A top-down methodology to depth map estimation controlled by morphological segmentation

Jean-Charles Bricola, Michel Bilodeau, Serge Beucher

A top-down methodology to depth map estimation controlled by morphological segmentation

Jean-Charles Bricola Michel Bilodeau Serge Beucher

MINES ParisTech, PSL - Research University,
CMM - Centre de Morphologie Mathématique,
35 rue St Honoré 77300 Fontainebleau, France

Abstract

Given a pair of stereo images and the transformation existing between the corresponding camera coordinate systems, the depth of a scene point can be computed from its projections on both images. Despite the difficulties related to the matching of such projections across homogeneous regions and the occlusion phenomenon, state of the art methods have already produced accurate results on classical stereo datasets. This article proposes a new way of approaching depth estimation. Instead of searching for dense pixel correspondences, a gross estimation of the disparities is initially performed at the region level, resulting in a regional disparity map which highlights the principal depth layers of the image. The disparity map is then systematically refined by considering finer partitions of the image. To this end, the watershed of the image colour gradient is selected in order to compute the image partitions alongside a meaningful hierarchy. We show that the ability to be driven by labelled markers enables the watershed algorithm to generate a co-segmentation of both stereo images given the regional disparities, which constitutes the main contribution of this paper. This co-segmentation allows one to reliably compute the disparities of pixels along the region contours. Finally, the contour disparities are transferred to the concerned regions after a careful analysis of their occlusion state with respect to each adjacent region. Though approximate, we show that the proposed method yields regional disparity maps which are close enough to ground truths in the view of performing the desired refinements. We also expose the perspectives of this methodology with respect to challenging stereo imagery, i.e. which is affected by noise or which contains a considerable amount of homogeneous regions.

1 Introduction

The problem of depth map computation from stereo dates back to the 70s and is traditionally solved by working out pixel correspondences between the images. Early approaches rely on feature points and contour matching in that respect [10, 16]. Based on a set of sparse measurements, the objective consists of finding the depth surface which best fits every sample.

Most current methods however concentrate on the establishment of dense correspondences. Local methods perform the matchings based on the dissimilarity costs of patches centred around the pixels forming the candidate match. There are two main difficulties: the choice of an appropriate dissimilarity measure, an exhaustive list of which can be found in [9], and the definition of the relevant matching support or weighting within the patch, which purpose is to ensure that more importance is given to the neighbours of the candidate pixel that belong to the same object. Aggregation of the matching costs may be performed to enforce consistency between scan-lines and is then followed by a refinement process. These methods are generally favoured when good runtime performance is required and produce very good results on standard datasets [7]. However, results are theoretically unpredictable across large and fully homogeneous regions due to the local nature of these approaches.

Global methods on the other hand aim at finding the best trade-off between the minimization of the matching costs, the preservation of features and the realization of smooth depth transitions, the smoothness being for example weighted by the negative of the image gradient in order to tolerate discontinuities along borders [8]. The dependencies of a pixel depth on its matching cost and neighbourhood depths can be modelled by Conditional Random Fields for which the associated energy is minimized by means of minimum cuts or belief propagation [19]. The resulting algorithms are usually much slower than the local ones [6]. It is however possible to approximate the minimization result by using semi-global approaches gathering and aggregating the optimal results along different search lines without affecting the end result [12].

The exploitation of regions within depth estimation algorithms is progressively gaining popularity. Related approaches may be distinguished between super-pixels and layer-based categories. In the first case, one is interested in matching super-pixels directly whilst the smoothness of depth is enforced across neighbouring regions having close colour distributions [20]. In the second case, the estimation process is guided by depth layers for which the planar equations are estimated [5, 19] or which serve as defining the object boundaries where sharp depth discontinuities are permitted [1]. Finally, regions are of a particular interest when dealing with occluded pixels [15, 14] as they enclose pixels which turn out to be visible in both images and which are likely to share similar depth when they are sufficiently close.

A region may be described according to its contour and the feature points it contains. However, when regions are perfectly homogeneous or noise interferes with the detection of feature points, contours provide the only clues of depth. The diffusion of contour disparities

to regions requires to take care of the state of a given contour with respect to each of its adjacent region: does it stand for an occlusion border or a physical frontier? This information is of vital importance in the context of regularization and has been little addressed so far, apart from [18] which exploits boundaries junctions to this end.

In this work, a depth map is generated by systematic depth refinements at every stage of the process. The reference image is first segmented using a marker driven watershed segmentation as presented in section 2. The markers are automatically extracted so as to highlight the salient objects in the scene and prevent region fusion caused by leaking gradients. The saliency is moreover controlled so as to obtain hierarchical segmentations of the same image. Section 3 presents the concept of regional disparity which measures the average displacement of a region across two stereo images and explains its relationship with the actual depth being searched for. Section 4 shows how the regional disparities are computed over fine levels of the segmentation using the results obtained for the coarse partition as a prior. Building up on the estimation of regional disparities obtained for the fine partition, a co-segmentation of both stereo images is produced (section 5) and the region related contour features are extracted (section 6). The results are presented for each step of the proposed method and perspectives are discussed in section 7.

