

# Disparity from stereo-segment silhouettes of weakly-textured images

Babette Dellen<sup>1</sup>  
bkdellen@bccn-goettingen.de

Florentin Wörgötter  
worgott@bccn-goettingen.de

Bernstein Center for Computational Neuroscience (BCCN) Göttingen, Max-Planck-Institute for Dynamics and Self-Organization, Göttingen, Germany  
BCCN Göttingen, University Göttingen, Göttingen, Germany

---

## Abstract

We propose a novel robust stereo algorithm for weakly-textured scenes. Unique correspondences existing between the silhouettes of corresponding image segments allow assigning accurate disparities to segment boundary points. This information as well as stereo from the weak texture inside segments, which is extracted using a region-constrained window-based matching algorithm, are fused and disparities are interpolated inside segments while considering potentially occluded areas derived from the depth-ordering of segments. The algorithm is applied to a set of weakly-textured images and it is demonstrated that stereo from segment silhouettes often provides sufficient information to reconstruct disparities in weakly- and non-textured image areas. The algorithm is applied to several real stereo images and its performance is evaluated quantitatively using images from the 2006 Middlebury dataset.

## 1 Introduction

In stereo vision, 3D information is reconstructed from two images of the same scene taken from different viewpoints. Different approaches to stereo disparity estimation have been extensively compared in several studies [4, 12, 13]. Most stereo algorithms perform well in textured image areas, but often fail when there is only weak texture, due to the correspondence problem. Here local matching fails, and, as a consequence, global methods do not deliver correct disparities either, simply because the energy functions used in global methods remain under-constrained. However, stereo from weakly textured images is important for many applications, which take place in urban or industrial settings, where little texture exists, requiring novel solutions to the stereo problem.

While being ill-suited for stereo analysis, weakly-textured image parts can easily be used for color-based segmentation and, in addition, it is often also possible to find unique *segment correspondences* between images. The goal of the present study is to recover disparity in weakly-textured image parts by using correspondent image segmentation together with an interpolation algorithm based on a spring-mass model. To this end, we constrain interpolation (the springs and the masses) by the disparities of the segment-edges as well as by

---

<sup>1</sup>BD is also with the Institut de Robòtica i Informàtica Industrial (CSIC-UPC), Barcelona, Spain  
© 2009. The copyright of this document resides with its authors.  
It may be distributed unchanged freely in print or electronic forms.

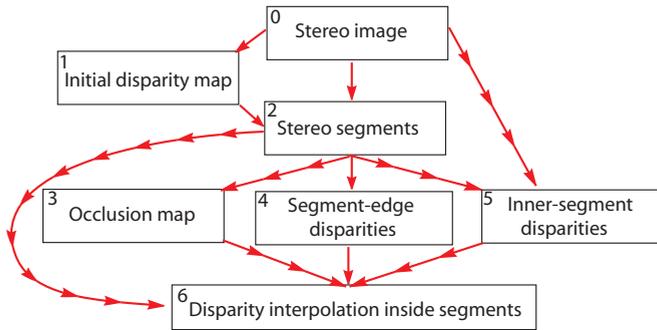


Figure 1: Block diagram of the algorithm. See text for details.

the vague, remaining stereo information inside a segment, which we can still recover using conventional stereo algorithms. This way we can regenerate rather accurate disparity information in regions that are usually quite resistant to stereo analysis, such as some images from the 2006 Middlebury stereo dataset, which are, for this reason, rarely being used for stereo-algorithm benchmarking. Furthermore, while the algorithm consists of several components, our analysis below shows that the same parameter set can be used for all images analyzed so far.

The paper is structured as follows. First, we provide an overview of the different algorithmic components of the method and present first results for weakly-textured Middlebury stereo images in Section 2. In Sections 3, 4, 5, 6, and 7 the different algorithmic steps are explained in greater detail. In Section 8, the algorithm is applied to two other scenes which contain very little texture and the results are evaluated quantitatively. Finally, in Section 9, the results are discussed and directions for future research are given.

## 2 Overview of the algorithm

The proposed method is based on the assumption that regional correspondences between stereo images provide robust and valuable information about the 3D structure of the visual scene that can be exploited to obtain dense stereo information even in untextured image areas. The stereo image is first decomposed into corresponding regions, i.e. stereo segments (see step 1-2 of Fig. 1). More details can be found in Section 3. A representative example of such a decomposition can be seen in Fig. 4(a-d). This part of the algorithm can be replaced by other methods, if desired, such as [11, 16].

