Mumford-Shah Meets Stereo: Integration of Weak Depth Hypotheses

Thomas Pock
Institute for Computer Graphics and Vision
Graz, University of Technology
Inffeldgasse 16, A-8010 Graz, Austria
pock@icg.tugraz.at

Christopher Zach
VRVis Research Center
Inffeldgasse 16, A-8010 Graz, Austria
zach@vrvis.at

Horst Bischof
Institute for Computer Graphics and Vision
Graz, University of Technology
Inffeldgasse 16, A-8010 Graz, Austria
bischof@icg.tugraz.at

Abstract

Recent results on stereo indicate that an accurate segmentation is crucial for obtaining faithful depth maps. Variational methods have successfully been applied to both image segmentation and computational stereo. In this paper we propose a combination in a unified framework. In particular, we use a Mumford-Shah-like functional to compute a piecewise smooth depth map of a stereo pair. Our approach has two novel features: First, the regularization term of the functional combines edge information obtained from the color segmentation with flow-driven depth discontinuities emerging during the optimization procedure. Second, we propose a robust data term which adaptively selects the best matches obtained from different weak stereo algorithms. We integrate these features in a theoretically consistent framework. The final depth map is the minimizer of the energy functional, which can be solved by the associated functional derivatives. The underlying numerical scheme allows an efficient implementation on modern graphics hardware. We illustrate the performance of our algorithm using the Middlebury database as well as on real imagery.

1. Introduction

In the last 20 years, the computation of visual correspondences in one or more stereo image pairs has been a challenging task for the vision community. Many different algorithms have been proposed in order to cope with typical stereo problems such as large untextured areas, occlusions and varying radiance.

In order to get faithful depth maps one has to find a model which can deal with both smooth objects and sharp depth discontinuities. On the other hand depth discontinuities often coincide with the edges of the image.

In their celebrated paper [17], Mumford and Shah proposed a variational segmentation model which intents to decompose an image into distinct region using piecewise smooth functions. Originally the model was developed for image segmentation but the concept of piecewise smooth functions also fits perfectly to depth maps.

The contribution of this paper is to unify the segmentation and depth processes in a theoretically consistent variational framework. More precisely, we propose a Mumford-Shah like functional which uses a common discontinuity set for both the image intensity function and the depth map. Weak edges at depth discontinuities will therefore be enforced as well as the image edges will mainly influence the depth map. In addition, we propose a robust data term which is insensitive to outliers and can deal with multiple depth hypotheses obtanied from different matching algorithms. The final depth map is computed by minimizing the energy functional. Therefore, we develop a simple and robust fixed point algorithm (merely two update equations have to be implemented) which can be computed within a few seconds on modern graphics hardware.

The structure of the paper is as follows. Section 2 reviews some recent developments of stereo algorithms, introduces the Mumford-Shah segmentation model and discusses different approximations of the Mumford-Shah functional. In Section 3 we develop the joint image and depth segmentation functional and we show how to robustly integrate multiple depth hypotheses. In Section 4 we propose a simple algorithm to compute the depth map. We also give some details concerning implementation issues.

In Section 5 we evaluate the performance of our method using the Middlebury database. In addition, we apply our method to challenging real world data sets showing the excellent performance of our method. In the last section we give some conclusions and suggest possible directions for future investigations.

2. Related Work

2.1. Computational Stereo

Calculating depth maps from a set of images is today still an active and challenging research topic. According to the well-known standard benchmark suite for computational stereo [21, 22], most of the high-quality methods are based on global 2D Markov random field optimization approaches like graph cuts [4] or loopy belief propagation [28]. The top-ranked methods include explicit treatment of potentially occluded regions and, often more importantly, a reasoning mechanism to detect image edges or segments reflecting depth discontinuities (e.g. [13, 27, 35]). Most of these methods use simple image similarity measures, often based only on (sampling invariant) single pixel comparisons [2]. The reason for their good performance on the benchmark dataset lies in their sophisticated depth extraction procedure incorporating multiple stages.

The approach proposed in [36] takes a very different route: a simple winner-takes-all depth extraction method is combined with an advanced image similarity score. The utilized large aggregation window yields very discriminative descriptors for image similarity. The foreground fattening effect is avoided by incorporating a weighted support approach, which implicitly leads to aggregation windows with irregular, context dependent shape. The application of the adaptive window shapes substantially improves the result near depth discontinuities. Different approaches using similar ideas are presented in [3] and [32].

Another line of methods aiming at increasing the accuracy near depth discontinuities without reducing the performance in smooth regions uses variable-sized support windows for the similarity score [14, 33]. Conceptually, a number of disparity hypotheses for every pixel is generated, and the final disparity value is selected from this hypotheses. In [33], the determination of the correct window size is highly coupled with the depth extraction procedure for highest efficiency.

Most of the computational stereo methods employ an optimization framework suited for Markov random field problems. There are only a few approaches based on variational principles. Several methods are based on optical flow approaches incorporating the epipolar constraint to reduces the search space (e.g. [19, 23, 25, 26]). These methods evolve a depth map in order to minimize an energy functional. One drawback of these approaches is the necessity

to provide derivatives of the similarity function. A suitable regularization scheme allows depth discontinuities based on image gradients. In [20] the probabilities of the disparity hypotheses for every pixel are evolved using non-linear diffusion methods, hence it is closer to the Markov random field formulations.