Hereafter, it is assumed that the images constituting the stereo pair are rectified, which means that the scene points projections are vertically aligned but have different abscissa. The measured difference in abscissa corresponds to the disparity and is inversely proportional to the depth [11]. The image which the depth if estimated for is referred to as the reference image, whilst the other image is denoted as the second image.

Symbol	Description
Morphological operators	
B and H are structuring elements	
H is an isotropic structuring element of elementary size	
$\delta_B(f)$	Dilation of f: $\delta_B(f)[x] = \sup_{x \in B} f[x]$
$\varepsilon_B(f)$	Erosion of f: $\varepsilon_B(f)[x] = \inf_{x \in B} f[x]$
$D_g^1(f)$	Geodesic dilation of f under mask g: $D_g^1(f) = \inf(\delta_H(f), g)$
$\mathtt{R}_g(f)$	Geodesic reconstruction of g from marker f: $\mathtt{R}_g(f) = D_g^{+\infty}(f) = D_g^1\left(\ldots\left(D_g^1(f)\right)\ldots\right)$
$\mathtt{R}_g^*(f)$	Dual geodesic reconstruction of g from marker f: $\mathtt{R}_g^*(f) = -\mathtt{R}_{-g}(-f)$

Table 1. Notation

2 Reference Image Segmentation

The watershed algorithm partitions an image into a series of connected components satisfying each an homogeneity criterion. They are obtained by flooding a topological surface, typically the colour gradient of the image, from a set of markers. Each marker gives birth to a lake which a unique label is assigned to and each lake in turn results in one region. As the flooding goes on and uniformly increases in altitude, watersheds are constructed so as to prevent lakes with different labels from merging. The interested reader will find the details of the watershed construction in [2] as well as practical details for an effective implementation based on hierarchical priority queues in [4].

The segmentation result is sensitive to the image gradient quality and the choice of appropriate markers. A colour gradient expressed as the supremum of the individual hue, saturation and luminance gradient magnitudes is used whilst the markers are computed from the h-minima of the colour gradient. The rest of this section focuses on the generation of segmentation markers.

2.1 Segmentation markers

In mathematical morphology, a marker is not limited to a single pixel but usually represents a connected component. Markers control the behaviour of the watershed segmentation: they should preferably lie at the bottom of the topological surface being flooded and discriminate every region to be segmented. We derive here a method of marker extraction which processes the h-minima of the colour gradient.

2.1.1 h-minima

Let g be a function mapping a point in 2d space to a real number. x belongs to the h-minima of g if and only if there exists a path $\Gamma : [0,1] \to \mathbb{R}$ from $x = \Gamma(0)$ to one of the minima of g, say $x_0 = \Gamma(1)$, such that $g(\Gamma(s_i)) \geq g(\Gamma(s_j))$ for every $s_i < s_j$ and $g(x) - g(x_0) < h$. Hence,