For each stereo segment, segment silhouettes are computed in both frames. Unique correspondences of silhouette-edge points are searched, and the respective disparities are calculated (see also Section 4 and step 4 of Fig. 1). While many segment edges represent real object boundaries, some edges are in fact caused by an occlusion and thus do not represent a “true” edge. These occluded edges need to be identified to avoid erroneous interpolation results.

Occluded areas are estimated by finding the approximate depth ordering of the scene (see Section 6 and step 3 of Fig. 1). The segment-center disparity for each stereo segment is computed from the distance between the left and right segment center. The stereo segments

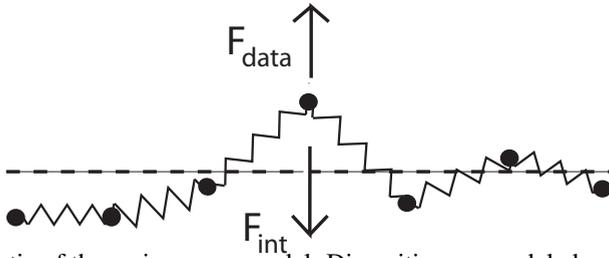


Figure 2: Schematic of the spring-mass model. Disparities are modeled as positions of mass points, connected to their nearest neighbors through interaction forces (springs). Data forces incorporate precomputed disparities of segment boundaries and inner-segment areas.

are ordered according to their segment-center disparity. An edge of a segment that is adjacent to an edge of another segment with larger disparity (thus, it is closer) is then assumed to be caused by an occlusion, and an occlusion map is obtained.

Additionally, texture inside segments can be exploited by applying a window-based stereo algorithm which operates strictly inside stereo segments (see Section 5 and step 5 of Fig. 1), thus occlusion problems do not arise. Confidence values are computed and only those disparities are used that have a high confidence, resulting usually in sparse disparity maps.

The sparse disparity information from both the segment-silhouette edges and inner-segment texture are used to interpolate disparity values inside segment areas (see Section 7 and step 6 of Fig. 1). For this purpose, we developed an interpolation algorithm based on a spring-mass model incorporating region constraints. Each pixel of the image is described by a mass point, connected via elastic springs to its nearest neighbor (see Fig. 2). However, if the neighbor belongs to a different segment, the connection is cut. The amplitude of the mass point represents its disparity value. The precomputed sparse disparity maps are used to define data forces which push the respective masses towards the precomputed value. A damping force is included to drive the system to a local minimum. The corresponding system of equations is numerically solved.

Results of the algorithm applied to the stereo-image pair *Baby1* (Fig. 3(a)) from the Middlebury stereo database (<http://vision.middlebury.edu/stereo/data/scenes2006/>) [15] are shown in Fig. 3(c,f). The basic structure of the scene could be recovered, as compared to ground truth (see Fig. 3(b)). Results of a state-of-the-art global stereo algorithm are shown in Fig. 3(d) for comparison. These results have been obtained using the stereo-matcher front end and the MRF library (<http://vision.middlebury.edu/MRF/code/>) [15], based on the works of [2, 3, 8, 15]. For optimization, we used the graph cuts algorithm with  $\alpha$  expansion and default parameters except for the disparity range, which was adjusted to fit the disparity range of the respective image. For less textured scenes, the graph-cuts algorithm delivers poor results, which we do not show in most cases, since it would be an unfair comparison, considering that these algorithms are designed for sufficiently textured images. The graph-cuts based stereo algorithm provides for example excellent results for more textured images, e.g. *Tsukuba*, *Venus*, *Cones*, and *Teddy* from the Middlebury dataset. Nevertheless, we listed the computed average error for the selected Middlebury images in Table 1 to provide the reader at least some orientation for evaluating our results. Results of our algorithm for the Middlebury images *Lampshade2*, *Plastic*, and *Wood1* are shown in Fig. 3(i,l,o). Please note that all areas for which no disparities could be given have been colored black, including

