The corresponding diffusion-reaction equation is then given by

$$u_t = \text{div}\,\left(\Psi_V'\left(|\nabla u|^2\right)\nabla u\right)$$
$$- \frac{1}{\alpha}\,m(f_l, f_r, u)\ . \quad (8)$$

Because the scalar-valued *diffusivity* $\Psi_V'\left(|\nabla u|^2\right)$ is a function of the unknown u, this PDE is nonlinear,

contrary to the linear PDEs that result from image-driven methods. A prominent example of isotropic disparity-driven regularisation is *Total Variation* [19] regularisation, used in [2, 21], where $\Psi_V = \sqrt{s^2 + \varepsilon^2}$.

3 PDE-Based Anisotropic Disparity-Driven Stereo

In [23] an anisotropic flow-driven regulariser for motion estimation was derived for the first time, but as we have seen, equivalent anisotropic disparity-driven ideas for variational stereo are still missing. However, such a smoothing strategy would have the favourable property that it allows smoothing along evolving disparity discontinuities, but not across. This can lead to the enhancement of meaningful edges, thus improving the estimation of discontinuities in the disparity field, without the problem of oversegmentation.

3.1 Adapting Anisotropic Flow-Driven Regularisation

Adapting the design ideas of Weickert and Schnörr [23] directly to our stereo setting results in the following regulariser: $V_{\mathrm{AD}}(\nabla u) = \mathrm{tr}\,\Psi_V(J)$, where Ψ_V is an increasing convex function and the argument $J := \nabla u \nabla u^\top$ is a symmetric, positive semidefinite 2×2 matrix. If J has the orthonormal eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$ with corresponding nonnegative eigenvalues λ_1 and λ_2, then $\Psi_V(J)$ is defined as the matrix with the eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$ and the eigenvalues $\Psi_V(\lambda_1)$ and $\Psi_V(\lambda_2)$:

$$J = \sum_{i=1}^{2} \lambda_i\, \mathbf{v}_i \mathbf{v}_i^\top$$

$$\Rightarrow \quad \Psi(J) := \sum_{i=1}^{2} \Psi_V(\lambda_i)\, \mathbf{v}_i \mathbf{v}_i^\top \ . \qquad (9)$$

Employing the regulariser $V_{\mathrm{AD}}(\nabla u)$ leads to the diffusion-reaction equation

$$u_t = \mathrm{div}\,\left(D(J)\,\nabla u \right) - \frac{1}{\alpha}\, m(f_l, f_r, u) \ , \quad (10)$$

with the diffusion tensor $D(J) := \Psi_V'(J)$. For anisotropic flow-driven optical flow, the argument J includes a coupling between the two flow components of the optical flow. In this manner the desired anisotropic behaviour is ensured because the eigenvectors of J are in general not parallel to the gradients of both flow components. In the stereo case, however, the eigenvalues and eigenvectors of J are trivial: $\lambda_1 = |\nabla u|^2, \lambda_2 = 0$ and $\mathbf{v}_1 = \frac{1}{|\nabla u|}\nabla u, \mathbf{v}_2 = \frac{1}{|\nabla u|}\nabla u^\perp$, where $\nabla u^\perp := (-u_y, u_x)^\top$ is a vector orthogonal to ∇u. With this the diffusion part of equation (10) comes down to

$$\mathrm{div}\,\left(D(J)\nabla u \right) \qquad (11)$$

$$= \quad \mathrm{div}\,\left(\Psi_V'(\nabla u \nabla u^\top)\nabla u \right) \qquad (12)$$

$$\overset{(9)}{=} \quad \mathrm{div}\left(\left[\frac{\Psi_V'(|\nabla u|^2)}{|\nabla u|^2}\nabla u \nabla u^\top \right. \right.$$
$$\left. \left. + \frac{\Psi_V'(0)}{|\nabla u|^2}\nabla u^\perp (\nabla u^\perp)^\top \right]\nabla u \right) \quad (13)$$

$$\overset{(*)}{=} \quad \mathrm{div}\,\left(\frac{\Psi_V'(|\nabla u|^2)}{|\nabla u|^2}|\nabla u|^2\nabla u + \mathbf{0} \right)(14)$$

$$= \quad \mathrm{div}\,\left(\Psi_V'(|\nabla u|^2)\nabla u \right) \ , \qquad (15)$$

where $(*)$ makes use of the facts that $\nabla u^\top \nabla u = |\nabla u|^2$ and $(\nabla u^\perp)^\top \nabla u = 0$. We conclude that for the stereo case, the use of the regulariser $V_{\mathrm{AD}}(\nabla u)$ yields the already presented disparity-driven isotropic behaviour of equation (8).