2.2. Variational Image Segmentation

Besides the computation of visual correspondence, another fundamental low level tasks is to segment the image domain into distinct surface patches belonging to distinct objects. A variety of ways to define the task of segmentation have been proposed e.g. [6, 10, 17], but what is common to these algorithms is that they try to minimize the same segmentation energy [15]. The calculus of variations provides a uniform framework where energy minimization finds a precise language by means of variational principles.

In their seminal paper [17] Mumford and Shah proposed a segmentation model which is based on piecewise smooth approximations of the intensity function. The model has successfully been utilized for several applications e.g. active contour models [9], edge detection [5], image regularization [31], image decomposition [24], inpainting [12] and registration [11]. Moreover, the Mumford-Shah functional has strong connections to concepts such as statistical estimation via maximum likelihood [8]. Its neuronal plausibility has been discussed in [18]. This emphasizes the universality of the Mumford-Shah functional for computer vision applications.

The original Mumford-Shah (MS) segmentation model is defined as

$$E_{\mathrm{MS}} = \int_{\Omega} (u - g)^2 \, \mathrm{d}\Omega + \alpha \int_{\Omega \backslash \Gamma} |\nabla u|^2 \mathrm{d}\Omega + \beta \, \mathrm{length}(\Gamma) \;,$$

where g is the observed image, u is a piecewise smooth approximation, Γ is a set containing edges in u, α and β are tuning parameters. It was shown that there exist a simple relation between (α, β) and $(scale, contrast)$ [17], in fact

$$scale = \sqrt{\alpha} \;, \quad contrast = \sqrt{2\beta/\sqrt{\alpha}} \;.$$

In its original setting, the Mumford-Shah functional is hard to minimize, due to the lack of convexity and regularity of the edge-length term. A first solution was given by Ambrosio and Tortorelli [1] by approximating the original functional by simpler elliptic variational problems. They proposed to replace the edge set Γ by means of a 2D function z and designed the so-called phase field energy $E_{z,\varepsilon}$ which additionally depends upon a scale parameter ε. The remarkable property associated with this formulation is that as $\varepsilon \to 0$, $E_{z,\varepsilon}$ approaches the length of Γ. However, this approximation works well only if the scale of ε is in the order of one pixel.

In [7] Chambolle presented a different approach based on a non-local approximation of the Mumford-Shah functional. The major advantage of this formulation is that the explicit computation of the edge set Γ is avoided by using a family of continuous and non-decreasing functions $f : [0, +\infty) \to [0, +\infty)$ satisfying

$$\lim_{t \to 0^+} \frac{f(t)}{t} = 1 \; , \; \lim_{t \to +\infty} f(t) = 1 \; .$$

While Chambolle uses functions of the form $f(t) = \arctan(t)$, it was later shown that $f(t) = log(1 + t)$ is a better choice, since it is less sensitive to local minima and requires a smaller number of iterations to converge [16]. In the discrete setting, the non-local approximation of the Mumford-Shah functional yields

$$
\begin{aligned}
D_{\mathrm{MS},\varepsilon} \;=\; & \varepsilon^2 \sum_{\boldsymbol{x} \in \Omega \cap \mathbb{Z}^2} |u(x) - g(x)|^2 \\
& + \; \varepsilon^2 \sum_{\boldsymbol{x} \in \Omega \cap \mathbb{Z}^2} \sum_{\xi \in \mathcal{N}(x)} \mathcal{A}_\xi \log\left(1 + \mathcal{B}_\xi |\nabla_\xi u(x)|^2\right) ,
\end{aligned}
\tag{1}
$$

where

$$\nabla_\xi u(x) = u(x + \varepsilon \xi) - u(x), \; \mathcal{A}_\xi = \beta \frac{\rho(\xi)}{a_\varepsilon |\xi|}, \; \mathcal{B}_\xi = \frac{\alpha}{\beta} \frac{a_\varepsilon}{|\xi| \varepsilon^2} \; ,$$

$a_\varepsilon = \varepsilon \log \frac{1}{\varepsilon}$, the convolution term $\rho(\xi)$ takes constantly the value $(\sqrt{2} - 1)/2$ and $\mathcal{N}(x)$ is the set of nearest neighbors such that $|\xi|_\infty = 1$. For further information we refer to [16].

3. Our Approach

In this section we describe the extension of the non-local approximation of the Mumford-Shah functional (1) to a joint color-depth segmentation functional. We propose a robust data term which can deal with multiple depth hypotheses. For minimization we derive the functional derivatives (Euler-Lagrange equations in the continuous setting).

3.1. Joint Color-Depth Segmentation

A color image can be described by three channels (e.g. RGB, LAB). Thus, the most obvious way is to extend the scalar valued variables u and g of (1) to vector valued variables $\mathbf{u} = (u_1, u_2, u_3)^T$ and $\mathbf{g} = (g_1, g_2, g_3)^T$. We denote the squared L_2 norm of a vector $\mathbf{x}$ as $\|\mathbf{x}\|_2^2$. We point out that more sophisticated norms can be used which may improve the segmentation results [5].