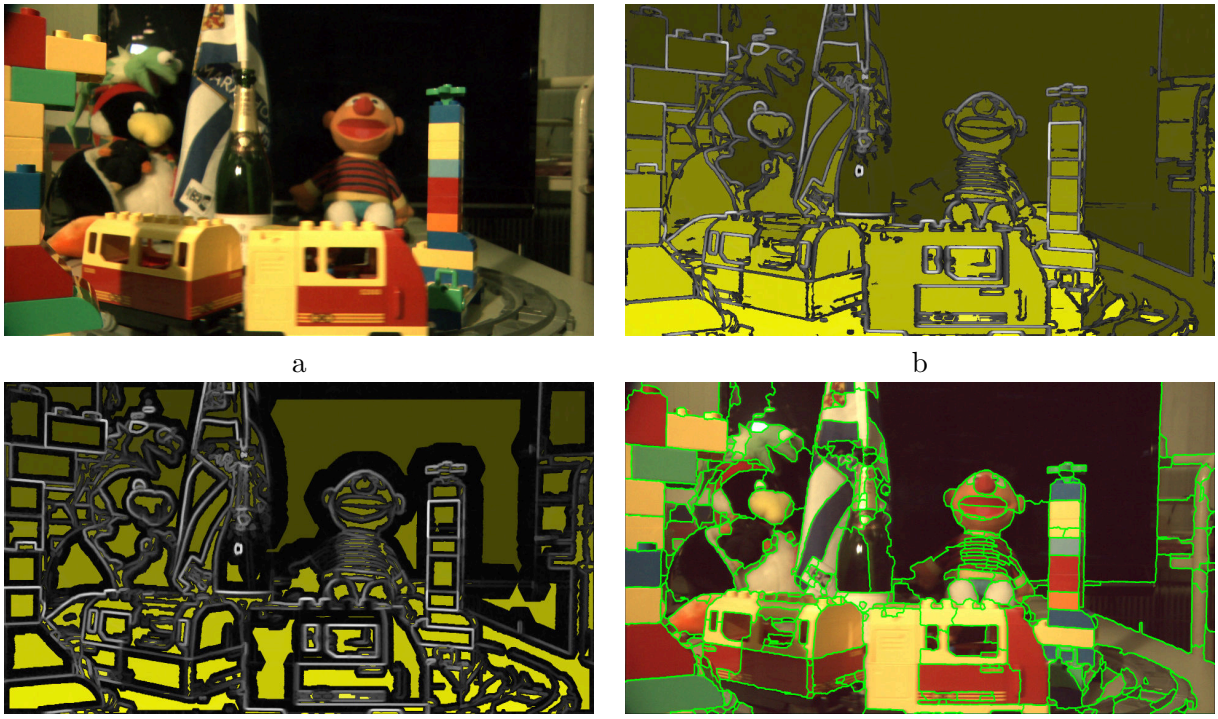


Figure 1. (a) Input image, (b) The h-minima of the input image colour gradient are represented over the image gradient. The main structures of the images are highlighted by the resulting indicator function, however there exists markers which target more than one region due to leakages in the gradient. (c) In order to circumvent this limitation, an adaptive erosion is applied on the h-minima. This way, small markers are preserved whilst markers split where the gradient leaks. (d) Choosing those markers leads to the watershed obtained for the coarse partition represented over the input image.

the h-minima of g are given by the following indicator function:

$$\mathcal{M}(g,h)[x] = \begin{cases} 1 & \text{if } \left(\mathtt{R}^*_g(g+h) - g\right)[x] > 0 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

The operator expressed in equation 1 is extensive with respect to h, which implies that $\mathcal{M}(g,h_i) \subset \mathcal{M}(g,h_j)$ for all $h_i < h_j$ as proved in [3]. In other words, increasing h yields markers which are bound to enforce the contrast between the regions they designate. Figure 1(b) shows the result of the h-minima for h sufficiently large. One can notice that gradient leakages cause undesired marker overlaps between two target regions.

2.1.2 Adaptive erosion on h-minima

Figure 1(c) shows the effect of applying an adaptive erosion on the h-minima of the gradient. The objective is to encourage markers splitting at locations where markers get thinner due to

gradient leakages. A distance function d is computed by construction of successive erosions applied on the binary mask corresponding to the h-minima of g, i.e. $\mathcal{M}(g, h)$. Thin structures belonging to the mask corresponds to valleys in function d. New markers are therefore given by the connected components resulting from the indicator function

$$R(d, \alpha)[x] = \begin{cases} 1 & \text{if } (d - \mathtt{R}_d(\alpha d))[x] > 0 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

where $0 \leq \alpha < 1$ controls the sensitivity to the valleys depth. Small values of α hence limit the splitting operation to the deepest valleys of d only. This erosion is said to be *adaptive* because it preserves small markers, contrary to fixed size erosions or openings.

2.1.3 Fine partition markers

In order to exploit hierarchical relationships between the image coarse and fine partitions, every region belonging to the fine partition must be included in only one region of the coarse partition. Lowering h within the computation of h-minima allows one to highlight new regions which are less contrasted, but when used with an adaptive erosion, an additional processing is required to satisfy the aforementioned constraint. It is essential to impose the watershed lines of the coarse partition on the fine partition: hence, the topological surface being flooded for the generation of the fine partition is an augmented colour gradient reaching the maximum possible value at locations covered by the coarse watershed. Finally, amongst the fine partition markers, at least one marker must be included in every cells forming the coarse partition. Such verification can be performed by a reconstruction of the coarse partition from the markers chosen for the fine partition. Cells which have not been reconstructed simply inherit from the corresponding coarse partition marker.

3 Regional Disparities

The regional disparity is a measure assigned to every region of a reference image partition. It quantifies the relative displacement of each region across the stereo image pair.

Let $\mathcal{R} = \{R_1, \cdots, R_n\}$ be the set of regions obtained for the reference image. $R_i(x, y)$ is an indicator value that equals 1 for every pixel (x, y) belonging to region R_i, 0 otherwise. Also let g_R and g_S be the colour gradients of the reference and second images respectively. For each region R_i, the translation $\mathbf{t}^{(i)}$ applied on g_S that provides the best overlap of the reference and second image gradients within R_i is estimated. As stereo images are assumed to be rectified, the translation only occurs along the x-axis, so $\mathbf{t}^{(i)} = (d^{(i)}, 0)$. Optimal translations are determined by equation 3.

$$d^{(i)} = \arg\min_{d} \sum_{(x,y)|R_i(x,y)=1} |g_R(x, y) - g_S(x - d, y)| \tag{3}$$

Assuming that all points in the second image have an abscissa smaller than or equal to their correspondence in the reference image, the regional disparity obtained for region R_i is $d^{(i)}$.