3.2 True Anisotropic Disparity-Driven Stereo

To finally model a highly adaptive anisotropic smoothing process for rectified stereo we will refrain from the design of a regulariser $V_{\mathrm{AD}}(\nabla u)$. In fact, we will directly model in the diffusion part of the diffusion-reaction equation (10).

In order to obtain truly anisotropic behaviour we need a more sophisticated structure detector than J. Inspired by the anisotropic diffusion filter from [22], we consider the *structure tensor* J_ρ [6] for stereo:

$$J_\rho := J_\rho(\nabla u_\sigma) := K_\rho * \left(\nabla u_\sigma \nabla u_\sigma^\top \right) \ , \quad (16)$$

where $u_\sigma := K_\sigma * u$, K_σ denotes a Gaussian kernel of standard deviation σ and $*$ is the convolution operator. We see that J_ρ extends J in two ways: (i) It regularises the disparity u by a Gaussian convolution of standard deviation σ and (ii) integrates

neighbourhood information by convolving the tensor entries with a Gaussian kernel of standard deviation ρ. Regularisation of the unknown u by Gaussian convolution with a *noise scale* σ was first proposed in the context of nonlinear diffusion to reduce staircaising artifacts and problems with noise, c.f. [5]. Despite the fact that ∇u_σ is a useful edge detector, the problem still remains that it is sensitive under noise for small σ, while an increased σ can lead to undesired cancellation effects. This can be overcome by an additional convolution of the tensor entries with an *integration scale* ρ.

The structure tensor J_ρ is a symmetric, positive semidefinite matrix with two orthonormal eigenvectors $\mathbf{w}_1, \mathbf{w}_2$, which give the directions of the local disparity structure. The corresponding non-negative eigenvalues, w.l.o.g. $\mu_1 \geq \mu_2 \geq 0$, give the average contrast along these directions. So we propose the following diffusion-reaction equation which makes use of the structural information contained in J_ρ:

$$u_t = \mathrm{div}\,\left(D\left(J_\rho\right)\,\nabla u\right) - \frac{1}{\alpha}\,m(f_l, f_r, u)\ ,\ \ (17)$$

with the diffusion tensor

$$D(J_\rho) := \Psi'_V(J_\rho) := \sum_{i=1}^{2} \Psi'_V(\mu_i)\,\mathbf{w}_i\,\mathbf{w}_i^\top\ .$$

$$(18)$$

We further propose to make use of the Perona-Malik [18] diffusivity with a contrast parameter $\tilde{\varepsilon} > 0$:

$$\Psi'_V(s^2) := \frac{1}{1 + s^2/\tilde{\varepsilon}^2}\ .\qquad (19)$$

This diffusivity is known to make backward diffusion possible and thereby enhance edges even more. With this choice we can now show that our method exhibits the described anisotropic behaviour:

– In *flat regions:*
$\mu_1 \approx \mu_2 \approx 0 \Rightarrow \Psi'_V(\mu_1) \approx 1,\ \Psi'_V(\mu_2) \approx 1$, which leads to homogeneous smoothing in both directions.
– At a *straight edge in w_2-direction:*
$\mu_1 \gg \mu_2 \approx 0 \Rightarrow \Psi'_V(\mu_1) \approx 0,\ \Psi'_V(\mu_2) \approx 1$, which leads to anisotropic smoothing in edge direction, but not across.
– At *corners:*
$\mu_1 \geq \mu_2 \gg 0 \Rightarrow \Psi'_V(\mu_1) \approx 0,\ \Psi'_V(\mu_2) \approx 0$, which prevents smoothing.