As mentioned in the introduction, our major aim is to unify the discontinuities of the color segmentation process with those of the depth process. The simplest extension is to treat the depth map in the same way as the color image. Therefore, we introduce a piecewise smooth depth map d

and an initial depth map d_0. The initial depth map could be provided by any local matching algorithm (e.g. using a simple correlation window).

We introduce two additional parameters. The first parameter $\gamma \in (0, 1)$ weights the influence of the color and depth gradients to the common discontinuity set. If $\gamma \to 1$ the discontinuities are mainly influenced by the color gradients. If $\gamma \to 0$ the discontinuities are mainly affected by the depth discontinuities. We found that $\gamma = 0.9$ yields good results in most cases. The second parameter δ is used to control the smoothness of the depth map.

In this setting, the discrete energy enabling joint color and depth segmentation is given by:

$$S_\varepsilon = \varepsilon^2 \sum_{x \in \Omega \cap \mathbb{Z}^2} \Psi(x) \; + \; \varepsilon^2 \sum_{x \in \Omega \cap \mathbb{Z}^2} \Phi(x) \, , \tag{2}$$

where the data term $\Psi(x)$ is given by

$$\Psi(x) = \gamma \, \|\mathbf{u}(x) - \mathbf{g}(x)\|_2^2 \; + \; (1 - \gamma)\delta \, |d(x) - d_0(x)|^2 \; . \tag{3}$$

The regularization term $\Phi(x)$ is given by

$$\Phi(x) = \sum_{\xi \in \mathcal{N}(x)} \mathcal{A}_\xi \log\left(1 + \mathcal{B}_\xi \mathcal{G}(x, \xi)\right) , \tag{4}$$

where joint color depth gradients are given by

$$\mathcal{G}(x, \xi) = \gamma \|\nabla_\xi \mathbf{u}(x)\|_2^2 \; + \; (1 - \gamma)|\nabla_\xi d(x)|^2 \; .$$

Optimizing this energy results in piecewise smooth color and depth images.

Fig. 1 shows the optimization result of our energy (2) applied to a stereo pair of the Middlebury database [22]. The initial depth map was obtained from a *winner takes all* strategy applied to a 3×3 sum of absolute intensity values (SAD) correlation window. As expected, the resulting approximations of the color image and the depth map yield piecewise smooth functions. One can also see that the discontinuities of the color image well correspond to the discontinuities of the depth map. Nevertheless, in areas of many wrong initial matches, our first model does not yield good results. The error norm measuring the deviations from the initial depth map is a quadratic one which does not allow for outliers. In the next section we show how to resolve this problem by means of a robust error norm.

3.2. Selection from Multiple Depth Hypotheses

If a set of n initial depth maps $d_0^{(i)}$ is provided, the depth segmentation procedure can be extended by simply summing up the according data terms. We presume, that the initial depth maps will often disagree, in particular near to depth discontinuities and occlusions. Consequently, the deviation of the depth process d from the initial maps
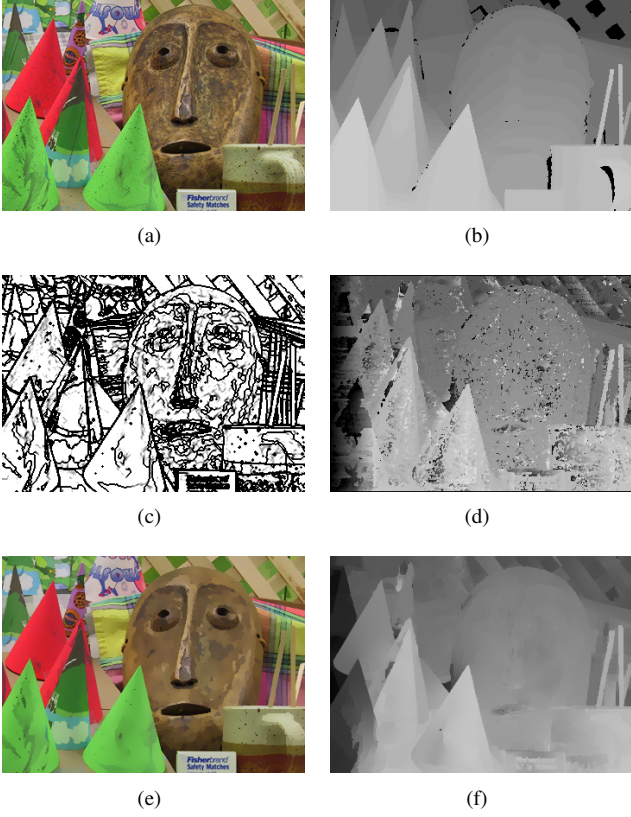
Figure 1. Optimization result of S_ε using the quadratic data term. (a) Left input image (b) Ground truth. (c) Joint color-depth edge set (black pixels indicate strong edges). (d) Initial Match obtained from 3×3 SAD. (e) Piecewise smooth approximation of the left input image. (f) Piecewise smooth approximation of the depth map.