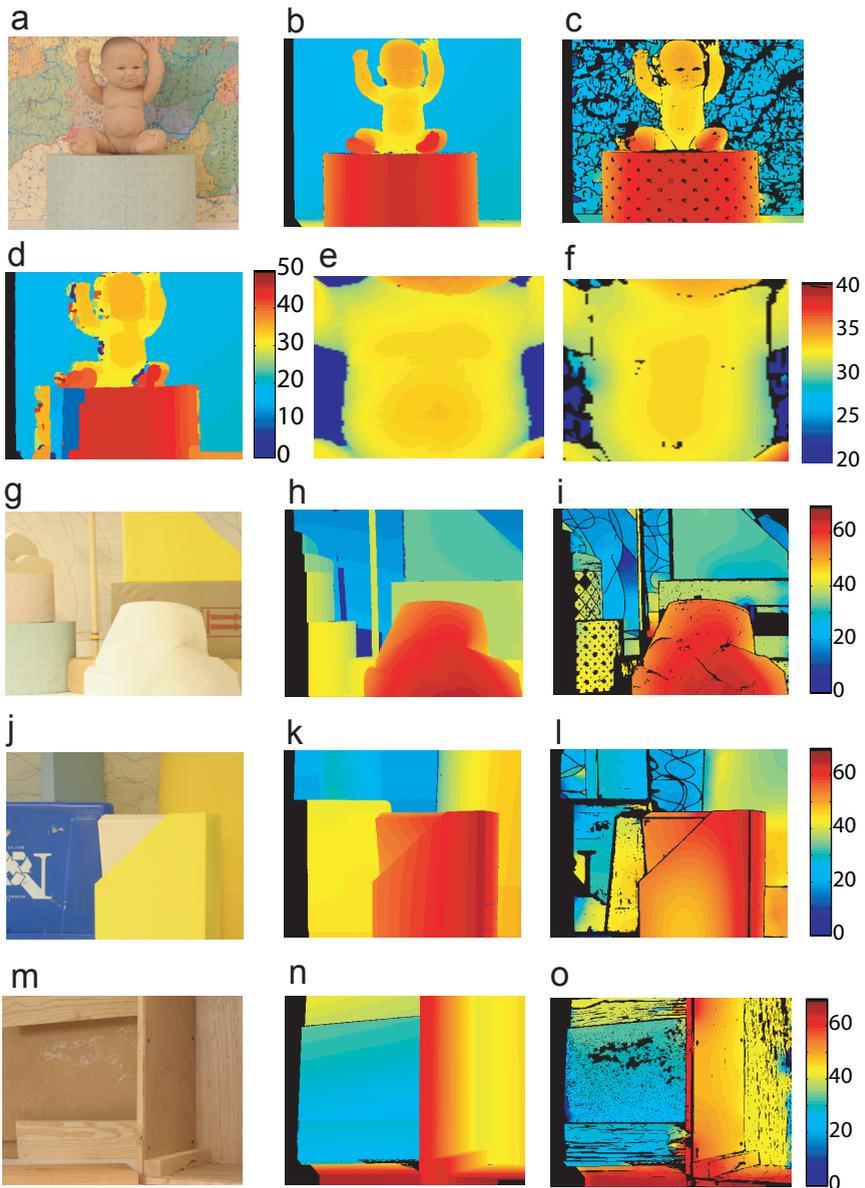


Figure 3: Results for the Middlebury dataset 2006 (a) *Baby1* left image. (c) Ground truth. (c) Estimated disparities (d) Result of graph-cuts based global stereo algorithm. (e) Enlarged part of the baby torso (ground truth). (f) Same area for our results. (g) *Lampshade2* left image. (h) Ground truth. (i) Estimated disparities. (j) *Plastic* left image. (k) Ground truth. (l) Estimated disparities. (m) *Wood1* left image. (n) Ground truth. (o) Estimated disparities. Images can be retrieved from <http://vision.middlebury.edu/stereo/data/scenes2006/>.

monocular parts.