3.1 Interpretation

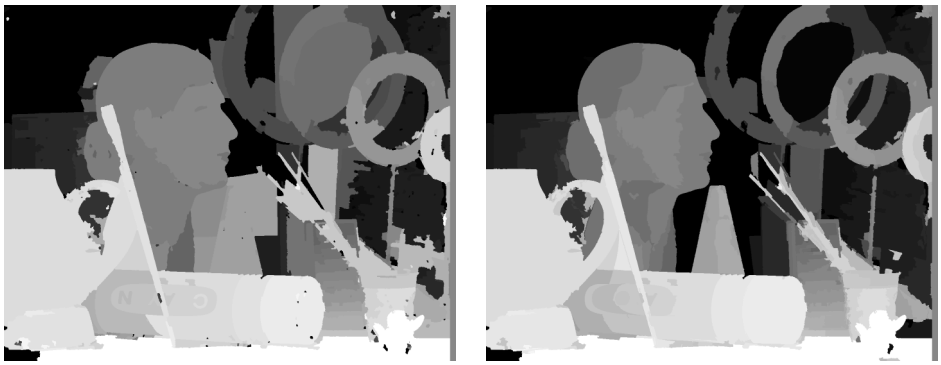
The link between regional disparities and pixel disparities can be explained as follows:

- Transferring the regional disparity of a region to its pixels is valid if the region is fronto-parallel to the camera and the region contours are representative of the region physical frontier.
- The regional disparity of regions representing holes within objects is likely to inherit from the disparities of occlusion contours, in particular if there is no gradient information inside the hole, and thus cannot be used to provide an estimate of the disparities inside those regions, but around the holes.
- The regional disparity of slanted regions belongs to the range of the actual pixel disparities. They may only serve as a gross estimate of the true disparities.

When regions undergo semi-occlusions preserving at least a piece of frontier contour, several optimal overlaps of the stereo gradients are plausible with respect to equation 3, which may lead to incorrect regional disparities. It is however possible to compute a rectified regional disparity map to take these semi-occlusions into account.

3.2 Rectified regional disparities

The rectification algorithm proposed below aims at detecting occlusion contours and rectifying the mistaken regional disparities accordingly. It first assigns to every region two disparities:



a b

Figure 2. Comparison between (a) brute and (b) rectified regional disparities which handle semi-occluded regions on the coarse partition of *Middlebury Art* image [13].

one that is representative of the disparity near the left contour of the region, the other of the right contour. To do so, regions obtained for the reference image are vertically split into left and right subregions and regional disparities for these subregions are estimated. Second, the following neighbourhood relationships between regions are defined: $\mathcal{N}_l(R_i)$ includes all regions $R_j \in \mathcal{R}$ that are adjacent to the left subregion of R_i and the same goes for $\mathcal{N}_r(R_i)$ with respect to the right subregion of R_i. Third, algorithm 1 is executed for every region R_i:

Algorithm 1 Regional disparities rectification

function RectifyDisparities($\mathcal{R}$, i)
  $d_l \leftarrow$ `getLeftContourDisparity`(R_i)
  $d_r \leftarrow$ `getRightContourDisparity`(R_i)
  if $d_l > d_r + \tau$ **then**
    for all $R_j \in \mathcal{N}_l(R_i)$ **do**
      $d_r^{(j)} \leftarrow$ `getRightContourDisparity`(R_j)
      if $|d_l - d_r^{(j)}| < \tau$ **then**
        $d_l \leftarrow d_r$
      end if
    end for
  else if $d_r > d_l + \tau$ **then**
    for all $R_j \in \mathcal{N}_r(R_i)$ **do**
      $d_l^{(j)} \leftarrow$ `getLeftContourDisparity`(R_j)
      if $|d_r - d_l^{(j)}| < \tau$ **then**
        $d_r \leftarrow d_l$
      end if
    end for
  end if
  return (d_l, d_r)
end function

The intuition behind this algorithm is straightforward: when a region is assigned to sig-

nificantly different left and right disparity values according to the chosen threshold τ (i.e. when algorithm condition 4 or 11 is fulfilled), either the highest disparity value is due to an occluding contour or the region is simply not fronto-parallel. In the case of an occlusion, one checks whether condition 7 or 14 holds. In the affirmative, the disparity that is the smallest between the left and right disparities is assigned to the entire region.

