A Top-Down Approach to the Estimation of Depth Maps Driven by Morphological Segmentations

Jean-Charles Bricola(✉), Michel Bilodeau, and Serge Beucher

CMM – Centre de Morphologie Mathématique,
MINES ParisTech – PSL Research University,
35 rue St Honoré, 77300 Fontainebleau, France
{jean-charles.bricola,michel.bilodeau,serge.beucher}@mines-paristech.fr

Abstract. Given a pair of stereo images, the spatial coordinates of a scene point can be derived from its projections onto the two considered image planes. Finding the correspondences between such projections however remains the main difficulty of the depth estimation problem: the matching of points across homogeneous regions is ambiguous and occluded points cannot be matched as their projections do not exist in one of the image planes.

Instead of searching for dense point correspondences, this article proposes an approach to the estimation of depth map which is based on the matching of regions. The matchings are performed at two segmentation levels obtained by morphological criteria which ensure the existence of an hierarchy between the coarse and fine partitions. The hierarchy is then exploited in order to compute fine regional disparity maps which are accurate and free from noisy measurements.

We finally show how this method fits to different sorts of stereo images: those which are highly textured, taken under constant illumination such as Middlebury and those which relevant information resides in the contours only.

Keywords: Watershed · Segmentation hierarchies · Disparity estimation · Joint stereo segmentation · Non-ideal stereo imagery

1 Introduction

The estimation of depth maps from stereoscopic data traditionally comes down to finding pixel correspondences between stereo images. For the sake of simplicity, we assume throughout this paper that stereo images are rectified such that any scene point projects with the same ordinate in the stereo image planes. The difference in abscissa is referred to as the disparity and is inversely proportional to the depth being searched for.

Modern approaches are based on dense pixel correspondences and resort to the framework described in [11]. The first step consists of computing the matching costs between a pair of pixels for each scanline and for each disparity belonging

J.A. Benediktsson et al. (Eds.): ISMM 2015, LNCS 9082, pp. 122–133, 2015.
DOI: 10.1007/978-3-319-18720-4_11

to the search domain. The cost is obtained by means of a dissimilarity measure between the patches centred at the pixels being under study. The reader may find an exhaustive list of such measures in [7]. The second phase of the estimation aims at ensuring a disparity consistency among the scanlines. Local approaches achieve this by diffusing the costs across the scanlines which results in fast algorithms and end with a refinement process. Global approaches generally associate an energy to the disparity map which takes account of both the disparity local costs and the consistency across neighbour pixels. This energy is eventually minimised to yield to final disparity map.

Many of these approaches seek the actual frontiers of objects within their estimation process. For instance, the gradient of the image is used in order to determine depth continuities in [6], whilst a geodesic distance across a relief defined by the image intensity values is used in the context of pixels matching [5] so as to weight the importance of pixels which are in the vicinity of the patch centre with respect to the scene. This leads to a strong interest in using regions for stereo image analysis, because they determine the frontiers of objects as well as the membership of an occluded pixel within a region that is only semi-occluded. Several region-based algorithms have already been made available which either exploit matchings across over-segmentations [15] or fit planes through object-oriented regions given a non-refined disparity map such as [4,14].

In this work, we propose a novel region-based stereo algorithm with the following contributions:

1. The disparity maps are estimated without relying on dense pixel matchings. This aspect is interesting for stereo imagery which is either poorly textured or subject to noise because the assumption that a majority of valid pixel matchings exists with respect to the usual aggregation step may no longer hold. Hence this approach does not assume the existence of a roughly good initial disparity estimate.
2. Disparities are estimated at the region level on a hierarchy of segmentations composed of the coarse level which highlights the objects in the scene and a fine level which is over-segmented. The computations are performed in waterfalls: first at the coarse level where depth planes are highlighted, then at the fine level for which a finer degree of precision is obtained. One can choose at which level to stop the algorithm depending on the speed required for its application.
3. Contours are taken into account within the estimation process after a careful reasoning on occlusion boundaries. Disparities along contours are often misused because they require the knowledge of object membership. This problem has little been raised, with the exception of [13] which exploits boundary junctions to this end.