### 3 Stereo Segments

Our first goal is to get stereo segments (see step 1-2 of Fig. 1). For this purpose, we use an image-segmentation method in which stereo segments are obtained through a 3D linking process. First, a spin variable  $\sigma_i$  is assigned to each pixel  $i$  of the stereo image. To incorporate constraints in form of local correspondence information, we distinguish between neighbors within a single frame (2D bonds) and neighbors across frames (3D bonds). We create a 2D bond  $(i, k)_{2D}$  between two pixels within the same frame with coordinates  $(x_i, y_i, z_i)$  and  $(x_k, y_k, z_k)$  if  $|(x_i - x_k)| \leq 1$ ,  $|(y_i - y_k)| \leq 1$ , and  $z_i = z_k$ . Across frames, we create a 3D bond  $(i, j)_{3D}$  between two spins  $i$  and  $j$  if  $|(x_i + d_{ij}^x - x_j)| \leq 0.5$ ,  $y_i = y_j$ ,  $z_i \neq z_j$ , and  $a_{ij} = 1$ . The values  $d_{ij}^x$  is the shift of the pixels between frames  $z_i$  and  $z_j$  along the axis  $x$ , obtained from an initial disparity map (step 1 of Fig. 1). The parameters  $a_{ij}$  are the respective amplitudes (or confidences). Disparities used for the linking process are obtained using the stereo-matcher front end and the MRF library from <http://vision.middlebury.edu/MRF/code/> [15], based on the works of [2, 3, 8, 15], using  $\alpha$  expansion for optimization, delivering an amplitude of 1 everywhere. However, this method can be replaced by any stereo algorithm of choice, since the segment-linking process is highly robust to errors in the initial disparity map and delivers correct results even for very sparse disparity maps, requiring only an initial disparity “guess”.

The spin model is now implemented such that neighboring spins with similar color have the tendency to align. We use a  $q$ -state Potts model [10] with the Hamiltonian

$$H = - \sum_{\langle ik \rangle_{2D}} J_{ik} \delta_{\sigma_i, \sigma_k} - \sum_{\langle ij \rangle_{3D}} J_{ij} \delta_{\sigma_i, \sigma_j} \quad , \quad (1)$$

with  $J_{ij} = 1 - \Delta/\bar{\Delta}$  and  $\Delta_{ij} = |g_i - g_j|$ , where  $g_i$  and  $g_j$  are the gray (color) values of the pixels  $i$  and  $j$ , respectively. The mean distance  $\bar{\Delta}$  is obtained by averaging over all bonds.

Here,  $\langle ik \rangle_{2D}$  and  $\langle ij \rangle_{3D}$  denote that  $i, k$  and  $i, j$  are connected by bonds  $(i, k)_{2D}$  and  $(i, j)_{3D}$ , respectively. The Kronecker  $\delta$  function is defined as  $\delta_{a,b} = 1$  if  $a = b$  and zero otherwise. The segmentation problem is then solved by finding clusters of correlated spins in the low temperature equilibrium states of the Hamiltonian  $H$ . The total number  $M$  of segments is then determined by counting the computed segments. It is usually different from the total number  $q$  of spin states, which is a parameter of the algorithm (here  $q = 10$ ).

We solve this task by implementing a clustering algorithm. In a first step, “satisfied” bonds, i.e. bonds connecting spins of identical spins  $\sigma_i = \sigma_j$ , are identified. Then, in a second step, the satisfied bonds are “frozen” with a some probability  $P_{ij}$ . Pixels connected by frozen bonds define a cluster, which are updated by assigning to all spins inside the same clusters the same new value [14]. In the method of superparamagnetic clustering proposed by Blatt *et al.* (1996) [1] this is done independently for each cluster. In this paper, we will employ the method of energy-based cluster updating (ECU), where new values are assigned in consideration of the energy gain calculated for a neighborhood of the regarded cluster [9, 17]. The algorithm is controlled by a single “temperature” parameter, and has been shown to deliver robust results over a large temperature range. After a 140 iterations, clusters are used to define stereo segments (see Fig. 4(a-b)).

## 4 Disparities of segment-silhouette points

Our second goal is the computation of disparities of segment-silhouette boundary points (see step 4 of Fig. 1). Given two corresponding segments, a unique correspondence for boundary pixels can be found for the outer segment boundaries as long as the pixel does not belong to a boundary segment oriented parallel to  $x$ -axis, i.e. the scanline. A pixel  $i$  with segment label  $s_i$  and position  $x_i$  is considered to be a left segment boundary pixel  $i_l$  if there is no pixel  $u$  with  $x_u < x_i$  and  $s_u = s_i$ . A pixel  $i$  with segment label  $s_i$  is considered to be a right segment boundary pixel  $i_r$  if there is no pixel  $u$  with  $x_u > x_i$  and  $s_u = s_i$ . For each scanline, we find the boundary pixels  $i_l$  and  $i_r$  for the left image and the boundary pixels  $j_l$  and  $j_r$  for the right image. The edge disparities are then defined as  $d_i^e = d_{i_l,i} \delta_{i_l,i} + d_{i_r,i} \delta_{i_r,i}$  with amplitudes  $a_i^e = \delta_{i_l,i} + \delta_{i_r,i}$  where  $d_{i_l} = i_l - j_l$  and  $d_{i_r} = i_r - j_r$ .