$d_0^{(i)}$ must be quantified in a robust manner, and the term $|d(x) - d_0(x)|^2$ in (3) needs a suitable adaption. We select the robust distance function $\phi(s) = \frac{s^2}{1+s^2}$ [30]. Hence, we replace (3) by the robust data term

$$\Psi(x) = \gamma \|\mathbf{u}(x) - \mathbf{g}(x)\|_2^2 \qquad (5)$$
$$+ (1-\gamma)\,\delta \sum_{i=1}^{n} \frac{\left|d(x) - d_0^{(i)}(x)\right|^2}{1 + \left|d(x) - d_0^{(i)}(x)\right|^2} .$$

Fig. 2 shows the optimization result for the previous example but now using the robust data term. We used two initial depth hypotheses. The first one is the same as in the previous example. The second one is obtained from a 3×3 sum of absolute differences of gradients (GRAD) correlation window, which is more invariant to illumination changes. One can see that the robust data term substantially improves the results. The reason for the improvement is twofold. First, the robust distance function allows for outliers. Indeed, Fig. 2(c) and Fig. 2(d) show the implicitly detected outliers of the robust error norm. The outliers cor-

respond to black pixels. Second, using multiple hypotheses allows the algorithm to select the most confident matches.
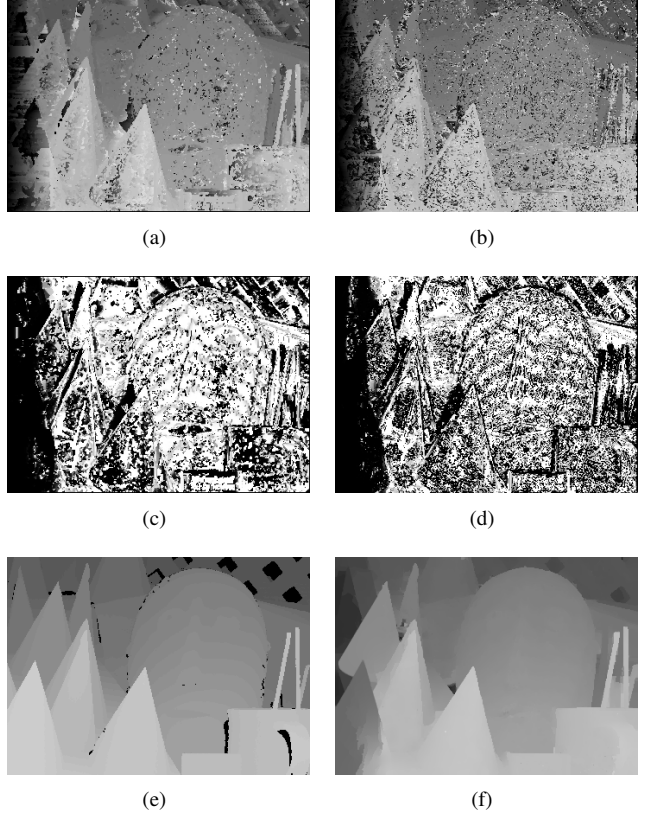

Figure 2. Optimization result using the robust data term of two depth hypotheses. (a) Initial depth map obtained from 3×3 SAD. (b) Initial depth map obtained from 3×3 GRAD. (c) Outlier detected in SAD matches. (d) Outlier detected in GRAD matches. (e) True depth map. (f) Depth map obtained from joint depth and color segmentation.

3.3. Minimization of the Energy

It is well known that discrete energies such as (2) can be minimized by finding the zeros of the functional derivatives. Since the energy S_ε is highly non-convex, only a *good* local minimum can be computed.

Taking the functional derivatives of (2) with respect to $\mathbf{u}$ and d one arrives at

$$\partial_t \mathbf{u}(x) = (\mathbf{u}(x) - \mathbf{g}(x)) \qquad (6)$$
$$- \sum_{\xi \in \mathcal{N}(x)} \mu_\xi(x) \nabla_\xi \mathbf{u}(x)$$

and

$$\partial_t d(x) = \sum_{i=1}^{n} \nu_i(x) \left(d(x) - d_0^{(i)}(x) \right) \qquad (7)$$
$$- \sum_{\xi \in \mathcal{N}(x)} \mu_\xi(x) \nabla_\xi d(x) ,$$

where the diffusion weights are given by

$$\mu_\xi(x) = \frac{\mathcal{A}_\xi \mathcal{B}_\xi}{1 + \mathcal{B}_\xi \mathcal{G}(x, \xi)}$$

and the weights derived from the robust data term

$$\nu_i(x) = \frac{\delta}{\left(1 + \left| d(x) - d_0^{(i)}(x) \right|^2 \right)^2} .$$

4. Implementation

Based on the functional derivatives (6) and (7) we apply a semi-implicit linearization technique and arrive at the following simple update equations.

$$\mathbf{u}^{k+1}(x) = \frac{\mathbf{g}(x) + \sum_\xi \mu_\xi^k(x) \, \mathbf{u}^k(x + \varepsilon\xi)}{1 + \sum_\xi \mu_\xi^k(x)} \qquad (8)$$

and

$$d^{k+1}(x) = \frac{\sum_i \nu_i^k(x) \, d_0^{(i)}(x) + \sum_\xi \mu_\xi^k(x) \, d^k(x + \varepsilon\xi)}{\sum_i \nu_i^k(x) + \sum_\xi \mu_\xi^k(x)} . \qquad (9)$$

These equations can be solved alternately using simple fixed point iterations. Note that at each iteration $k + 1$ the update equations solve a linear diffusion equation whose nonlinear coefficients $\mu_\xi(x)$ and $\nu_i(x)$ depend on the previous iterate k. This scheme is commonly known as *lagged diffusivity fixed point iteration* and was introduced in [34] by Vogel and Oman for total variation denoising. The scheme works robustly and almost linear convergence is achieved.

