

# MGM: A Significantly More Global Matching for Stereovision

Gabriele Facciolo, Carlo De Franchis, Enric Meinhardt

▶ **To cite this version:**

Gabriele Facciolo, Carlo De Franchis, Enric Meinhardt. MGM: A Significantly More Global Matching for Stereovision. BMVC 2015, 2015, Swansea, United Kingdom. 10.5244/C.29.90 . hal-01240853

**HAL Id: hal-01240853**
**https://enpc.hal.science/hal-01240853**

Submitted on 10 Dec 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# MGM: A Significantly More Global Matching for Stereovision

Gabriele Facciolo[1]
facciolo@cmla.ens-cachan.fr
Carlo de Franchis[2]
carlo.de-franchis@ens-cachan.fr
Enric Meinhardt[2]
enric.meinhardt@cmla.ens-cachan.fr

[1] IMAGINE/LIGM,
École Nationale des Ponts et Chaussées,
France
[2] CMLA,
École Normale Supérieure de Cachan,
France

**Abstract**

Semi-global matching (SGM) is among the top-ranked stereovision algorithms. SGM is an efficient strategy for approximately minimizing a global energy that comprises a pixel-wise matching cost and pair-wise smoothness terms. In SGM the two-dimensional smoothness constraint is approximated as the average of one-dimensional line optimization problems. The accuracy and speed of SGM are the main reasons for its widespread adoption, even when applied to generic problems beyond stereovision. This approximate minimization, however, also produces characteristic low amplitude streaks in the final disparity image, and is clearly suboptimal with respect to more comprehensive minimization strategies.

Based on a recently proposed interpretation of SGM as a min-sum Belief Propagation algorithm, we propose a new algorithm that allows to reduce by a factor five the energy gap of SGM with respect to reference algorithms for MRFs with truncated smoothness terms. The proposed method comes with no compromises with respect to the baseline SGM, no parameters and virtually no computational overhead. At the same time it attains higher quality results by removing the characteristic streaking artifacts of SGM.

## 1 Introduction

Stereovision estimates the depth of a scene from two or more images taken from slightly different viewpoints. This is done by computing the apparent motion of the scene points between the views. In the case of *stereo-rectified* image pairs [10] this motion is referred to as *disparity*.

Stereo matching methods are traditionally divided [20] into *local* and *global methods*. Local methods estimate the disparity independently for each pixel by comparing features (usually a window around the pixel) of the left and right image. Local methods are computationally cheap, however if the comparison window falls on an ambiguous area (lacking texture or with repetitive patterns) the estimated disparity will likely be incorrect. *Global methods* cope with these ambiguities by imposing the smoothness of the disparity map, which permits to derive reasonable estimates even in the ambiguous areas.

© 2015. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

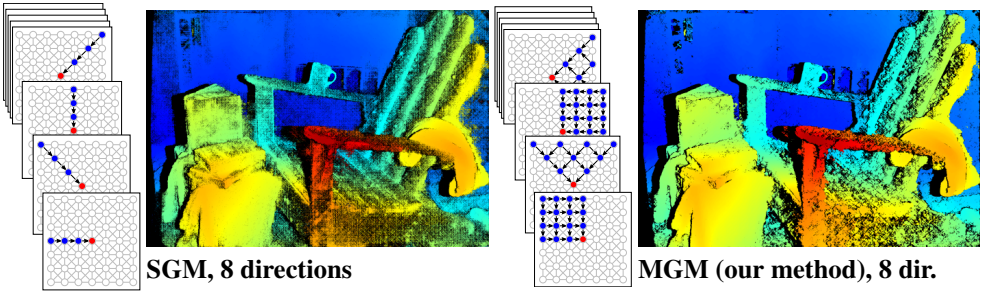



Figure 1: Results of our baseline implementation of SGM and the proposed method MGM for the *Adirondack* pair (figure 5). The diagrams on the left depict for each method the information used by the recursive update of the red pixel in each of the 8 scans of the algorithms.