## 5 Disparities from inner-segment texture

Our third goal is the computation of disparities from weak texture inside stereo segments (see step 5 of Fig. 1), using a region-constrained window-based matching algorithm. For each pixel  $i$  of the left image  $I_L$ , we define a rectangular window  $R_i$  around pixel  $i$ , and compute its correlation coefficient  $c_{i,j} = a_{i,j}/(b_i b_j)$  with the rectangular window  $R_j$  of each pixel  $j$  of the right image  $I_R$  along the  $x$ -axis (scanline), with

$$a_{i,j} = \sum_{u \in R_j} \sum_{v \in R_i} [I_R(x_u, y_u) - m_j] [I_L(x_v, y_v) - m_i] \delta_{s_u, s_i} \delta_{s_v, s_i} \quad (2)$$

$$b_i = \sum_{v \in R_i} [I_L(x_v, y_v) - m_i] \delta_{s_v, s_i} \quad (3)$$

$$b_j = \sum_{u \in R_j} [I_R(x_u, y_u) - m_j] \delta_{s_u, s_i} \quad , \quad (4)$$

where  $s_i$ ,  $s_u$ , and  $s_v$  are the segment labels of pixels  $i$ ,  $u$ , and  $v$ , respectively, and

$$m_i = \sum_{v \in R_i} I(x_v, y_v) \delta_{s_v, s_i} / \sum_{v \in R_i} \delta_{s_v, s_i} \quad (5)$$

$$m_j = \sum_{u \in R_j} I(x_u, y_u) \delta_{s_u, s_i} / \sum_{u \in R_j} \delta_{s_u, s_i} \quad . \quad (6)$$

For each pixel  $i$  we find the pixel  $j_{\max} = \arg[\max_j(c_{i,j})]$  for which the correlation coefficient  $c_{i,j}$  is maximal. The disparity value of pixel  $i$  is defined as

$$d_i^c = x_i - x_{j_{\max}} \quad , \quad (7)$$

and the respective confidence value is given by  $c_{i,j_{\max}}$ . The amplitude  $a_i^c$  is further defined by

$$a_i^c = \theta \left( \sum_{u \in R_{j_{\max}}} \sum_{v \in R_i} \delta_{s_v, s_i} \delta_{s_u, s_i}, \tau_1 \right) \theta(c_{i,j_{\max}}, \tau_2) \quad , \quad (8)$$

where  $\tau_1$  and  $\tau_2$  are thresholds, and  $\theta$  is a step function with  $\theta(f, \tau) = 1$  if  $f$  is larger than  $\tau$  and zero otherwise.

## 6 Potentially occluded areas

Our fourth goal is the identification of potentially occluded areas (see step 3 of Fig. 1). These occlusions are computed from the ordering of segments. For each stereo segment, we compute the disparity of the center of mass of the segment

$$d_{cm,i} = \frac{\sum_{u \in S_{i,left}} x_u / \sum_{u \in S_{i,left}} 1 - \sum_{u \in S_{i,right}} x_u / \sum_{u \in S_{i,right}} 1}{2}. \quad (9)$$

Then we compute the Occlusion Map according to

$$O_i = \theta \left( \sum_{j \in N_j} \theta(d_{cm,j} - d_{cm,i}, \tau_3), 0 \right) \quad (10)$$

where  $N_j$  is rectangular neighborhood around pixel  $j$  of  $5 \times 5$  pixels and  $\tau_3 = 5$  is a parameter. The disparities of occluded pixels are excluded from the interpolation process.