Our minimization procedure is as follows. The initial depth maps $d_0^{(i)}$ are computed using a weak matching algorithm. The color image $\mathbf{u}$ is initialized with the source image $\mathbf{g}$. In our examples we used the LAB color space which has been designed to correlate with human color discrimination performance. The initial depth map d is set to the median of the single depth hypotheses. Finally the fixed point algorithm is iterated until convergence.

Instead of adjusting the parameters α and β directly, we specify *scale* and *contrast* and use the relations (1) for conversion. We are using normalized image coordinates, i.e. $\Omega = [0, 1] \times [0, 1]$. The grid size parameter is set to one pixel, i.e. $\varepsilon = 1 / \max(w, h)$, where w and h are the width and height of the image.

Update equations (8) and (9) are perfectly suited for an accelerated implementation on programmable graphics processing units (GPU). Every iteration of the update equations can be performed on the GPU using a single pass. Since GPUs currently cannot update elements in-place, the implementation uses alternating render targets, and the resulting numerical procedure resembles Jacobi iterations (rather than the faster Gauss-Seidel scheme). All reported results were obtained using one GPU core of a Nvidia GeForce 7950 GX2 graphics card. The achieved performance was approximately 500 iterations per second for a 450×375 color image. We fixed the maximum number of iterations to 2500 and therefore the stereo results were obtained after approximately 5 seconds.

5. Results

5.1. Benchmark Datasets

The datasets from the Middlebury stereo vision page [21] currently constitute the standard benchmark suite for computational stereo approaches. In our experiments, the initial matches consist of four depth maps: the first map is generated using absolute differences of the gradients within a 3×3 window. The other depth maps are calculated using local adaptive support windows [36]. The employed window shapes are 5×5, 7×7 and 9×9 pixels. Depth map extraction is performed by a winner-takes-all method. All results were computed using the following parameters: $\gamma = 0.9$, $\delta = 1.0$, $scale = 10/512$, $contrast = 7$.

Fig. 3 displays the obtained depth maps and the difference to the ground truth (with error threshold set to 1) using our approach with constant parameter settings. Table 1 gives details about stereo results of the Middlebury data set for both the initial and the final depth maps. From this we can see an average reduction of the error by 11.5% using a threshold of 1.0 and 13.96% using a threshold of 0.5.

The most difficult stereo pair for our algorithm is the *Tsukuba* data set, where our approach performs substantially inferior than on the other datasets. In contrast, the *Teddy* and *Cones* image pairs are handled very well, and the obtained results can easily compete with those generated by sophisticated global methods. The intrinsic subpixel estimation and smoothing property of our approach yields to highly accurate depth maps, which is indicated by the evaluation results using an error threshold of 0.5.

5.2. Real Datasets

In this section we present the results for two real-world datasets not part of a benchmark suite. Since both datasets are captured in an outdoor setting, we employ the normalized cross correlation as similarity score to compensate for changes in exposure time. The initial weak depth maps are obtained using different support window sizes (5, 7, 9

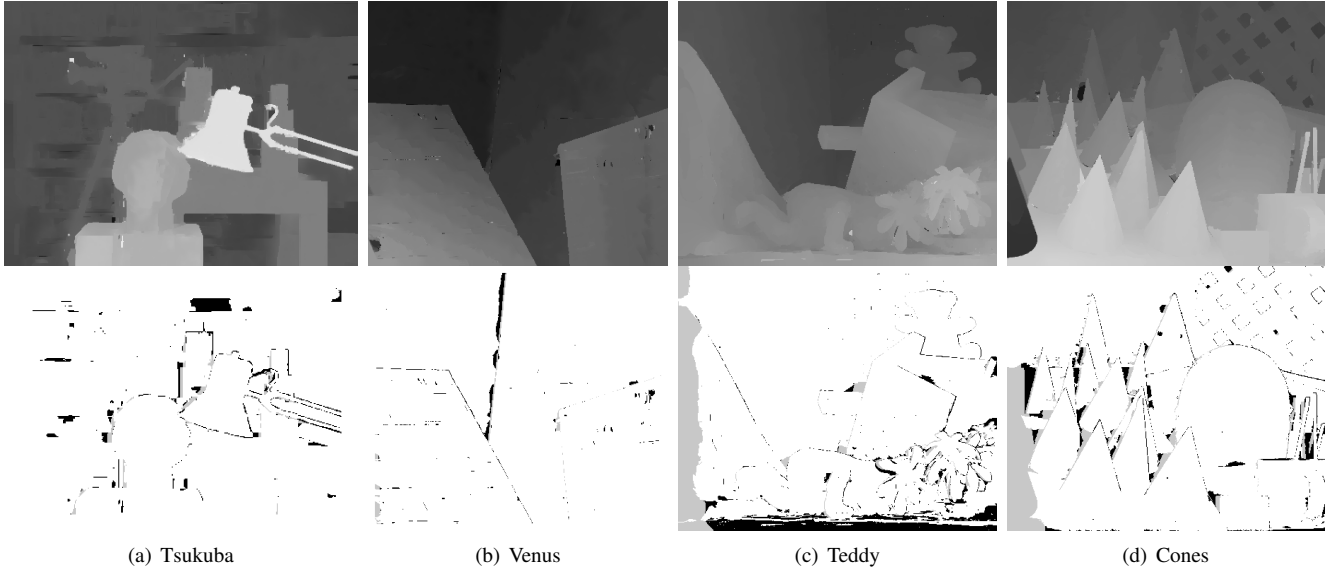| | (a) Tsukuba | (b) Venus | (c) Teddy | (d) Cones |