4 Refinement of regional disparity maps

The computation of regional disparities on fine partitions significantly improves the quality of the resulting disparity maps when input images contain a lot of texture. Fine partitions though are prone to include parasite regions subject to total occlusions. In order to estimate the regional disparity for each of these regions, one has to take account of the hierarchical relationship that exists between coarse and fine regions. For every occluded region, the matching support is transferred to the close non-occluded fine regions which are part of the same gross region. The occlusion areas are detected using the regional disparity map computed for the coarse partition. Finally a relaxation process is applied on the resulting regional disparities in order to enforce their smooth evolution across each coarse region.

4.1 Evaluation

Figure 3 shows the regional disparity maps obtained for the fine partitions on classical stereo datasets. If these disparity maps were used as the final result, the corresponding overall error rate with respect to ground truth would equal 4.27%, 6.92% and 9.20% with an error tolerance of 2 pixels for the disparities obtained on the Tsukuba, Cones and Teddy test cases respectively.

As expected, the most important errors occur along homogeneous slanted regions and reveal the limits of regional disparities. Nonetheless, the latter provide the necessary information for the extraction of depth-ordered layers and the generation of equivalent stereoscopic segmentations described in section 5. Those two ingredients play a central role in the computation of region contour disparities and their attribution to the concerned regions as shown in section 6.

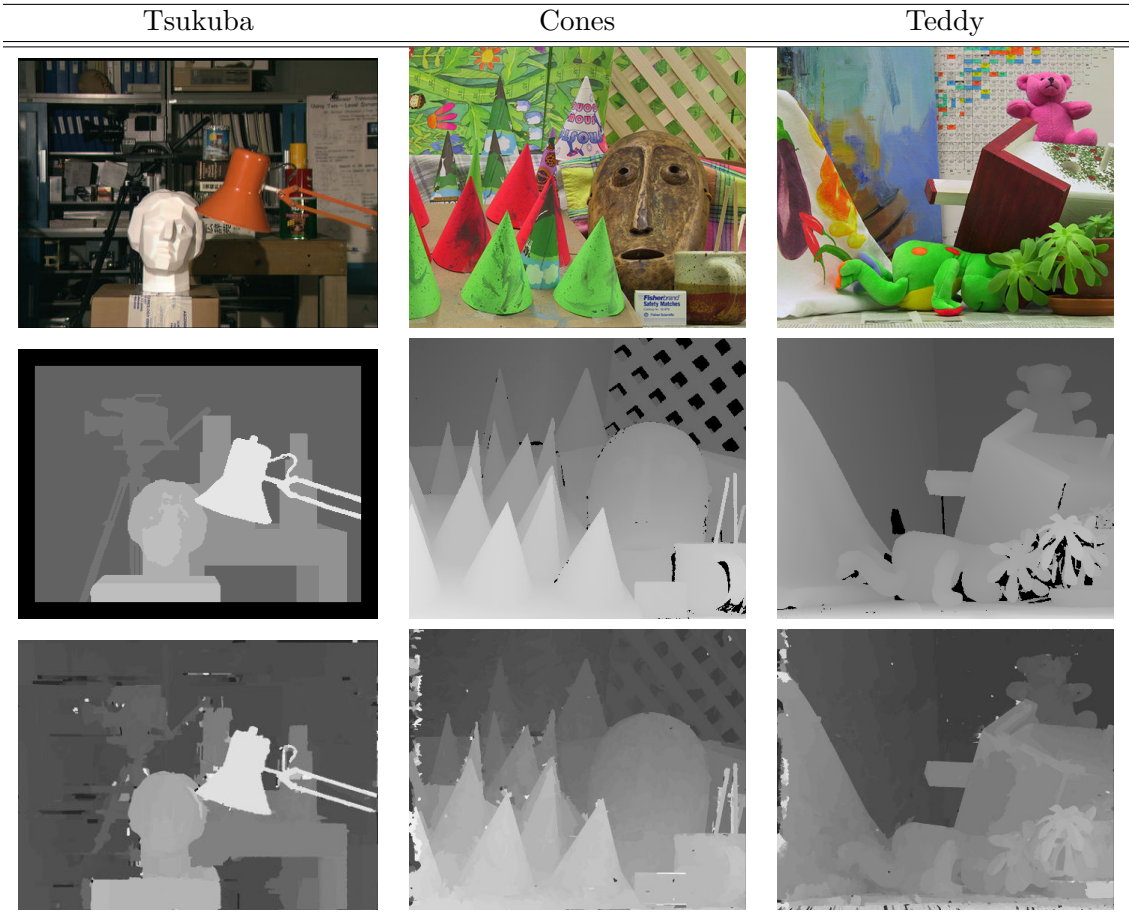
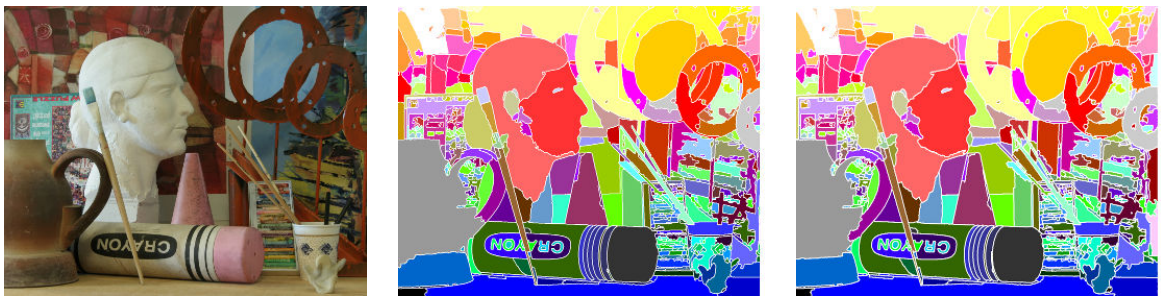