The mechanism of the depth estimation system is presented in section 2. The method exploits the advantages offered by a marker-driven watershed. Markers are used to produce the coarse and fine segmentations and are transferred across stereo images in order to produce equivalent segmentations which facilitate the establishment of contour point correspondences. These aspects are discussed in section 3.

2 Top-Down Estimation of Disparities

The proposed method relies on the concept of *regional disparity*. We define the regional disparity as a measure attributed to a region of the reference image. This measure represents the average disparity of the pixels that compose the region and is obtained by searching for a displacement which optimally superimposes the region with the target image.

In this approach, regional disparities are initially computed for a coarse partition of the reference image. At this level, region matchings (section 2.1) are quite stable because of their singularity. The resulting disparity map constitutes a gross approximation of the true disparity map, with the most important imprecisions across regions not being fronto-parallel to the camera. At the over-segmented level, matching errors are more frequent, but disparity variation is better captured because regions are smaller. A relaxation process (section 2.2) that relies on the disparity map obtained at the coarse level is applied onto the disparities obtained for the fine partition in order to correct any mistaken measure and ensure some smoothness across the coarse segments. Finally, a regularization of the depth map based on a linear estimation process called kriging (section 2.3) is performed so as to obtain the final disparity maps.

2.1 Region Matchings and Regional Disparities

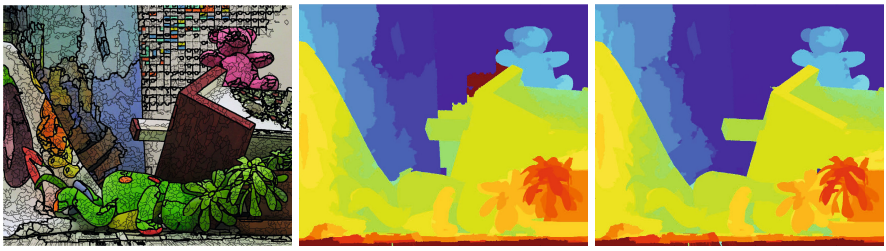
A region $\mathbf{R}$ is represented as an indicator function R such that $R[x, y] = 1 \Leftrightarrow (x, y) \in \mathbf{R}$. We denote $\mathbf{R}^{(d)}$ the region obtained by shifting $\mathbf{R}$ of d pixels along the horizontal image axis and define its indicator function by $R^{(d)}[x, y] = R[x-d, y]$. A partition is a set that contains all the regions extracted from an image. We describe the coarse partitions of the reference and the second images of the stereo pair as P_{C1} and P_{C2} respectively. We also let P_{F1} to be the fine partition obtained for the reference image. The disparity $d^\star$ is assigned to a region of the reference image if its superimposition onto the second image shifted from $d^\star$ pixels minimises a dissimilarity cost chosen according to the image acquisition setup.

A popular dissimilarity measure employed on images acquired under the same illumination conditions is the Sum of Absolute Differences (SAD) between the image intensities. It is possible to evaluate the SAD for each region of the reference image and choose the disparity that yields the superimposition with the minimum cost. However, when a region undergoes a semi-occlusion in the second image, the true superimposition is likely to come at a high cost because it partially compares pixels with an occluding object. That results in errors as can be seen in figure 1(b). To circumvent that issue, it is best to compute a mean of absolute differences for each region resulting from the intersections of the reference image partition and a coarse partition of the second image of the stereo pair at every possible disparity and search for the region and the disparity that minimises the mean.

The Jaccard distance [9] measures the overlap between two regions by computing the ratio between their intersection area and union area. The distance between two regions is then expressed as:

$$c_{\text{Jaccard}}(\mathbf{R}_i, \mathbf{R}_j) = 1 - |\mathbf{R}_i \cap \mathbf{R}_j| / |\mathbf{R}_i \cup \mathbf{R}_j| \quad (1)$$

The asymmetric version of this distance which replaces $|\mathbf{R}_i \cup \mathbf{R}_j|$ by $|\mathbf{R}_i|$ in equation 1 can be used to discard the dissimilarity costs obtained across region intersections which cover less than a reasonable threshold of the reference image region. Doing so yields the coarse regional disparity map represented in figure 1(c).