Global methods model the correspondence estimation as a global minimization usually associated with a Markov Random Field (MRF) formulation. The resulting energy functional is a sum of unary matching terms and smoothness terms that force adjacent pixels to have similar disparities. While this problem is NP-hard, strong local minima can be computed [9] and many techniques have been developed to that effect, such as move-making approaches [4, 15] and message-passing methods [7, 14]. Their effectiveness has been reviewed in [24]. Most of these techniques, however, are too slow when applied to images of a reasonable size.

Because of the ever increasing size of the problems, fast approximate algorithms producing reasonable solutions are currently preferred to global techniques [1, 5, 9, 12, 16, 19, 22, 26]. Among these approximate algorithms, *semi-global matching* (SGM) [12] is nowadays one of the preferred choices for stereovision applications [11, 29] because of its efficiency and good performance, and it has even been applied to problems beyond disparity estimation, for example, to approximately solve the Potts model [22].

In SGM the two-dimensional smoothness constraint is efficiently approximated as the average of one-dimensional line optimization problems. This approximation reduces at each pixel to an optimization on a star-shaped graph (usually with 8 cardinal directions) centered at the pixel. This structure favors the occurrence of piecewise constant solutions along these directions. However, two adjacent scan lines share little information, so if the data term is weak, their solutions can differ, yielding the well known streaking artifacts of SGM.

The success the SGM algorithm relies on several heuristics. In this paper we clarify some of them by analyzing SGM in the light of its recently proposed interpretation as a min-sum Belief Propagation algorithm by Drory *et al*. [6]. This leads us to propose a new version of the algorithm that overcomes some of its limitations. Our principled interpretation of SGM reduces by a factor five its energy gap with respect to reference optimization algorithms for MRFs with truncated smoothness terms.

Extensive evaluation shows that the proposed algorithm removes all streaking artifacts, improves the visual quality of the result, and computes efficiently approximate solutions to large MRF problems. All this is achieved while keeping the flexibility of SGM, without introducing new parameters and with a computing overhead of at most 20%.

**Previous work.** The strategies for finding a minimum of the global energy function differ. Many methods simplify the 2D graph by reducing it to simpler subproblems. But these sim-

plifications are critical as they determine the flow of information. Dynamic programming approaches [17] perform the optimization in one dimension for each scan line individually, which commonly leads to streaking effects. This is avoided by tree-based dynamic programming approaches. Some methods derive, from the 2D graph, a single tree that spans the entire image [27]. Others construct trees that vary their grid structure with the position of the pixel [3]. The Fast-PD algorithm [16] exploits information from the original MRF problem and its dual resulting in a remarkable speed-up with respect to alpha-expansion. In [5] a very efficient block coordinate descent is used to approximately solve MRFs with truncated smoothness terms. In SGM [12] the optimization is restricted to a star-shaped graph centered at the current pixel, which sometimes results in streaking artifacts over poorly textured areas. Increasing the number of paths (*i.e.* 16 directions) may suppress them but doubles the computational burden [12].

The technique we propose aims at increasing the support of SGM's graph without extra overhead, by leveraging its connection with message passing algorithms [6]. An on-line demo of the proposed method is available on the project website[1].

## 2 The baseline Semi-Global Matching

The Semi-Global Matching (SGM) algorithm [12] for dense stereo matching is an efficient tool for approximate energy minimization for 2D MRF.

The stereo matching problem is formulated as finding the disparity map $D$ that minimizes the global energy defined on the graph $G = (I, \mathcal{E})$

$$E(D) = \sum_{\mathbf{p} \in I} C_{\mathbf{p}}(D_{\mathbf{p}}) + \sum_{(\mathbf{p},\mathbf{q}) \in \mathcal{E}} V(D_{\mathbf{p}}, D_{\mathbf{q}}), \tag{1}$$

where $C_{\mathbf{p}}(d)$ is a unary data term that represents the pixel-wise cost of matching $\mathbf{p}$ for disparity $d \in \mathcal{D}$ (where $\mathcal{D} = \{d_{min}, \cdots, d_{max}\}$ is the search space). The pairwise terms $V(\cdot,\cdot)$ enforce smoothness of the solution by penalizing changes of neighboring disparities on the edge set $\mathcal{E}$ (which is usually the 8-connected or 4-connected image graph). SGM considers a particular type of pairwise term of the form

$$V(d,d') = \begin{cases} 0 & \text{if } d = d' \\ P1 & \text{if } |d - d'| = 1 \\ P2 & \text{otherwise} \end{cases} . \tag{2}$$