## 7 Disparity interpolation within segments

Our final goal is the fusion of disparity information derived in the previous steps and the subsequent interpolation of disparities inside segments to obtain a dense disparity map (see step 6 of Fig. 1). We assign a mass  $m_i$  to each pixel  $i$  of the left image. All masses within a segment are connected by springs with a spring constant  $k$ . Each mass is described by “position”  $x_i$  and “velocity”  $v_i = dx_i/dt$ . The position corresponds to the instantaneous disparity value of the mass, scaled by factor  $f_s = 5/d_{\max}$ , where  $d_{\max} = 1.5\max(d_i^e)$ . We obtain the following dynamical system

$$dx_i/dt = v_i \quad (11)$$

$$dv_i/dt = -\gamma v_i + F_{\text{edges}} + F_{\text{inner}} + F_{\text{int}} \quad (12)$$

with the data forces  $F_{\text{edges}} = k^e a_i^e (f_s d_i^e - x_i)$  and  $F_{\text{inner}} = g^c a_i^c (f_s d_i^c - x_i)$ , and an interaction force  $F_{\text{int}} = k \sum_{l \in N_i} (x_l - x_i) \delta_{s_i, s_l}$ , where the sum goes over all four nearest neighbors  $N_i$  of mass  $i$  which are within the same segment. The constant  $\gamma = 0.5$  determines the amount of damping of the system. The system is initialized with random values and the developed using a 4th order Runge-Kutta technique with 1200 iteration steps and a step size of  $1/6$  time units. To remove outliers from the inner disparity data, we define  $g^c = k^c (1 - |x_i - d_i^c|)$  if  $g^c > 0$  and zero otherwise. Spring-mass models have been employed previously for 3D reconstruction and global optimization in early vision [5, 6], but not in combination with image segmentation.

## 8 Results

We apply the algorithms to various stereo images with parameter choices  $\tau_1 = 40$ ,  $\tau_2 = 0.92$ ,  $k^c = 0.25$ ,  $k^e = 5$ ,  $k = 10$ , and window-size for the region-constrained correlation-based matching algorithm of  $11 \times 11$  pixels. For the computation of stereo segments, the same temperature  $T = 0.27$  was used for all images and parameters to obtain the initial disparity map were not changed. Disparities of segments  $i$  for which  $\sum_{j \in S_i} (a_j^c + a_j^e) = 0$  are set to zero