(a) (b) (c)

Fig. 1. (a) Coarse and fine segmentations obtained for the reference image of *Teddy*. (b) Regional disparities obtained for the coarse partition without a constraint partition P_{C2} and (c) with a constraint partition P_{C2}. Semi-occluded regions are less affected by measurement errors when using a constraint partition.

When images are poorly textured and acquired under different lighting conditions, the contours appear to remain the most pertinent criterion for matching regions. To this end, the gradients of the stereo images are compared using the SAD taking note of the following observation: a contour that separates two regions always constitutes the physical frontier of one of the regions but might constitute an occlusion border for the other region. For this reason, the regional disparities are computed across subregions of the reference image illustrated in figure 2: every region $\mathbf{R}_i$ of the reference image is split into two subregions denoted by $\mathbf{R}_i{-}\text{L}$ and $\mathbf{R}_i{-}\text{R}$ along its vertical skeleton. Regions which are likely to be semi-occluded are those for which the disparities of their subregions differ significantly so that the highest disparity equals the disparity of a neighbour region, the one which is occluding. The following rectification is therefore applied on the regional disparities: if $d^\star(\mathbf{R}_i{-}\text{R}) \gg d^\star(\mathbf{R}_i{-}\text{L})$ and $d^\star(\mathbf{R}_i{-}\text{R}) \simeq d^\star(\mathbf{R}_j{-}\text{L})$ such that $\mathbf{R}_i{-}\text{R}$ and $\mathbf{R}_j{-}\text{L}$ are neighbours, the disparity $d^\star(\mathbf{R}_i{-}\text{L})$ is transferred to $\mathbf{R}_i{-}\text{R}$. And vice-versa.

2.2 Relaxation of Regional Disparities on the Fine Partition

A relaxation process is applied to the disparities initially measured on the fine partition. The process consists of detecting wrong disparity measures, ensuring that disparities evolve smoothly across neighbour regions whilst permitting discontinuities at objects frontiers. For that reason, the relaxation process is applied independently on each cluster of fine regions composing the same coarse region.

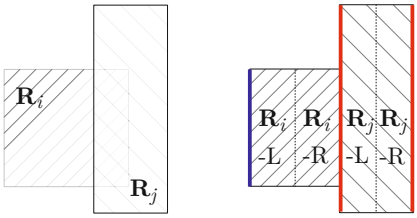


Fig. 2. Interpretation of regional disparities attributed to subregions when comparing the image gradients. Every region $\mathbf{R}_i$ is split along its vertical skeleton into left and right sub-regions $\mathbf{R}_i{-}\mathrm{L}$ and $\mathbf{R}_i{-}\mathrm{R}$. A region that is semi-occluded, in this example $\mathbf{R}_i$, has a piece of contour having a disparity equal to the disparity of the occluding region $\mathbf{R}_j$ which is likely to result in $d^\star(\mathbf{R}_i{-}\mathrm{R}) \simeq d^\star(\mathbf{R}_j{-}\mathrm{L})$. The disparity $d^\star(\mathbf{R}_i{-}\mathrm{L})$ is related to the physical frontier of $\mathbf{R}_i$ and constitutes the sole disparity representative of the displacement of $\mathbf{R}_i$.

The relaxation is implemented in the framework of Markov fields. The field's nodes represent the fine regions enclosed in the coarse region under study and the edges model the adjacency relationships between these regions. An objective function is assigned to every Markov field and is expressed as a sum of pairwise terms $P_{i,j}$ which grow quadratically with the difference between disparities assigned to neighbour regions $\mathbf{R}_i$ and $\mathbf{R}_j$, denoted as $d^\circ(\mathbf{R}_i)$ and $d^\circ(\mathbf{R}_j)$ respectively, and a sum of unary terms U_i which penalize the assignment of a disparity $d^\circ(\mathbf{R}_i)$ that strongly contradicts the initial measure $d^\star(\mathbf{R}_i)$. It is of course essential to determine the reliability $\alpha_i \in [0,1]$ of a disparity measure $d^\star(\mathbf{R}_i)$ and modulate the unary terms accordingly. To this end, we use the disparity map obtained for the coarse partition to localise both the fine regions which are likely to be severely occluded in the second image of the stereo pair and those whose disparity is significantly too far from the coarse regional disparity d_C. The measures attributed to these fine regions are assumed to have a low reliability. The pairwise and unary terms are defined by equations 2 and 3 respectively.