It imposes a small penalty $P1$ for small jumps in disparity (up to one pixel), which are common on slanted surfaces. The second penalty term $P2$ (with $P2 > P1$) accounts for larger disparity jumps. The penalty $P2$ can be further adapted [3, 12, 22] depending on the image content to align the disparity discontinuities with the discontinuities in the image.

The SGM algorithm computes an approximate solution to the NP-hard problem (1). The strategy adopted by SGM consists in dividing the grid-shaped problem into multiple one-dimensional problems defined on scan lines, which are straight lines that run through the image in the 4 or 8 cardinal directions. For each cardinal direction $\mathbf{r}$, SGM computes a matrix of costs $L_{\mathbf{r}}$. The costs $L_{\mathbf{r}}(\mathbf{p}, d)$ are computed recursively from the edges of the image along a path in the direction $\mathbf{r}$:

$$L_{\mathbf{r}}(\mathbf{p}, d) = C_{\mathbf{p}}(d) + \min_{d' \in \mathcal{D}} (L_{\mathbf{r}}(\mathbf{p} - \mathbf{r}, d') + V(d, d')). \tag{3}$$

[1] http://dev.ipol.im/~facciolo/mgm

This recursion is in fact a dynamic programming algorithm that solves the problem restricted to the directed graph induced by the scan line $\{\mathbf{p} - i\mathbf{r}\}, i = 0, 1, 2\ldots$ Because of the special form of the smoothness potential (2), $L_{\mathbf{r}}(\mathbf{p}, \cdot)$ can be computed with just 7 instructions per disparity. These costs are computed in each direction $\mathbf{r}$ and are then added to obtain an aggregated cost volume

$$S(\mathbf{p}, d) = \sum_{\mathbf{r}} L_{\mathbf{r}}(\mathbf{p}, d). \tag{4}$$

The final disparity for each pixel is selected based on a Winner-Take-All (WTA) evaluation of $S(\mathbf{p}, \cdot)$.

## 2.1 SGM as Min-Sum Belief Propagation

Belief Propagation (BP) [18] can be seen as an energy minimization algorithm on a graph. The Min-sum BP is a *message-passing* algorithm, that computes each node's belief (energy min marginal) by sending messages along the edges of the graph. A message from node $\mathbf{q}$ to node $\mathbf{p}$ is defined recursively from the messages to node $\mathbf{q}$ as:

$$m_{\mathbf{q}\to\mathbf{p}}(d) = \min_{d' \in \mathcal{D}} (C_{\mathbf{q}}(d') + \sum_{(\mathbf{q},\mathbf{k}) \in \mathcal{E}, \mathbf{k} \neq \mathbf{p}} m_{\mathbf{k}\to\mathbf{q}}(d') + V(d, d')). \tag{5}$$

And the state belief of a node $\mathbf{p}$ is computed from the messages as

$$B(\mathbf{p}, d) = C_{\mathbf{p}}(d) + \sum_{(\mathbf{q},\mathbf{p}) \in \mathcal{E}} m_{\mathbf{q}\to\mathbf{p}}(d). \tag{6}$$

Upon convergence $\arg\min_d B(\mathbf{p}, d)$ yields the estimated solution.

On a tree, BP computes the exact minimum of the energy in two passes sending message from the leafs to the root and back. On a general graph however, BP is implemented as an iterative algorithm, updating messages according to some schedule. The sequential schedule [24, 25], for instance, propagates messages in raster order updating the nodes immediately. Messages are typically initialized to all-zero.