4   Numerical Solution of the PDE

What needs to be mentioned is how to solve the diffusion-reaction equation (17) in its steady-state where $u_t = 0$. As is proposed in [3], we use a coarse-to-fine warping approach. This multiscale approach is achieved by using a downsampling of the image pair by a factor $\eta \in (0, 1)$, yielding $[L, \ldots, 0]$ warping level, depending on the image size and η. On each level, we compute disparity increments via a linearised approach that is applicable because the increments are usually small. This strategy allows to handle large disparities correctly. Moreover, due to the PDE-based nature of our approach, we can speed up the computation by following the idea of [4] and using a nonlinear multigrid scheme to solve the problem at each warping level. On each grid level, we apply a Gauss-Seidel solver with alternating line relaxation to the resulting linear system of equations. Occurring spatial derivatives of the image data are approximated by central finite differences of fourth order and spatial derivatives of the disparity by second order approximations.

5   Experiments

We evaluate our presented PDE-based anisotropic disparity-driven stereo method against the graph cuts approach of Kolmogorov and Zabih [13] (available for download at www.cs.cornell.edu/~rdz/graphcuts.html) and the isotropic disparity-driven method of Slesareva *et al.* [21], adapted to the rectified stereo case. This is achieved by using the trivial fundamental matrix, which yields a horizontal epipolar direction. Furthermore, we made use of the mentioned multigrid solver [4], i.e., our approach just replaces the isotropic disparity-driven regularisation of [21] by our new anisotropic disparity-driven method. However, we will see that this may give drastic improvements.

To reduce the amount of parameters to be estimated for our method, we choose some standard settings for our experiments: A coarsening factor $\eta = 0.95$ for the multiscale approaches and regularisation parameters $\varepsilon = 0.001,\ \tilde{\varepsilon} = 0.1$. For our anisotropic method using the structure tensor, we estimate a value for σ and set $\rho := 2\sigma$.

For our first experiment, we tested the three approaches on a grey value version of the 'Plastic' image pair from the Middlebury stereo page (vision.middlebury.edu/stereo), which is shown together with the grey value coded ground truth disparity in the top row of Figure 1. To make a quantitative analysis of results possible, we employ two different error measures. They reflect how well a disparity estimate $\mathbf{u} = (\mathbf{u}_i)$ matches the given ground truth $\mathbf{u}^{\mathrm{gt}} = (\mathbf{u}_i^{\mathrm{gt}})$, for images with $i = 1, \ldots, N$ pixels. The first measure is the *average absolute disparity error* (AADE) of [21] and the second one is the *bad pixel error* (BPE) of [20], which gives the percentage of pixels which deviate more than a threshold $\delta_d > 0$ from the ground truth. These measures are defined as follows:

$$\mathbf{AADE}\,(\mathbf{u}, \mathbf{u}^{\mathrm{gt}}) = \frac{1}{N} \sum_{i=1}^{N} \left| \mathbf{u}_i - \mathbf{u}_i^{\mathrm{gt}} \right|, \qquad (20)$$

$$\mathbf{BPE}\,(\mathbf{u}, \mathbf{u}^{\mathrm{gt}}) = \frac{100}{N} \sum_{i=1}^{N} \mathbf{T}\left(\left| \mathbf{u}_i - \mathbf{u}_i^{\mathrm{gt}} \right| > \delta_d \right), \quad (21)$$

where $\mathbf{T}(b) = 1$ if $b = \mathrm{true}$, and 0 else. As proposed in [20], we set $\delta_d = 1$.

The achieved results and colour-coded error maps (green $\equiv$ error $< \delta_d$, yellow $\equiv \delta_d \leq$ error $< 3\delta_d$ and red $\equiv$ error $\geq 3\delta_d$) for the three methods can be found in the middle and lower row of Figure 1. In Table 1, we collected the corresponding error measures and computation times, also for the 'Teddy' pair that we will present in Figure 2 and other Middlebury pairs. For the latter we do not give disparity estimates due to space limitations.

Concerning the results for the 'Plastic' pair, c.f., Figure 1 and Table 1, one sees that due to the piecewise smooth ground truth the variational approach of Slesareva *et al.* and also our PDE-based method easily outperforms the GC approach. The mentioned drawback of the strict regulariser used in the GC approach becomes obvious: The smoothly varying disparity of the folder in the foreground is not recovered well, which one impressively sees in the corresponding error maps. In addition it becomes clear that our new anisotropic disparity-driven method brings quite some benefits compared to its isotropic counterpart. This can mainly be seen in the much better estimation of the background in the upper right part and the folder in the foreground. The improvements are most striking in regions where there are strong edges in the disparity

Table 1: Error measures (AADE, BPE) and computation times for experiments of Figures 1, 2 and others. Experiments were conducted on a standard PC (3.2 GHz Intel Pentium 4, 256 MB RAM). For 'Teddy' only the non-occluded regions were evaluated in the error measures, for the rest only the reliable regions.