$$P_{i,j} \propto \left(d^\circ(\mathbf{R}_i) - d^\circ(\mathbf{R}_j)\right)^2 \tag{2}$$

$$U_i = \alpha_i |d^\star(\mathbf{R}_i) - d^\circ(\mathbf{R}_i)| + (1-\alpha_i)|d_C - d^\circ(\mathbf{R}_i)| \tag{3}$$

Finally, the disparity assignments which minimise the objective function are found using the minimum cut algorithm described in [10]. The effect of the relaxation process is illustrated in figures 3(b) and 3(c).

2.3 Regularisation of Disparity Maps

The regularisation of disparity maps is performed by ordinary kriging [2]. Given a set of samples for which the values are known as well as a variogram modelling the variability of the values taken by two points in terms of their distance, the kriging computes an unbiased estimator which minimises the variance of the estimation error.

Since there is no prior information regarding the expected depth map, we assume that the variability between two pixels in terms of disparity is proportional to their euclidean distance provided that these pixels are included in the same coarse segment. The kriging is applied independently on every coarse segment. The seeds originate from points belonging to the watershed of the coarse partition and from pics or holes enclosed in a region (cf. figure 3(e)). A correspondence in the second image is searched for each candidate seed. Only those having a disparity equal to the regional disparity of the fine region in which they are enclosed are eventually retained within the refinement process which yields the result presented in figure 3(f).

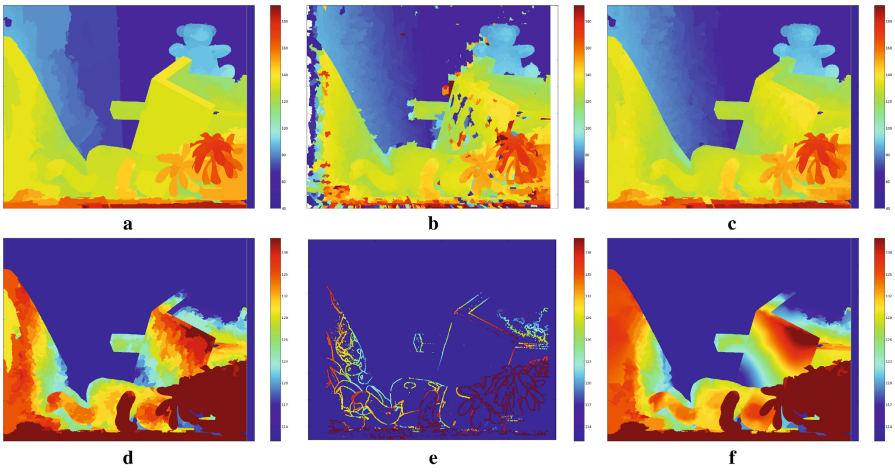


Fig. 3. The top-down approach to depth map computation. (a) Regional disparities of P_{C1} which serve as a basis for detecting wrong measures on the (b) brute regional disparity map of P_{F1}. (c) Regional disparities issued from the relaxation process on P_{F1} and (d) its visualisation with a restricted range of disparities. (e) Contour disparities along the watershed of the coarse segmentation and internal points disparities. (f) Interpolation result obtained by kriging.

3 Joint Stereo Segmentations Using Markers Transfer

The watershed-based segmentation [3] controlled by markers plays an essential role in this approach. This section is devoted to the mechanisms employed for extracting markers which delineate the salient objects in the scene, even if the image gradients suffer from leakages, and for computing equivalent segmentations between the images composing the stereo pair given the regional disparities and matching criteria presented in section 2.