Figure 3. Depth images and error maps generated for the well-known stereo vision benchmark datasets.

Table 1. Stereo results using the Middlebury stereo vision benchmark datasets.

	Tsukuba		Venus		Teddy		Cones		Avg. Rank	
threshold	*1.0*	*0.5*	*1.0*	*0.5*	*1.0*	*0.5*	*1.0*	*0.5*	*1.0*	*0.5*
initial	14.1	28.8	19.8	28.1	17.7	25.1	8.67	14.3	29.0	27.1
final	2.86	18.3	1.10	3.45	6.63	11.2	3.67	7.52	13.6	9.4

and 11 pixels width and height). All initial depth maps are computed using a GPU-accelerated plane-sweep approach to depth estimation using a triplet of source images with a small baseline in-between.

The first dataset depicted in Fig. 4 represents an outdoor statue featuring several depth discontinuities. The starting depth map d^0 is the median of the depth hypotheses (Fig. 4(b)) illustrates the typical foreground fattening effect induced by larger support windows. The depth discontinuities in the final depth result (Fig. 4(c)) are much closer to the true ones. In Fig. 4(d) the corresponding 3D geometry is displayed as a colored point cloud. Since the edges in the color image are rather weak, a low setting of *contrast* is required to preserve the depth discontinuities.

Another triplet of source images is shown in Fig. 5(a)–(c). The columns in front of the facade are good indicators for depth discontinuities displayed in Fig. 5(g) (with the exception of the small region above the entrance, where the color image and the initial depth map do not provide sufficiently large edges to separate the column and facade regions).

6. Conclusion

We presented a novel computational stereo approach, which combines a highly successful image segmentation method with a robust voting scheme to integrate several weak depth images into a common depth estimate. Selective spatial smoothing as induced by the joint color/depth segmentation guides the voting procedure in order to remove outliers present in all supplied weak depth maps.

Although the proposed energy functional and the numerical scheme may appear somewhat complex at the first glance, the essence of the method is very simple and can be effectively accelerated by graphics processing units.

There are several directions for potential future enhancements. Using global values of *scale* and *contrast* for the entire images yields non optimal segmentation results in many real-world cases. A scheme to determine these values adaptively depending on the local image content is expected to prove beneficial [31].

Currently we do not consider the certainty of the depth hypotheses as outlined e.g. in [29]. A possible direction is the incorporation of a certainty measure based on the matching cost distribution into the robust data term (5).

Acknowledgements
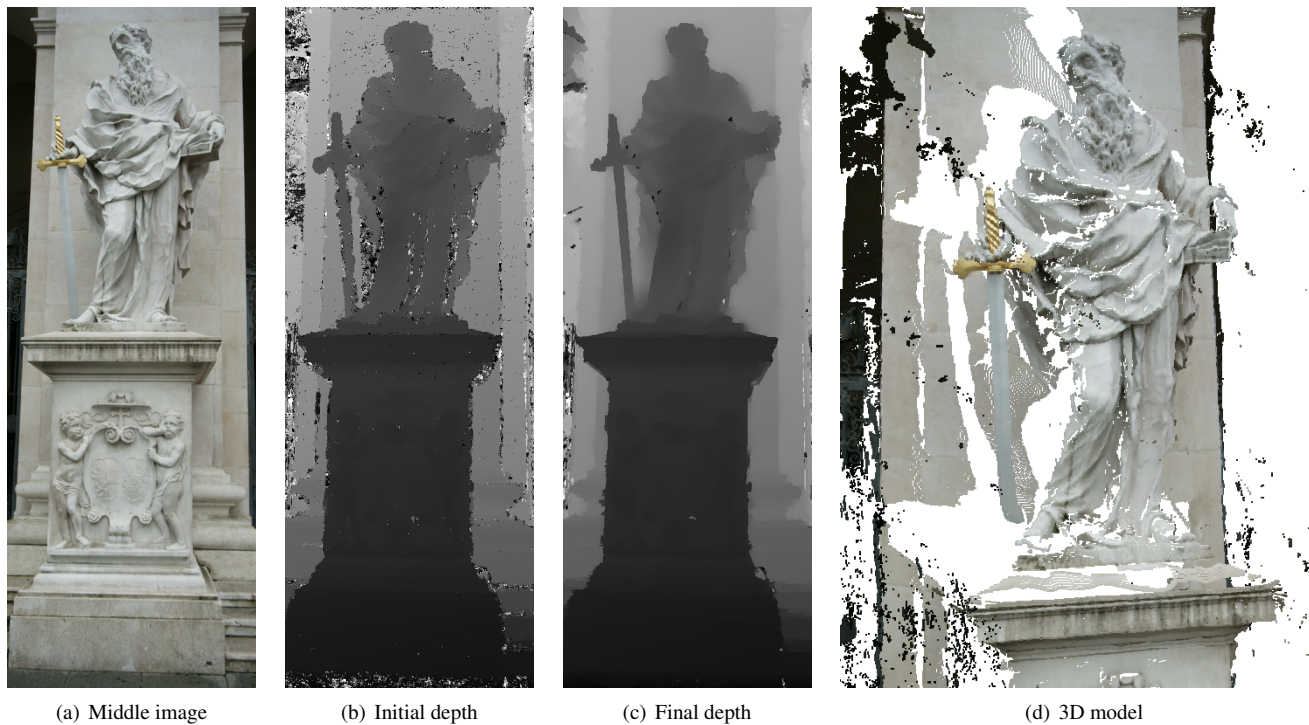
(a) Middle image   (b) Initial depth   (c) Final depth   (d) 3D model