Pair Max. disp.			GC [13]	Isotropic [21]	Our method
Plastic 66	AADE BPE Time [s]		7.60 57.13 190.25	1.21 24.37 9.51	1.37 19.45 23.82
Teddy 59	AADE BPE Time [s]		1.49 13.46 106.08	0.64 10.37 10.39	0.61 9.22 21.61
Laundry 78	AADE BPE Time [s]		6.19 35.48 133.59	3.22 37.18 11.34	2.95 34.25 21.69
Bowling1 77	AADE BPE Time [s]		4.79 53.41 204.91	4.63 30.35 9.40	3.36 24.41 20.26

field, as can be expected for our anisotropic method. Regarding the BPE, our method gives an improvement of about 20% compared to the method of Slesareva *et al.* and even 65% compared to the GC approach. If we evaluate the given computation times, we see that the more complex anisotropic method leads to an average increase of about 100% compared to the isotropic method with multigrid. However, the GC approach is still far behind, especially for pairs with large disparities. For the 'Plastic' pair the increase in computation time is about 800% compared to our approach and even 2000% compared to the method of Slesareva *et al.*, which impressively shows the efficiency of the employed multigrid solver.

As a second experiment, we compared our results for 'Teddy', c.f., Figure 2 and Table 1, with the official ranking of the Middlebury page for $\delta_d = 0.5$. With the isotropic method of Slesareva *et al.* one currently obtains rank 11 out of 46, which we can improve to rank 8 with our new anisotropic method. As can be seen in the error maps of Figure 2, the main improvements of our method lie in the better estimation of the floor in the lower part of the image. Small improvements are also visible at the right side of the teddy and at the back of the stuffed animal on the floor. However, another insight of this experiment is that some very recent
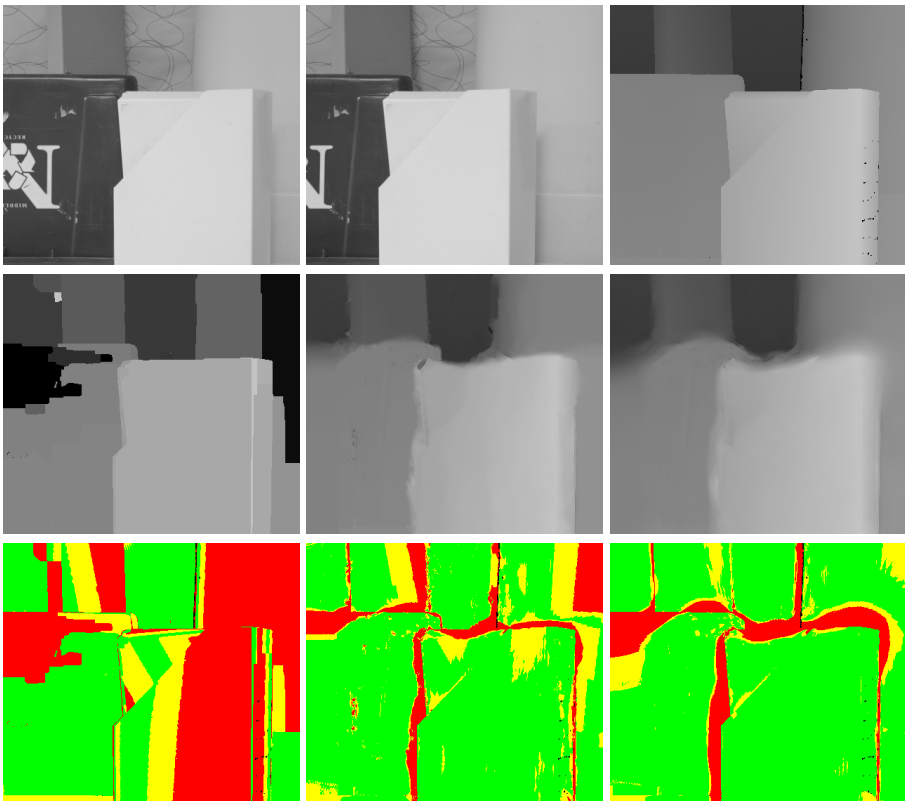
Figure 1: **First row, from left to right:** Left image of *'Plastic'* pair (423×370 pixels). Right image. Ground truth disparity magnitude, non-reliable pixels are marked in black. **Second row, from left to right:** Disparity magnitude for GC approach [13] ($\lambda = 10$, automatically estimated). Same for rectified stereo version of Slesareva *et al.* [21] ($\alpha = 7, \sigma_{pre} = 0.35, \gamma = 60, L = 93$). Same for our method ($\alpha = 90, \sigma_{pre} = 0.45, \gamma = 100, \sigma = 4.5, \rho = 9, L = 93$). **Third row, from left to right:** Error map for GC approach. Same for rectified stereo version of Slesareva *et al.* [21]. Same for our method.