3.1 Markers Extraction

The h-minima of a function g are defined as a binary function $M(g,h)$ which exclusively equals 1 at any point satisfying $\mathtt{R}^*_g(g+h)-g>0$, where the elevation

h is a positive constant and $\mathtt{R}^{*}_{g}(g+h)$ stands for the dual geodesic reconstruction of $g+h$ on top of the original function g. Taking the h-minima of a colour gradient yields markers which segment regions based on their frontier contrast. Although h has been fixed in all our experiments, it is worth mentioning that this elevation can be dynamically determined using a method similar to [12].

Markers originating directly from the h-minima have a severe inconvenient because two catchment basins of g merge as soon as a point with the smallest altitude on the watershed of g has been flooded. This favours premature fusions of h-minima as h increases when the gradient is subject to leakages. One way of preventing a flood at such leaking passages is to analyse the shape of the lakes resulting from the flooding induced by the h-minima. To this end, an adaptive erosion is applied on the h-minima and yields the marker set resulting from the indicator function in equation 4:

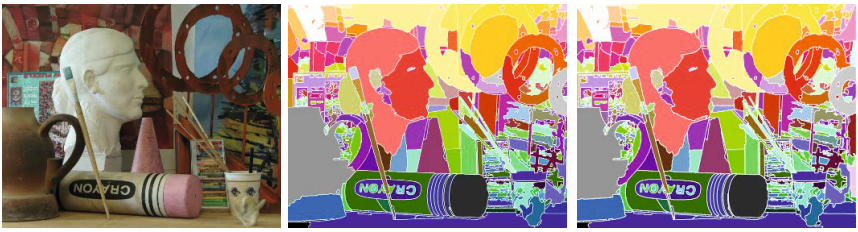
$$M_{\alpha}(g,h)[x,y] = \begin{cases} 1 & \text{if } (\mathcal{D} - \mathtt{R}_{\mathcal{D}}(\alpha\mathcal{D}))[x,y] > 0 \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

where $\mathcal{D}$ is the distance function computed by successive erosions on the marker set issued from $M(g,h)$, $\alpha \in [0,1[$ is a scaling factor controlling the intensity of the erosion and $\mathtt{R}_{\mathcal{D}}(\alpha\mathcal{D})$ stands for the geodesic reconstruction of the rescaled distance function under the original distance function. The adaptive erosion splits markers at narrow valleys with respect to their distance function. New markers are contained in the original h-minima and all h-minima can be reconstructed from the new marker set.

3.2 Stereo Equivalent Segmentations

The morphological co-segmentation consists of obtaining equivalent partitions between the images composing a stereo pair. In that context, the watershed segmentation driven by markers remains the tool of choice. The first experiments are presented in [3], where the idea is to propagate the markers obtained for the reference image to the second image. We now propose two mechanisms for obtaining equivalent segmentations of stereo images thanks to the transfer of labels onto image markers. The first one which is asymmetric relies on the regional disparities directly without being concerned about the matching criteria which makes it ideal for non-ideal stereo imagery. The second one which is symmetric revisits the matching criteria discussed in section 2 and in the current form applies only to images taken under the same illumination conditions. However, its mechanism identifies regions which are occluded in the reference image and attributes them a specific label which does not exist in the reference image partition.

Asymmetrical Transfer. In this algorithm, the transfer of markers is guided using the regional disparities presented in section 2. The algorithm first consists of estimating the equivalent partition of the second image, followed by the labelling of the gradient minima from which the watershed is eventually constructed.



(a) (b) (c)

Fig. 4. Asymmetric co-segmentation of (a) *Art* reference image from Middlebury database. (b) Reference image segmentation. (c) Equivalent segmentation of the second image. The labels attributed to the markers are preserved throughout the co-segmentation process and yields a matching between regions.