Equation (5) is reminiscent of SGM. Indeed, Drory *et al.* [6] established the connection between SGM and the min-sum BP algorithm. In their interpretation the SGM recursive update formula (3) is computing the state belief of the node $\mathbf{p}$ during the $\mathbf{r}$-oriented scan of the image. To make it explicit:

$$B_{\mathbf{r}}(\mathbf{p}, d) = L_{\mathbf{r}}(\mathbf{p}, d) = C_{\mathbf{p}}(d) + \overbrace{\min_{d' \in \mathcal{D}} (L_{\mathbf{r}}(\mathbf{p} - \mathbf{r}, d') + V(d, d'))}^{m_{(\mathbf{p}-\mathbf{r})\to(\mathbf{p})}(d)}. \tag{7}$$

They also show that the aggregate of state beliefs for all 8-directions given by (4), corresponds (modulo a correction) to the min-marginals for the star-shaped graph centered at $\mathbf{p}$ (shown in figure 1). The authors show that compared to BP, in SGM the aggregate $S$ overcounts the data term by a factor $N_{dir} - 1$ ($N_{dir} = 8$ in the case of 8-directions). They propose to subtract the excess data-terms directly from the aggregated costs as

$$S_{oc}(\mathbf{p}, d) = \sum_{\mathbf{r}} L_{\mathbf{r}}(\mathbf{p}, d) - (N_{dir} - 1) C_{\mathbf{p}}(d). \tag{8}$$

Then, selecting the WTA over $S_{oc}(\mathbf{p}, \cdot)$ equals to the exact minimizer of the energy restricted to the star-shaped graph.

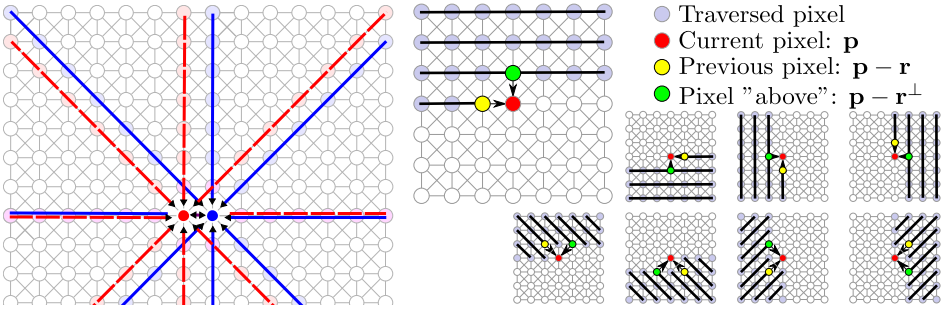



Figure 2: (a) Star-shaped graphs associated in SGM to two adjacent pixels. (b) Depiction of the 8 image traversals and the corresponding recursion directions $\mathbf{r}$ and $\mathbf{r}^{\perp}$ in MGM.

## 2.2 SGM and streaking artifacts

In the light of this result we can understand the occurrence of the streaking artifacts in SGM. Although the optimization is exact over the star-shaped graph, the graphs for two adjacent pixels (as shown in figure 2(a)) are loosely related as they share only the nodes on a single line (plus 4 intersection points). In this case, when the data term on the horizontal line is weak, *i.e.* all disparity hypothesis are equally plausible, the messages from the vertical directions, which are completely unrelated, can and will produce different results for each pixel. This also means that the smoothness constraint is poorly enforced by SGM because the messages are restricted to the 8 paths of the graph.

# 3 More Global Matching

Our main contribution can be summarized as a change in the recursive update formula (3). In the spirit of the belief update formula (6) we propose to update $L_{\mathbf{r}}$ using information from more than one direction. Concretely our strategy injects information from the 2D problem in the processing of SGM's 1D paths (see figure 1). This is efficiently done by incorporating messages from the nodes visited in the previous scanline (*i.e.* the pixel above).