probabilistic approaches are still able to outperform variational or PDE-based approaches, even on test pairs with piecewise smooth ground truth. This can be explained by the more sophisticated model assumptions made in these approaches, like explicit occlusion handling [13].

As a third experiment we reconstructed the *'Portal'* scene (available for download at cmp.felk.cvut.cz/~cechj/GCS), using the estimated disparities as hightfields. The scene is part of a larger set of rectified real-world scenes, collected by Jan Cech and Radim Sara. This specific scene, c.f., top row of Figure 3, shows the portal of a church with many details around the

door and on the arch. The estimated disparity magnitudes for the GC approach and for our method are also given. They were used in the reconstructions depicted in the bottom row of Figure 3. One clearly sees that the reconstruction with the GC approach is not satisfactory. All smoothly slanted surfaces are estimated in a stair-like manner, originating from the strict regularisation. One furthermore experiences unpleasant outliers at the right border. Our method solves these problems: We get a very accurate reconstruction, with sharp discontinuities and lots of fine details, e.g., the frets at the top of the portal and even the door handle are estimated well. Concerning the computation time, the GC approach
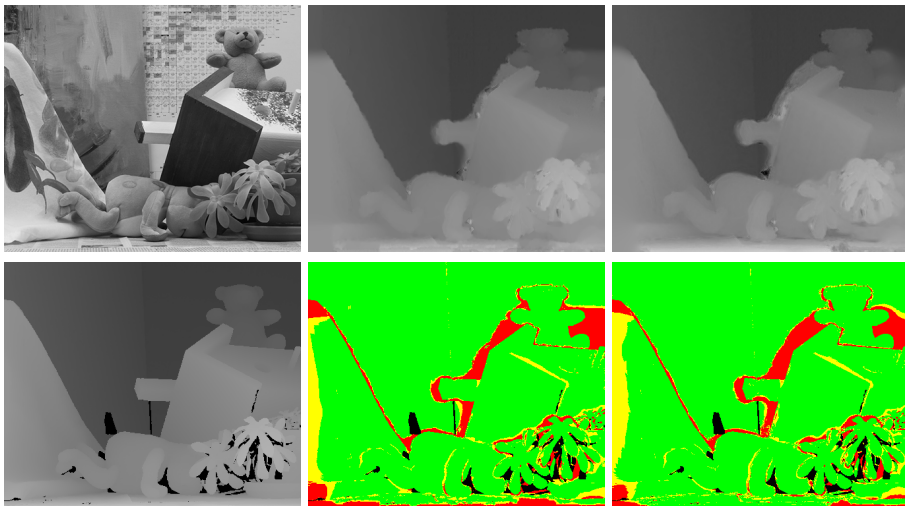
Figure 2: **First row, from left to right:** Left image of *'Teddy'* pair (450×375 pixels). Disparity magnitude for rectified stereo version of Slesareva *et al.* [21] ($\alpha = 5.5, \sigma_{pre} = 0.5, \gamma = 7.5, L = 94$). Same for our method ($\alpha = 20, \sigma_{pre} = 0.45, \gamma = 5.5, \sigma = 2.5, \rho = 5, L = 94$). **Second row, from left to right:** Ground truth disparity magnitude, non-reliable pixels are marked in black. Error map for rectified stereo version of Slesareva *et al.* [21]. Same for our method.

needed 199.96 s for the disparity estimation using 35 discrete depth levels, whereas our method only needed 33.27 s.