Given the partitioning of the reference image as a label map $\mathcal{L}_1$, such that $\mathcal{L}_1[x, y] = i \Leftrightarrow R_i[x, y] = 1$, the label map $\mathcal{L}_2$ is estimated according to equation 5.

$$\mathcal{L}_2[x, y] = \arg\max_i \{d^\star(\mathbf{R}_i) \times R_i[x + d^\star(\mathbf{R}_i), y]\} \tag{5}$$

Hence, each region $\mathbf{R}_i$ belonging to the reference image partition is shifted according to its regional disparity. The shifts are of different intensities implying therefore the existence of some overlaps between the shifted regions. In the case of an overlap, only the region which has the smallest depth can be visible to the camera. Hence $\mathcal{L}_2[x, y]$ is set to the label of the region which remains visible at (x, y). The other regions at that point can be marked as being occluded.

The equivalent segmentation is obtained by computing the watershed controlled by the minima of the gradient of the second image. The labels estimated in $\mathcal{L}_2$ are transferred to the markers, i.e. if (x, y) belongs to the minima, then the label $\mathcal{L}_2[x, y]$ is attributed to that point. The preservation of labels through the co-segmentation process yields the matching between the regions of the stereo partitions, as shown in figure 4.

The asymmetrical transfer has one limitation: it is impossible to represent regions in the second image which do not appear in the reference image. Such regions are directly merged in the equivalent segmentation to regions that are visible in the reference image. So another way to tackle the co-segmentation problem is to focus on a symmetrical transfer.

Symmetrical Transfer. The problem now comes down to relabelling the markers obtained independently for each image of the stereo pair. To achieve this, a sequence of back-and-forth label transfers is performed until no label changes after one of the transfer. Let M_{C1} and M_{C2} be the set of markers chosen for computing the coarse segmentations of the reference and the second image respectively. We define as $c(\mathbf{m}_i, \mathbf{m}_j)$ the cost of transferring the label of $\mathbf{m}_i \in M_{C1}$ to $\mathbf{m}_j \in M_{C2}$. A single way transfer consists of:

1. Defining the affinity cost $c(\mathbf{m}_i, \mathbf{m}_j^{(d)})$ between $\mathbf{m}_i$ and $\mathbf{m}_j$ shifted by d pixels. This cost is initialized to the mean of absolute differences between the image intensities across each intersection. If $\mathbf{m}_i \cap \mathbf{m}_j^{(d)} = \emptyset$, the cost is set to $+\infty$.
2. Searching for a set of markers $\mathbf{J}_i^{(d)} = \{\mathbf{m}_k^{(d)}\}$ for any $\mathbf{m}_k \in M_{C2}$ such that the Jaccard distance, as expressed in equation 1, between this union of markers and $\mathbf{m}_i$ is minimised.
3. Resetting $c(\mathbf{m}_i, \mathbf{m}_j^{(d)})$ to $+\infty$ if $\mathbf{m}_j^{(d)} \notin \mathbf{J}_i^{(d)}$ or $c_{\text{Jaccard}}(\mathbf{m}_i, \mathbf{J}_i^{(d)})$ is too high.
4. Computing the final affinity cost $c(\mathbf{m}_i, \mathbf{m}_j) = \min_d c(\mathbf{m}_i, \mathbf{m}_j^{(d)})$
5. Transferring the label of $\mathbf{m}_i$ to $\mathbf{m}_j$ if $c(\mathbf{m}_i, \mathbf{m}_j)$ is reasonably small.

In this procedure, we use the symmetric Jaccard distance, because the markers are chosen with the same flooding criterion and yield segmentations at the same scale of precision. However such markers can split between two stereo images as can be noticed in figure 5. For that reason, step 2 has been introduced in the matching procedure in order to prevent a penalty that would be due to a low Jaccard distance taken between the markers individually. Our procedure hence favours the transfer of a unique label to markers that have split. Step 5 ensures that a transfer can only occur when the regions that are covered by markers are similar in texture and colour. Costs exceeding 0.08 as a mean of the absolute differences of image intensities scaled between 0 and 1 do not generally lead to pertinent matchings.