Let us consider the left-to-right direction. The image is traversed in raster order (left-to-right, top-to-bottom) and SGM updates each node $\mathbf{p}$ using only the beliefs from the node on its left $L_{\mathbf{r}}(\mathbf{p}-\mathbf{r},\cdot)$. Instead we propose to access as well the beliefs from the node directly above $\mathbf{p}$ (indicated by the direction $\mathbf{r}^{\perp}$) . Thus our proposed recursion is:

$$L_{\mathbf{r}}(\mathbf{p},d) = C_{\mathbf{p}}(d) + \sum_{\mathbf{x}\in\{\mathbf{r},\mathbf{r}^{\perp}\}} \frac{1}{2} \min_{d'\in\mathcal{D}} (L_{\mathbf{r}}(\mathbf{p}-\mathbf{x},d') + V(d,d')). \tag{9}$$

As a result of this multiple recursion, the belief at a given pixel is influenced by its entire upper-left quadrant (as illustrated in figure 1). In comparison SGM recursion only sees information from the line of pixels to its left.

For each propagation direction $\mathbf{r}$ we compute $L_{\mathbf{r}}$ using an adequate traversal order (depicted in figure 2(b)). The resulting beliefs are then combined using the over-counting corrected formula (8), and the disparity is estimated by WTA. Compared to SGM, MGM only requires a few extra operations per pixel.

Each pass of MGM can be seen as the first iteration of a sequential belief propagation algorithm [24, 25] with a particular scheduling order dictated by the propagation direction (see figure 2(b)). Since the messages from the non visited nodes (below $\mathbf{p}$) are initialized to 0, the update rule (9) corresponds exactly to (5). However, extension of this correspondence beyond the first iteration is non-trivial, as messages from different passes may be needed. The exploration of this relationship is left for future work.

Since MGM introduces dependency among the scan lines, these cannot be processed in parallel in the same way as it is done with SGM. Parallelization in MGM is achieved diagonal-by-diagonal. That is, during the image traversal in raster order the pixels of a diagonal going from top-right to bottom-left can be processed in parallel, because they only depend on their top and left neighbors (which were computed by the previous diagonal). Applying this strategy we observed a nearly linear speed-up on a 16-core test machine (see the supplementary material for details).

# 4 Experiments

We first evaluate the effectiveness of our method as a fast approximate minimization tool for pairwise MRF. Then the improvement of the proposed method is quantified in terms of performance on a stereo benchmark.

## 4.1 Evaluation for MRF minimization

We evaluated our method on pairwise MRFs from the Middlebury energy minimization benchmark [24]. For the stereo-matching problem the MRF graphs consist of regular 4-connected grids. We compared with well known optimization algorithms provided with the benchmark.

**Experimental setup.** The disparity map from images $I_1$ and $I_2$ was estimated by minimizing

$$E(D) = \sum_{\mathbf{p}\in I} C_{\mathbf{p}}(D_{\mathbf{p}}) + \lambda \sum_{(\mathbf{p},\mathbf{q})\in\mathcal{E}} V(D_{\mathbf{p}}, D_{\mathbf{q}}), \tag{10}$$

where $\mathcal{E}$ is the edge set of the 4-connected image graph and the disparities $D_{\mathbf{p}}$ are discretized at one pixel precision. We used the absolute differences of intensities for the data term $C_{\mathbf{p}}(d) = |I_1(\mathbf{p}) - I_2(\mathbf{p}+d)|$ (summed over all channels for color images). This matching cost, although not robust to radiometric changes, provides a common ground for comparison with the methods in [24]. The smoothness term $V(\cdot,\cdot)$ is the $l_1$-norm truncated at 2, which amounts to taking $P1 = \lambda$ and $P2 = 2\lambda$ in (2). For simplicity the fine tuning of the matching costs and variable weights were not considered here (essentially disabling the intensity cues and the Birchfield & Tomasi costs [2] for all the methods). However different matching costs and intensity cues can be easily incorporated in the proposed method as in SGM [8, 12]. For instance, in section 4.2 we use census [28] as matching cost, which is robust to radiometric changes. In this section the parameter $\lambda$ is chosen for each image but is unchanged across algorithms.