6 Conclusions and Outlook

In this paper, we filled the gap in existing smoothing strategies for stereo vision. We have first shown that a straight-forward adaptation of anisotropic ideas from optical flow computations [23] does not work for stereo, as the smoothing process remains isotropic. As a remedy, we presented a novel PDE-based anisotropic disparity-driven method, based on anisotropic diffusion filters. Our experiments clearly show that such a method can help to considerably improve the results compared to previous isotropic approaches, such as [21]. This again demonstrates that it pays off to replace existing isotropic approaches by the additional degrees of freedom that come from anisotropy. Comparing to very recent probabilistic approaches, we have seen that our method is indeed competitive as we are ranking among the best 20% of all featured methods in the official Middlebury ranking. Furthermore, the

application of highly efficient multigrid schemes [4] is still possible, resulting in moderate run times in the order of a few seconds for standard test images. This is in general much less than the computation times for the tested GC approach [13].

It is evident that our method still leaves space for some improvements. If one takes a closer look at the error maps in Figures 1 and 2, one realizes that errors mostly occur at occluded regions, e.g., at the left border of the folder in Figure 1 or at the left border of the house in Figure 2. In [2], the authors present a variational approach with explicit occlusion handling, which gives favourable results at occlusions. Incorporating such concepts, we aim to develop a PDE-based approach of even better quality.
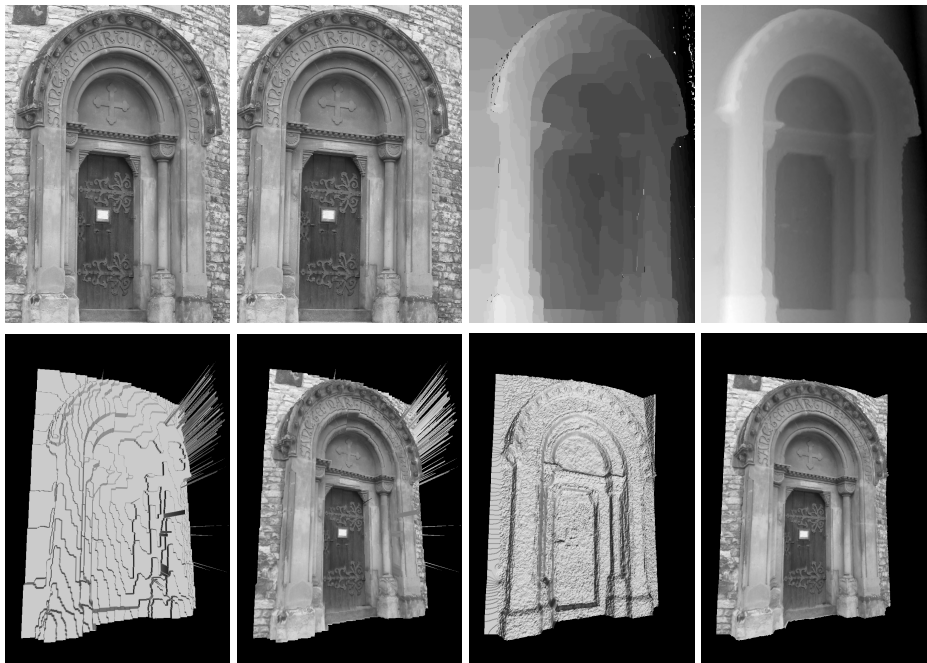
Figure 3: **First row, from left to right:** Left image of the *'Portal'* image pair (greyscale version, cropped and resized to 435×615 pixels to remove a black border stemming from the rectification). Right image. Disparity magnitude for GC approach [13] ($\lambda = 6$, automatically estimated). Same for our method ($\alpha = 40, \sigma_{pre} = 0.5, \gamma = 3, \sigma = 2.5, \rho = 5, L = 97$). **Second row, from left to right:** Reconstruction using GC approach. Same with texture mapping. Reconstruction using our method. Same with texture mapping.

References

[1] L. Alvarez, R. Deriche, J. Sánchez and J. Weickert. Dense disparity map estimation respecting image derivatives: a PDE and scale-space based approach. *Journal of Visual Communication and Image Representation,* 13, 3–21, 2002.