(a) (b) (c)

Fig. 5. (a) Illustration of the generalised Jaccard distance. Here, $\mathbf{m}_i$ is a marker of the reference image. The set $\mathbf{J}_i^{(d)}$ that minimises $c_{\text{Jaccard}}(\mathbf{m}_i, \mathbf{J}_i^{(d)})$ is $\{\mathbf{m}_1^{(d)}, \mathbf{m}_2^{(d)}, \mathbf{m}_3^{(d)}\}$. (b)-(c) Using this distance within the symmetric transfer algorithm enables the label transfer from a marker of the source image to its corresponding markers in the target image.

4 Experimental Results and Evaluation

In this section, the results obtained on the Middlebury database are analysed in terms of precision. We also present an application of the proposed method to a particularly challenging stereo pair subject to a considerable amount of noise and homogeneous regions.

(a) (b) (c)

Fig. 6. Symmetric co-segmentation of *Teddy* reference image. (a) Reference image markers. (b) Reference image segmentation. (c) Second image segmentation. Regions of the second image which are occluded in the reference image now appear with new labels.

Accuracy. The precision evaluated on the Middlebury database [1] is presented in table 1. Pixels which are badly matched are those having a disparity which differ from a certain threshold with respect to the ground truth [11]. The main source of inaccuracy arises from highly slanted regions, like the ground in Teddy (cf. figure 3(c)). Searching for an optimal superimposition between such regions cannot be reasonably done by performing a simple shift which accounts for the observed instability. One should consider more complex geometrical transformations. Another source of imprecision comes from regions undergoing a severe image border occlusion, like in *Cones* or *Teddy*. Although these regions are detected thanks to the co-segmentation and then merged to neighbour regions according to the image gradient prior to the kriging process (cf. figure 3(f)), the linear variability model doesn't seem to be sufficient to guess the true disparities. Nevertheless, it is interesting to note that methods having the same level of precision usually produce disparity maps which are perceptually less appealing than ours obtained with the present method at the coarse level of the segmentation. The human sensibility to the contrast between the different depth planes is indeed predominant and this is not taken by the accuracy measurement into account.

Micro-stereopsis Imagery. Our method is compared to the semi-global estimation algorithm of [8] on a stereo pair having the following characteristics: low disparity range, acquisition under different illuminations, many homogeneous regions, noise, semi-transparent objects. The establishment of pixel correspondences is therefore particularly difficult and yields to ambiguities across homogeneous regions as shown in figure 7(c). The most pertinent disparities actually arise from the object contours but this information tends to be diffused from either side of the contour in [8]. Our regional disparity map is computed at a coarse segmentation level using the matching criterion based on gradient superimposition and occlusion reasoning from subregions disparities. We obtain the result shown in figure 7(d) which answers to the two aforementioned problems.

Table 1. Percentage of pixels which are badly matched with respect to different tolerances on the Middlebury database

	Tolerance ± 0.5 px.			Tolerance ± 1 px.			Tolerance ± 2 px.		
	NonOcc	*All*	*Disc*	*NonOcc*	*All*	*Disc*	*NonOcc*	*All*	*Disc*
Tsukuba	10.1	10.7	23.7	5.14	5.58	17.3	3.15	3.40	12.3
Venus	9.06	9.50	15.5	2.11	2.46	10.9	0.65	0.80	4.94
Teddy	14.2	20.2	28.6	7.38	15.8	20.8	4.25	7.76	10.7
Cones	13.5	18.8	23.0	6.84	11.9	14.0	3.80	8.04	8.36
Average	16.4			9.50			5.38		

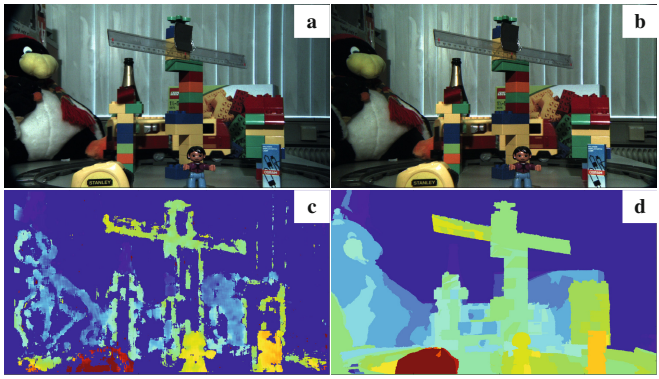


Fig. 7. Micro-stereopsis imagery. (a) reference and (b) second images of the stereo pair, (c) disparity map obtained with [8], (d) our coarse regional disparity map.