Figure 4. The result for a statue dataset ($\gamma = 0.9$, $\delta = 1.0$, $scale = 10/512$, $contrast = 2$).

(a) Left image   (b) Middle (key) image   (c) Right image   (d) Segmentation result

(e) Final edges   (f) Initial depth   (g) Final depth   (h) 3D model
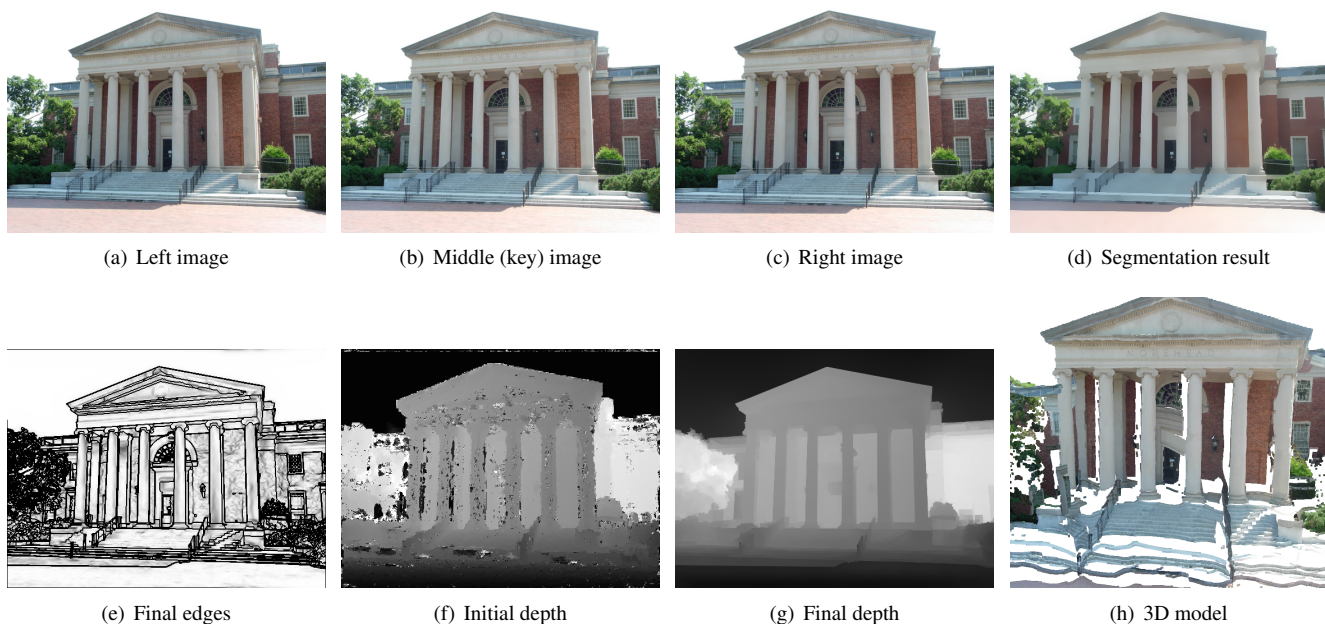
Figure 5. The result for a facade dataset. ($\gamma = 0.9$, $\delta = 1.0$, $scale = 7/512$, $contrast = 7$)

References

[1] L. Ambrosio and V. Tortorelli. Approximation of functionals depending on jumps by elliptic functionals via Γ-convergence. *Comm. Pure Appl. Math.*, 43:999–1036, 1990. 2

[2] S. Birchfield and C. Tomasi. A pixel dissimilarity measure that is insensitive to image sampling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(4):401–406, 1998. 2

[3] Y. Boykov, O. Veksler, and R. Zabih. A variable window ap-

proach to early vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 20(12):1283–1295, 1998. 2

[4] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 23(11):1222–1239, 2001. 2

[5] A. Brook, R. Kimmel, and N. Sochen. Variational restoration and edge detection for color images. *J. Mathematical Imaging and Vision*, 18:247–268, 2003. 2, 3