[2] R.B. Ari and N.A. Sochen. Variational stereo vision with sharp discontinuities and occlusion handling. In *Proceedings of the 2007 IEEE International Conference on Computer Vision,* Rio de Janeiro, Brazil, IEEE Computer Society Press, 1–7, 2007.

[3] T. Brox, A. Bruhn, N. Papenberg and J. Weickert. High accuracy optical flow estimation based on a theory for warping. In T. Pajdla, J. Matas, eds.: *Computer Vision – ECCV 2004,* Part IV. Volume 3024 of Lecture Notes in Computer Science. Springer, Berlin, 25–36, 2004.

[4] A. Bruhn, J. Weickert, T. Kohlberger and C. Schnörr. A multigrid platform for real-time motion computation with discontinuity-preserving variational methods. *International Journal of Computer Vision,* 70, 257–277, 2006.

[5] 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.

[6] W. Förstner and E. Gülch. A fast operator for detection and precise location of distinct points, corners and centres of circular features. In *Proceedings of the ISPRS Intercommission Conference on Fast Processing of Photogram-*

metric Data, Interlaken, Switzerland, 281–305, 1987.

[7] T. Froehlinghaus and J. Buhmann. Regularizing phase based stereo. In *International Conference on Pattern Recognition,* Part I. 451–455, 1996.

[8] I. Galić, J. Weickert, M. Welk, A. Bruhn, A. Belyaev and H.P. Seidel. Towards PDE-based image compression. In N. Paragios, O. Faugeras, T. Chan and C. Schnörr, eds.: *Variational, Geometric and Level-Set Methods in Computer Vision.* Volume 3752 of Lecture Notes in Computer Science. Springer, Berlin, 37–48, 2005.

[9] W.E.L. Grimson. Computational experiments with a feature based stereo algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 7, 17–34, 1985.

[10] B. Horn and B. Schunck. Determining optical flow. *Artificial Intelligence,* 17, 185–203, 1981.

[11] H. Kim and K. Sohn. Hierarchical disparity estimation with energy-based regularisation. In *Proceedings of the IEEE International Conference on Image Processing.* 373–376, 2003.

[12] A. Klaus, M. Sormann and K. Karner. Segment-based stereo matching using belief propagation and a self-adapting dissimilarity measure. In *Proceedings of the 18th International Conference on Pattern Recognition,* Part III. 15–18, 2006.

[13] V. Kolmogorov and R. Zabih. Multi-camera scene reconstruction via graph cuts. In A. Heyden, G. Sparr, M. Nielsen and P. Johansen, eds.: *Computer Vision - ECCV 2002,* Part III. Volume 2352 of Lecture Notes in Computer Science. Springer, 82–96, 2002.

[14] C. Lei, J. Selzer, and Y.-H. Yang. Region-tree based stereo using dynamic programming optimisation. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition.* 2378–2385, 2006.

[15] G. Li and S.W. Zucker. Differential geometric consistency extends stereo to curved surfaces. In A. Leonardis, H. Bischof and A. Pinz, eds.: *Computer Vision - ECCV 2006,* Part III. Volume 3953 of Lecture Notes in Computer Science, Springer, 44–57, 2006.

[16] A. Mansouri, A. Mitiche and J. Konrad. Selective image diffusion: application to disparity estimation. In *Proceedings of the 1998 IEEE International Conference on Image Processing.* Volume 3., Chicago, IL, 284–288, 1998.

[17] D. Marr and T. Poggio. Cooperative computation of stereo disparity. *Science,* 194, 283–287, 1976.

[18] P. Perona and J. Malik. Scale space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 12, 629–639, 1990.

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

[20] D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision,* 47, 7–42, 2002.

[21] N. Slesareva, A. Bruhn and J. Weickert. Optic flow goes stereo: A variational method for estimating discontinuity-preserving dense disparity maps. In W. Kropatsch, R. Sablatnig and A. Hanbury, eds.: *Pattern Recognition.* Volume 3663 of Lecture Notes in Computer Science, Springer, Berlin, 33–40, 2005.

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

[23] J. Weickert and C. Schnörr. A theoretical framework for convex regularisers in PDE-based computation of image motion. *International Journal of Computer Vision,* 45, 245–264, 2001.