5 Conclusion

We have presented an approach to the estimation of depth map controlled by the matching and the superposition of regions. Our depth maps are estimated at two different scales of segmentation: first at a coarse level, then at a fine level. The final disparity maps are obtained by means of an interpolation process, the kriging, which relies on disparity emitter points having a disparity equal to the regional disparity on the fine partition. Points belonging to the watershed are also taken into account as soon as their membership to the appropriate region has been established. The approach strongly relies on the use of the watershed-based segmentation driven by markers in order to obtain the segmentation hierarchies and equivalent segmentations across stereo pairs.

Our method offers good results in terms of precision and perception. It also paves the way to the processing of non-ideal stereo images. Being able to choose the level of precision is of great interest for applications concerned by the process running-time. One could for instance restrain the refinement of disparities to regions having a low average depth only. Furthermore, the availability of efficient

watershed implementations based on hierarchical queues, the few nodes involved in each Markov field during the relaxation process and the fact that kriging only comes down to solving systems of linear equations lead to interesting perspectives for computationally efficient implementations of this global depth estimation method. Future work focuses on the exploitation of such mechanisms in the view of processing stereoscopic video sequences.

Acknowledgements. This work has been performed in the project PANORAMA, co-funded by grants from Belgium, Italy, France, the Netherlands, and the United Kingdom, and the ENIAC Joint Undertaking.

References

1. Middlebury stereo database, http://vision.middlebury.edu/stereo/
2. Armstrong, M.: Basic linear geostatistics. Springer (1998)
3. Beucher, S.: Segmentation d'Images et Morphologie Mathématique. Ph.D. thesis, Ecole Nationale Supérieure des Mines de Paris (1990)
4. Bleyer, M., Gelautz, M.: A layered stereo matching algorithm using image segmentation and global visibility constraints. ISPRS Journal of Photogrammetry and Remote Sensing 59(3), 128–150 (2005)
5. De-Maeztu, L., Villanueva, A., Cabeza, R.: Near real-time stereo matching using geodesic diffusion. IEEE Transactions on Pattern Analysis and Machine Intelligence 34(2), 410–416 (2012)
6. Fua, P.: A parallel stereo algorithm that produces dense depth maps and preserves image features. Machine Vision and Applications 6(1), 35–49 (1993)
7. Goshtasby, A.: Similarity and dissimilarity measures. In: Image Registration. Advances in Computer Vision and Pattern Recognition, Springer, London (2012)
8. Hirschmuller, H.: Stereo processing by semiglobal matching and mutual information. IEEE Transactions on Pattern Analysis and Machine Intelligence (2008)
9. Jaccard, P.: Bulletin de la société vaudoise des sciences naturelles. Tech. rep. (1901)
10. Prince, S.: Models for grids. In: Computer Vision: Models, Learning, and Inference. Cambridge University Press (2012)
11. Scharstein, D., Szeliski, R.: A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. International Journal of Computer Vision (2002)
12. Vilaplana, V., Marques, F., Salembier, P.: Binary partition trees for object detection. IEEE Transactions on Image Processing 17(11), 2201–2216 (2008)
13. Yamaguchi, K., Hazan, T., McAllester, D., Urtasun, R.: Continuous markov random fields for robust stereo estimation. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) ECCV 2012, Part V. LNCS, vol. 7576, pp. 45–58. Springer, Heidelberg (2012)
14. Yang, Q., Wang, L., Yang, R., Stewénius, H., Nistér, D.: Stereo matching with color-weighted correlation, hierarchical belief propagation, and occlusion handling. IEEE Transactions on Pattern Analysis and Machine Intelligence 31(3) (2009)
15. Zitnick, C.L., Kang, S.B.: Stereo for image-based rendering using image oversegmentation. International Journal of Computer Vision 75(1), 49–65 (2007)