Figure 3. Regional disparity maps obtained for the fine partitions of three classical images from Middlebury stereo database [17]. Top row: input images, middle row: ground truth (black pixels correspond to unknown disparities), bottom row: regional disparity maps.



a b c

Figure 4. Morphological co-segmentation of the *Middlebury Art* image induced by regional disparities. (a) Reference image, (b) Coarse partition obtained for the reference image, (c) Equivalent partition obtained for the second image of the stereo pair. The use of labelled markers immediately yields a matching between the regions of the stereo pair.

5 Morphological Co-Segmentation

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 on the matter were presented in [2]. The key idea is to propagate the markers obtained for the reference image to the second image. The rectified regional disparities can therefore be used in that respect as shown in figure 4.

Given the partition map $\mathcal{L}_R$ of the reference image, such that $\mathcal{L}_R(x,y) = i \Leftrightarrow R_i(x,y) = 1$, the second image partition $\mathcal{L}_S$ is estimated according to the computed regional disparities. Regions of the reference image are propagated to the second image according to equation 4

$$\mathcal{L}_S(x,y) = \arg\max_i \left\{ d^{(i)}\ R_i(x - d^{(i)}, y) \right\} \tag{4}$$

In other words, when several pixels overlap due to the translations of different intensities applied on the reference image regions, the pixel which has the lowest depth (and hence the highest disparity with respect to the considered stereoscopic configuration) is therefore the only one to be eventually transposed to partition $\mathcal{L}_S$. Those which have not been propagated are therefore marked as being occluded in the second image of the stereo pair.

Once the partition has been obtained for the second image of the stereo pair, an adaptive erosion is applied on each cell of the partition so as to create markers which do not overlap with image borders. The markers obtained after this erosion are however not relabelled. The watershed segmentation is then computed using these markers and the colour gradient of the second image. Finally, the marker labelling inherited from the region transfer immediately yields the matching between the stereo regions.

Symbol	Description
Transfer of contour disparities	
h	one the 8 directions of the square image grid
B_h	Structuring element of direction h
D_R	Regional disparity map of reference image, s.t. $D_R(x,y) = d^{(i)} \Leftrightarrow R_i(x,y) = 1$
W_R	Binary watershed function of the reference image $W_R(x,y) = 1$ iff (x,y) belongs to watershed
W_C	Contour disparities available along the watershed $W_C(x,y)$ equals the contour disparity at pixel (x,y) if available, $\perp$ otherwise

Table 2. Notation

6 Contour disparities

This section focuses on the extraction and use of disparities which can be recovered along the watershed of the reference image.

6.1 Contour point matching

The matching of contour points is performed for each scanline of equation $y = y_s$ independently and relies on dynamic programming. Let $\mathcal{L}_R$ and $\mathcal{L}_S$ be the labelled partitions and $\mathcal{I}_R = \{x_1, \ldots, x_m\}$ and $\mathcal{I}_S = \{x'_1, \ldots, x'_n\}$ denote the ordered lists of the intercepted contour points abscissa of the reference and second images respectively.

The cost of matching $x \in \mathcal{I}_R$ to $x' \in \mathcal{I}_S$ depends on the corresponding points descriptors which originate from the co-labelled partitions and is expressed by equation 5:

$$c(x,x') = \begin{cases} 0 & \text{if for all } h,\ \delta_{B_h}\left(\mathcal{L}_R(x,y_s)\right) = \delta_{B_h}\left(\mathcal{L}_S(x',y_s)\right) \\ & \text{and } \varepsilon_{B_h}\left(\mathcal{L}_R(x,y_s)\right) = \varepsilon_{B_h}\left(\mathcal{L}_S(x',y_s)\right) \\ +\infty & \text{otherwise} \end{cases} \tag{5}$$