Four test images were considered. Three from the benchmark [24]: *Tsukuba* (16 depth labels, $\lambda = 20$), *Venus* (20 depth labels, $\lambda = 20$), *Teddy* (60 depth labels, $\lambda = 10$), and one stereo pair from the multi-view dataset by Strecha *et al.* [23]: *Fountain* (143 depth labels, $\lambda = 8$).

| | Tsukuba(16 labels) | | | Teddy(60 labels) | | | Venus(20 labels) | | | Fountain(143 labels) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Method** | **Egap%** | $\mathcal{B}(\%)$ | **t(s)** | **Egap%** | $\mathcal{B}(\%)$ | **t(s)** | **Egap%** | $\mathcal{B}(\%)$ | **t(s)** | **Egap%** | $\mathcal{B}(\%)$ | **t(s)** |
| TRW-S | **ref.** | **4.5** | 17 | **ref.** | 20.5 | 483 | **ref.** | **4.7** | 51 | ref. | **16.2** | 2170 |
| Expansion[4] | 0.09 | 4.6 | 5.9 | 0.13 | 21.2 | 58 | 0.07 | **4.7** | 12 | **-0.05** | 16.3 | 196 |
| BP-S | 1.78 | 8.2 | 4.7 | 0.68 | **20.1** | 382 | 0.68 | 6.3 | 18 | 1.32 | 16.9 | 1697 |
| SGM4 | 48.3 | 6.6 | 1.0 | 21.7 | 24.2 | 3.7 | 31.4 | 7.4 | 1.8 | 29.0 | 14.5 | 14 |
| ocSGM4 | 41.9 | **6.3** | 1.0 | 18.2 | 21.9 | 3.8 | 25.6 | 6.9 | 1.8 | 26.5 | **13.6** | 14 |
| MGM4 | **7.5** | 6.7 | 1.2 | **5.5** | **21.4** | 4.5 | **4.2** | **5.8** | 2.0 | **10.7** | 15.8 | 17 |

Table 1: Results. **Egap%** is the energy gap ($\frac{E-E_{ref}}{E_{ref}} \times 100$) of the solution with respect to the reference solution of TRW-S (Tree-Reweighted Message Passing [14]), **t(s)** is the time in seconds to compute the solution, and $\mathcal{B}(\%)$ is the percentage of pixels that differ more than 1 pixels from the ground truth. Two other reference methods are also included: alpha-expansion algorithm [4], and BP-S a sequential Belief-Propagation algorithm from [24]. The running times correspond to a single core of an Intel Core 2 Duo CPU @1.8GHz.

**Evaluation.** We compared five minimization algorithms: Three reference MRF optimization techniques: *TRW-S* (Tree-Reweighted Message Passing) [14], the *alpha-expansion* graph-cut algorithm [4], and BP-S a sequential Belief-Propagation algorithm from [24]. Our own baseline implementation of SGM4 [12], one variant incorporating the over-counting correction ocSGM4 [6], and the proposed algorithm MGM4. Since the energy (10) is defined on a 4-connected graph, these variants only explore 4 cardinal directions.

We assessed the performance by computing the energy gap, *i.e.*, the relative gap between the energy of the current solution and the energy of a reference (top performing) strategy, the bad pixel ratio, *i.e.*, the percentage of pixels that differ more than one level from the ground-truth, and the time in seconds for computing the solution.

**Results and discussion.** The results are reported in table 1 and figure 3. We note in table 3 that, in terms of energy, SGM4 and the over-counting corrected ocSGM4 [6] produce rather distant solutions from the reference optimum. Looking at figure 3 we can see that this error comes mainly from the regularity term (see supplementary material), which confirms that in SGM the smoothness term of (10) is only weakly enforced. Our algorithm (MGM4), on the other hand, consistently produces better minima than SGM4 and ocSGM4. It yields approximations that are within 10% of the optimum produced by TRW-S (the reference method), an improvement by a factor five with respect to SGM4 or ocSGM4. Another experiment (see supplementary material) performed on a large set of images (all the stereo pairs from the Middlebury page: 2005, 2006, and 2014) confirms that on average MGM yields a systematic improvement of 40% in the energy minima with respect to SGM.