[6] J. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 8(6):679-698.*, 8(6):679–698, 1986. 2

[7] A. Chambolle. Finite differences discretization of the Mumford-Shah functional. *Math. Modelling and Numerical Analysis*, 33:261–288, 1999. 3

[8] T. Chan and J. Shen. *Image Processing and Analysis*. SIAM, Philadelphia, 2005. 2

[9] T. Chan and L. Vese. Active contours without edges. *IEEE Trans. Image Processing*, 10(2):266–277, February 2001. 2

[10] D. Comaniciu and P. Meer. Mean shift: A robust approach toward feature space analysis. *IEEE Trans. Pattern Anal. Machine Intell.*, 24(5):603–619, 2002. 2

[11] M. Droske. *On Variational Problems and Gradient Flows in Image Processing*. PhD thesis, Universiät Duisburg-Essen, 2005. 2

[12] S. Esedoglu and J. Shen. Digital inpainting based on the Mumford-Shah-Euler image model. *Eur. J. Applied Mathematics*, 13:353–370, 2002. 2

[13] A. Klaus, M. Sormann, and K. Karner. Segment-based stereo matching using belief propagation and a self-adapting dissimilarity measure. In *International Conference on Pattern Recognition (ICPR)*, 2006. 2

[14] J. Little. Accurate early detection of discontinuities. In *Vision Interface*, pages 97–102, 1992. 2

[15] J. M. Morel and S. Solimini. *Variational Models in Image Segmentation*. Birkhäuser, Boston, 1995. 2

[16] M. Morini and M. Negri. Mumford Shah functional as Γ-limit of discrete Perona-Malik energies. *Math. Models and Methods in Applied Sciences*, 13:785–805, 2003. 3

[17] D. Mumford and J. Shah. Optimal approximation by piecewise smooth functions and associated variational problems. *Comm. Pure Appl. Math.*, 42:577–685, 1989. 1, 2

[18] J. Petitot. An introduction to the Mumford-Shah segmentation model. *J Physiol Paris*, 97(2-3):335–342, 2003. 2

[19] J.-P. Pons, R. Keriven, and O. Faugeras. Modelling dynamic scenes by registering multi-view image sequences. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 822–827, 2005. 2

[20] D. Scharstein and R. Szeliski. Stereo matching with nonlinear diffusion. *Int. Journal on Computer Vision*, 28(2):155–174, 1998. 2

[21] D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *Int. J. Comput. Vision*, 47(1-3):7–42, 2002. 2, 5

[22] D. Scharstein and R. Szeliski. High-accuracy stereo depth maps using structured light. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 195–202, 2003. 2, 3

[23] J. Shah. A nonlinear diffusion model for discontinuous disparity and half-occlusions in stereo. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 34–40, 1993. 2

[24] J. Shen. Piecewise $H^{-1} + H^0 + H^1$ images and the Mumford-Shah-Sobolev model for segmented image decomposition. *Applied Mathematics Research Express*, 4:143–167, 2005. 2

[25] N. Slesareva, A. Bruhn, and J. Weickert. Optic flow goes stereo: A variational method for estimating discontinuity-preserving dense disparity maps. In *Proc. 27th DAGM Symposium*, pages 33–40, 2005. 2

[26] C. Strecha and L. Van Gool. PDE-based multi-view depth estimation. In *1st International Symposium od 3D Data Processing Visualization and Transmission*, pages 416–425, 2002. 2

[27] J. Sun, Y. Li, S. Kang, and H.-Y. Shum. Symmetric stereo matching for occlusion handling. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, 2005. 2

[28] J. Sun, H. Y. Shum, and N. N. Zheng. Stereo matching using belief propagation. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 25(7):787–800, 2003. 2

[29] R. Szeliski and D. Scharstein. Symmetric sub-pixel stereo matching. In *European Conference on Computer Vision (ECCV)*, pages 525–540, 2002. 6

[30] S. Teboul, L. Blanc-Féraud, G. Aubert, and M. Barlaud. Variational approach for edge-preserving regularization using coupled PDE's. *IEEE Trans. Image Processing*, 7(3):387–397, March 1998. 4

[31] W. Vanzella, F. Pellegrino, and V. Torre. Self adaptive regularization. *IEEE Trans. Pattern Anal. Machine Intell.*, 26(6):804–809, June 2004. 2, 6

[32] O. Veksler. Stereo matching by compact windows via minimum ratio cycle. In *IEEE International Conference on Computer Vision (ICCV)*, pages 540–547, 2001. 2

[33] O. Veksler. Fast variable window for stereo correspondence using integral images. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, 2003. 2

[34] C. Vogel and M. Oman. Iterative methods for total variation denoising. *SIAM Journal on Scientific Computing*, 17(1):227–238, 1996. 5

[35] Q. Yáng, L. Wang, R. Yang, H. Stewnius, and D. Nistér. Stereo matching with color-weighted correlation, hierarchical belief propagation and occlusion handling. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, 2006. 2

[36] K.-J. Yoon and I.-S. Kweon. Adaptive support-weight approach for correspondence search. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 28(4):650–656, 2006. 2, 5