The optimal matching cost accumulated along the scanline is then obtained by evaluating the recursive relation defined by equation 6 at $C(m,n)$ for any occlusion cost $c_{occl} > 0$:

$$\begin{aligned} C(i,j) &= \min\{\, C(i-1,j-1) + c(x_i,x'_j), \\ &\qquad C(i-1,j) + c_{occl}, \\ &\qquad C(i,j-1) + c_{occl} \,\} \\ C(0,0) &= c(x_0,x'_0) \end{aligned} \tag{6}$$

Backtracing the arguments of the minimization that yields the accumulated cost $C(m,n)$ provides the matching result from which contour disparities are retrieved for every points

a b

Figure 5. Visualization of contour disparities obtained for *Teddy* under the full range of disparities (a) and a restricted range of disparities (b). One can notice the evolution of disparities along the contours of the shed's roof which were previously not captured by the regional disparities.

which have not been classified as occlusions in one of the stereo images.

Of course, the proposed algorithm enforces the ordering constraint. This property is desirable, but for contour points that belong to the same depth layer only. It is however possible to perform the matchings for each depth layers independently, given the regional disparity maps.

Figure 5 shows an example of these contour disparities computed over the Teddy image. In comparison with figure 3, the evolution of depth along the borders of the shed is now clearly visible. The result is however sensitive to the quality of the watershed: in particular false contours should not be used. Note that such false contours can be removed by using the first level of hierarchical algorithms, such as the waterfalls [2].

6.2 Occlusion reasoning and disparity transfer

Before performing any regularization, an essential question remains: which are the regions concerned by the previously computed contour disparities? A piece of contour always represents the physical frontier of at least one of the adjacent regions. If one of the adjacent regions were occluded, then its regional disparity should be inferior to the one of the non-occluded region.

Based on that observation, we devise an algorithm which transfers the contour disparities of the watershed to the neighbour pixels of the appropriate region(s). Table 2 describes the input variables used in the proposed algorithm. In order to detect the surrounding pixels of the watershed that belong to the region of lowest depth, we compute for each direction h the

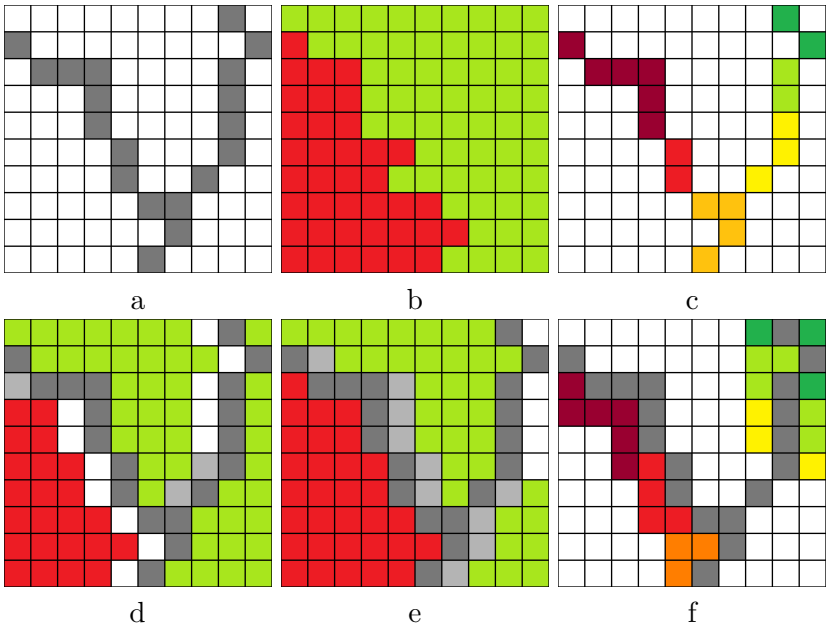


Figure 6. Disparity transfer from contour to regions. (a) Binary watershed W_R highlighted with gray pixels, (b) Regional disparity map D_R for which red pixels have a higher disparity than green pixels, (c) Contour disparities W_C, (d) The directional transfer of disparities $P_{\leftarrow}$ and (e) $P_{\rightarrow}$ occurs for pixels shown in white, (f) The resulting transfer inside appropriate regions yields the measurements enclosed in D_{contour}.

detection map P_h given by the indicator function described in equation 7:

$$\begin{aligned}
I_1(x,y) = 1 \quad &\Leftrightarrow \quad W_C(x,y) \geq 0 \\
F_1 \quad &= \quad D_R - \varepsilon_{B_h}\left(\varepsilon_{B_h}\left(D_R\right)\right) \\
F_2 \quad &= \quad \delta_H\left(\delta_H\left(D_R\right)\right) - \varepsilon_H\left(\varepsilon_H\left(D_R\right)\right) \\
I_2 \quad &= \quad \delta_{B_h}\left(I_1\right) - W_R \\
I_3(x,y) = 1 \quad &\Leftrightarrow \quad F_1(x,y) > 0 \\
I_4(x,y) = 1 \quad &\Leftrightarrow \quad F_2(x,y) = 0 \\
P_h \quad &= \quad I_2 \cap \left(I_3 \cup I_4\right)
\end{aligned} \tag{7}$$

such that $I_1, \ldots, I_4$ represent binary functions. The map holding the transferred contour disparities is finally given by equation 8.