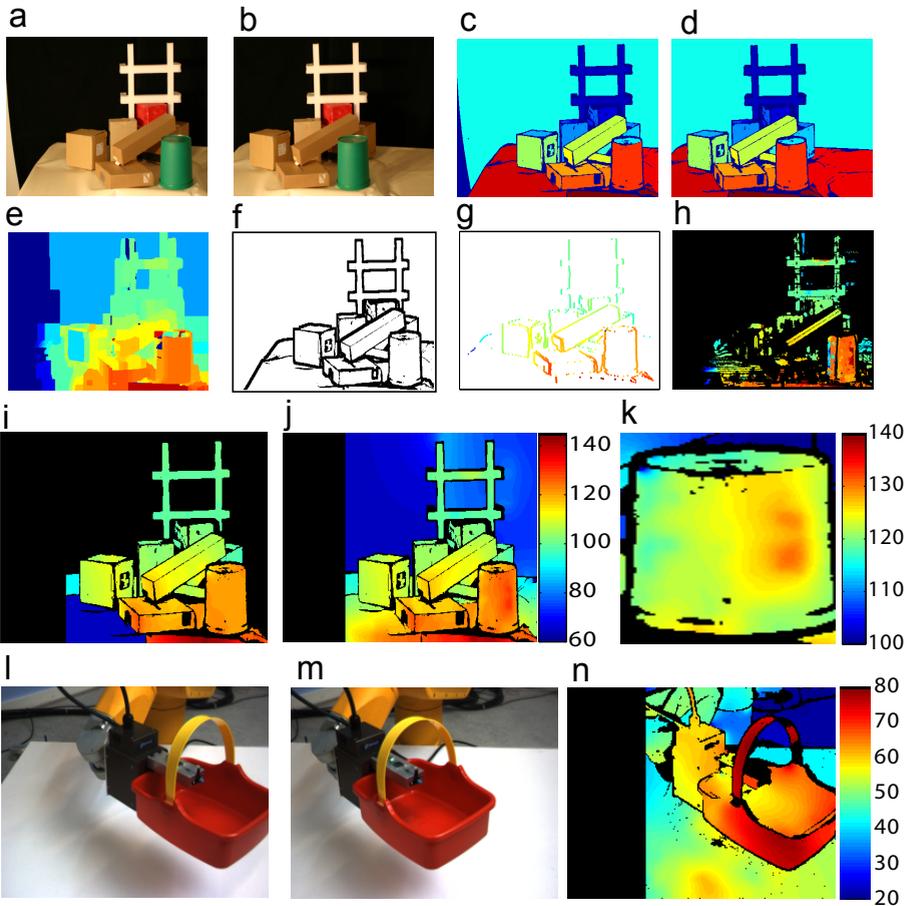


Figure 4: Results for real weakly-textured stereo pairs. (a) Left image of *Cluttered Scene*. (b) Right Image. (c) Segmentation of left image. (d) Segmentation of right image. (e) Initial disparity map (graph cuts). (f) Estimated occluded areas. (g) Edge disparities. (h) Disparities in inner-segment areas. (i) Interpolation result using edge disparities only. (j) Interpolation result using both edge and inner-segment disparities. (k) Result of trash can (enlarged). (l) Left image of *Basket scene*. (m) Right image of *Basket scene*. (n) Estimated disparities. Monocular parts have been colored black.

Sequence	$\bar{e}$	$\rho$	< 0.5	< 1	< 1.5	< 2	< 2.5	< 3	$\bar{e}_{gc}$	$\rho_{gc}$
Baby1	0.81	75	72	85	89	92	93	95	2.7	100
Lampshade2	2.0	85	49	65	75	79	83	85	8.44	100
Plastic	3.97	89	27	37	44	50	54	58	11.84	100
Midd1	1.39	67	62	75	79	83	85	87	1.76	100
Bowling2	1.44	81	64	77	82	86	88	89	1.91	100
Wood1	1.91	82	53	64	71	76	79	82	3.55	100

Table 1: Mean disparity errors, densities, and percentage of pixels having an disparity error smaller than a threshold of 0.5-3 pixels for six stereo images from the 2006 Middlebury dataset together with the mean disparity errors  $\bar{e}_{gc}$  and the respective densities  $\rho_{gc}$  of the graph-cuts based global stereo algorithm.

and excluded from further analysis (color-coded in black). The algorithm is very robust and the same parameter set can “always” be applied.

We show the results of the different stages of the algorithm on the example of the *Cluttered Scene* stereo pair, containing many objects which contain no or only little texture (see Fig. 4(a-b)). Stereo segments are extracted, i.e. a pixel in the left image and pixel in the right image belonging to the same stereo segment are assigned the same label. The color-coded results are depicted in Fig. 4(c-d). The initial disparity map utilized to obtain segment correspondences is shown in Fig. 4(e). From the stereo segments, segment-edge disparities are computed, cleaned from occlusions (shown in Fig. 4(f)), and presented in Fig. 4(g). In Fig. 4(h), the estimated inner-segment disparities which are above threshold are shown. Next, the edge disparities, inner-segment disparities, and segments are used to define the forces of the spring-mass model. In Fig. 4(i), the interpolation results using only the edge-disparity map are shown, demonstrating that texture is not absolutely required to recover the basic structure of the scene. The final interpolation result using both the edge-disparity map and the inner-segment disparity map is shown in Fig. 4(j). The 3D shape of most of the objects could be recovered while preserving object boundaries, as can be seen looking at the inset shown in Fig. 4(k). These results have been achieved despite the large disparities of *Cluttered Scene*, which pose an additional challenge to most stereo algorithms.

We further applied the algorithm to an extremely sparsely textured lab scene, showing an robot arm handling a plastic basket (see Fig. 4(l-m)). The stereo image pair is characterized by large disparities, caused by a drastic viewpoint change. Nevertheless, our algorithm could find a disparity map that evidently captures the based 3D structure of the scenario, shown in Fig. 4(n). We also applied the algorithm to six images from the Middlebury stereo database [12] (see Fig. 3). The qualitative results have already been described in Section 2. We further evaluated the mean disparity error  $\bar{e}$  in pixels and the density of the computed disparity maps. The mean error ranged from 0.81 to 3.97 pixels. The results are summarized in Table I.