$$D_{\text{contour}} = \begin{cases} \frac{\sum_h P_h \, \delta_{B_h}(W_C)}{\sum_h P_h} & \text{where } \sum_h P_h > 0 \\ \perp & \text{otherwise} \end{cases} \tag{8}$$

7 Conclusion

We have proposed a methodology of depth map estimation which, instead of trying to establish dense pixel correspondences from scratch, reasons about the displacements of gross regions first, refines the analysis at a finer degree of segmentation and deduces the disparities of contour points. The regularization step that follows is outside the scope of this paper but is intended as future work.

In this approach, the morphological segmentation is the prime mechanism for obtaining segmentation hierarchies and equivalent segmentations of stereo images. The latter drives the matching of contour points and is bound to remain robust in the presence of noise or many homogeneous regions, as opposed to features points or dense correspondences, which is the primary motivation of the region-based approach. However, in the case of more ideal stereo imagery, nothing prevents the use of regional disparities as a prior for feature points extraction which could add valuable information to the regularization process.

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] T. Aydin and Y. S. Akgul. Stereo depth estimation using synchronous optimization with segment based regularization. *Pattern Recognition Letters*, 31(15):2389–2396, 2010. 2

[2] S. Beucher. *Segmentation d'Images et Morphologie Mathématique*. PhD thesis, Ecole Nationale Supérieure des Mines de Paris, 1990. 4, 12, 13

[3] S. Beucher. Maxima and minima: a review. Technical report, Center of Mathematical Morphology - Mines ParisTech, 2013. 5

[4] S. Beucher and N. Beucher. Hierarchical queues: general description and implementation in mamba image library. Technical report, Center of Mathematical Morphology - Mines ParisTech, 2011. 4

[5] M. Bleyer and M. Gelautz. A layered stereo matching algorithm using image segmentation and global visibility constraints. *ISPRS Journal of Photogrammetry and Remote Sensing*, 59(3):128–150, 2005. 2

[6] M. Bleyer, C. Rother, P. Kohli, D. Scharstein, and S. Sinha. Object stereo – joint stereo matching and object segmentation. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 3081–3088. IEEE, 2011. 2

[7] L. De-Maeztu, A. Villanueva, and R. Cabeza. Near real-time stereo matching using geodesic diffusion. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 34(2):410–416, 2012. 2

[8] P. Fua. A parallel stereo algorithm that produces dense depth maps and preserves image features. *Machine vision and applications*, 6(1):35–49, 1993. 2

[9] A. Goshtasby. Similarity and dissimilarity measures. In *Image Registration*, Advances in Computer Vision and Pattern Recognition, pages 7–66. Springer London, 2012. 2

[10] M. J. Hannah. *Computer Matching of Areas in Stereo Images*. PhD thesis, Stanford, CA, USA, 1974. 2

[11] R. Hartley and A. Zisserman. *Multiple view geometry in computer vision*. Cambridge University Press, 2004. 3

[12] H. Hirschmuller. Stereo processing by semiglobal matching and mutual information. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(2):328–341, 2008. 2

[13] H. Hirschmuller and D. Scharstein. Evaluation of cost functions for stereo matching. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pages 1–8. IEEE, 2007. 8

[14] S. Huq, A. Koschan, and M. Abidi. Occlusion filling in stereo: Theory and experiments. *Computer Vision and Image Understanding*, 117(6):688–704, 2013. 2

[15] S. B. Kang, R. Szeliski, and J. Chai. Handling occlusions in dense multi-view stereo. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, volume 1, pages I–103. IEEE, 2001. 2

[16] Y. Ohta and T. Kanade. Stereo by intra-and inter-scanline search using dynamic programming. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (2):139–154, 1985. 2

[17] D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International journal of computer vision*, 47(1-3):7–42, 2002. 11

[18] K. Yamaguchi, T. Hazan, D. McAllester, and R. Urtasun. Continuous markov random fields for robust stereo estimation. In *Computer Vision–ECCV 2012*, pages 45–58. Springer, 2012. 3

[19] Q. Yang, L. Wang, R. Yang, H. Stewénius, and D. Nistér. Stereo matching with color-weighted correlation, hierarchical belief propagation, and occlusion handling. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 31(3):492–504, 2009. 2

[20] C. L. Zitnick and S. B. Kang. Stereo for image-based rendering using image over-segmentation. *International Journal of Computer Vision*, 75(1):49–65, 2007. 2