## 9 Discussion

We presented a novel robust method for disparity computation using stereo-segment silhouettes. Despite the elaborated structure of the algorithm, the results for various stereo-image pairs could be obtained with the same parameter set. For the Middlebury stereo images,

disparity maps with an average error between 0.81 and 3.97 pixels were obtained. Even for very weakly textured scenes such as *basket*, the algorithm succeeded in capturing the basic 3D structure of the scene. We also demonstrated that the algorithm is applicable to stereo images characterized by large disparities, which often pose a problem to other stereo algorithms. However, the algorithm is designed for stereo images containing little or no texture and as such is not expected to perform well in textured areas, also because color similarity alone cannot lead to a satisfactory segmentation result.

Color segments have been used before to improve disparity estimation. Yang *et al.* (2008) applied first a window-based stereo algorithm to the stereo-image pair and then fitted planes to segments [18]. Other methods assume that planes in disparity space coincide with color-segments, improving pixel assignment to their respective disparity plane, e.g. [7]. By contrast, our method is based primarily on the computation of disparities from stereo-segment silhouettes, which, to our knowledge, have not been utilized before. Texture cues are used then as an additional cue to improve results, but are not mandatory. Furthermore, disparities are interpolated using a spring-mass model and not via a surface-fitting procedure.

## Acknowledgment

The work has received support from the German Ministry for Education and Research (BMBF) via the Bernstein Center for Computational Neuroscience (BCCN) Göttingen under Grant No. 01GQ0430 and the EU Project PACO-PLUS under Contract No. 027657. We thank Norbert Krüger for providing the *basket* stereo image.

## References

- [1] M. Blatt, S. Wiseman, and E. Domany. Superparametric clustering of data. *Physical Review Letters*, 76(18):3251, 1996.
- [2] Y. Boykov and Vladimir Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (9):1124–1137, 2004.
- [3] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 23(11):1222–1239, 2001.
- [4] M. Z. Brown, D. Burschka, and G. D. Hager. Advances in computational stereo. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 25(8):993–1008, 2003.
- [5] B. Dellen and F. Wörgötter. Simulating dynamical systems for early vision. In *Int. Conf. Comp. Vis. Theory and Applications (Visapp)*, volume 2, pages 525–528, 2009.
- [6] B. Dellen, G. Alenya, S. Foix, and C. Torras. 3d object reconstruction from swiss-ranger sensor data using a spring-mass model. In *Int. Conf. Comp. Vis. Theory and Applications (Visapp)*, volume 2, pages 368–372, 2009.
- [7] A. Klaus, M. Sormann, and K. Karner. Segment-based stereo matching using belief propagation and a self-adapting dissimilarity measure. In *Intl. Conf. Pattern Recognition*, pages 15–18, 2006.

- [8] V. Kolmogorov and R. Zabih. What energy functions can be minimized via graph cuts? *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(2):147–159, 2004.
- [9] R. Opara and F. Wörgötter. A fast and robust cluster update algorithm for image segmentation in spin-lattice models without annealing – visual latencies revisited. *Neural Computation*, 10:1547–1566, 1998.
- [10] R. B. Potts. Some generalized order-disorder transformations. *Proc. Cambridge Philos. Soc.*, 48:106–109, 1952.
- [11] Carsten Rother, Tom Minka, Andrew Blake, and Vladimir Kolmogorov. Cosegmentation of image pairs by histogram matching - incorporating a global constraint into mrfs. In *CVPR '06: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 993–1000, 2006.
- [12] 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.
- [13] S. M. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski. A comparison and evaluation of mulit-view stereo reconstruction algorithms. In *IEEE Conf. Compu. Vision and Pattern Recognition*, pages 519–528, 2006.
- [14] R.H. Swendsen and S. Wang. Nonuniversal critical dynamics in monte carlo simulations. *Physical Review Letters*, 76(18):86–88, 1987.
- [15] R. Szeliski, R. Zabih, D. Scharstein, O. Veksler, V. Kolmogorov, A. Agarwala, M. Tappen, and C. Rother. A comparative study of energy minimization methods for markov random fields. In *Ninth European Conference on Computer Vision*, volume 2, pages 16–29, 2006.
- [16] Alexander Toshev, Jianbo Shi, and Kostas Daniilidis. Image matching via saliency region correspondences. In *IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis*, 2007.
- [17] C. von Ferber and F. Wörgötter. Cluster update algorithm and recognition. *Physical Review E*, 62:1461–1664, 2000.
- [18] Q. Yang, C. Engels, and A. Akbarzadeh. Near real-time stereo for weakly-textured scenes. In *British Machine Vision Conference*, 2008.