In terms of efficiency MGM can be 20% slower than SGM, however the results are well worth the time: they are more regular, present less artifacts, and better approach the minima of the underlying energy. For problems with larger label sets, such as *Fountain*, MGM (with its current non-optimized implementation) already provides an approximate solution in a tenth of the time needed by a classic technique. We can safely say that MGM can be used as an accurate and rather inexpensive approximation for solving problems with truncated smoothness terms or the Potts interaction penalty [22].

In the horizontal plane of the *Fountain* image (figure 3) we see that SGM4 produced smoother results than TRW-S and MGM4. The piecewise constant solutions are expected in this experiment because the smoothness term (2) is a truncated $l1$-norm. With it, it is




Figure 3: Qualitative result comparison for the images *Tsukuba* and *Teddy* from [24], plus a low resolution *Fountain* from [23]. Note that in the result of MGM4 there are no streaking artifacts as in the baseline SGM4. The last column (MGM4 $P1 = \lambda/4$) and the detail magnified in the last figure are commented in the text.

more convenient to jump three or more disparity levels once in a while (with penalty $2\lambda$) rather than changing disparity at each row (each with penalty $\lambda$). This serves as evidence that SGM4 does not actually enforce the smoothness defined in (10). Indeed, SGM4 only enforces the regularity by lines with a weak bond between parallel lines. On the other hand, this issue is easily solved with MGM, since it actually behaves more accordingly to the energy (10). For instance, lowering the value of the parameter $P1$ (for instance $P1 = \lambda/4$, and keeping $P2 = 2\lambda$) prevents this effect. The results are shown in the last column of figure 3.

## 4.2 Evaluation on Stereo Pairs with Ground Truth

**Experimental setup.** We evaluated our method using 37 full resolution stereo paris from the Middlebury datasets (2005, 2006, and 2014) [13, 21]. The images of the 2005 and 2006 datasets have a resolution of 1.4MP and the disparity range is about 150 pixels. The 2014 dataset contains 5MP images and the disparity range goes up to 500 pixels.

**Evaluation.** We were interested in comparing the performance of the proposed algorithm (MGM) with a baseline implementation of SGM and ocSGM (the over-counting corrected SGM). We evaluated the bad pixel ratio of each method. A bad pixel is a point whose disparity differs from the ground truth by more than one pixel. The three algorithms use the same settings: 8 propagation directions and as matching cost $C_{\mathbf{p}}$ the Hamming distance of census transform [28, 29] on a $5 \times 5$ neighborhood (normalized by the number of channels). The parameters $P1$ and $P2$ were set for all images to $P1 = 8$ and $P2 = 32$ (as in

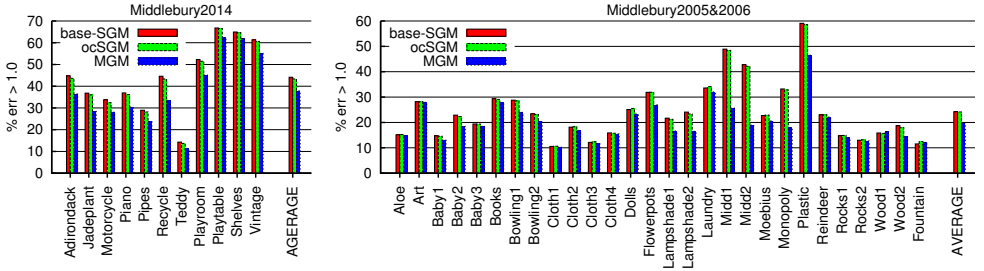



Figure 4: Bad pixel ratio (% of pixels with error > 1) on the two sets of test images. Our method yields the lowest average errors on the two datasets.

OpenCV's SGBM). No intensity cues (adapting $P2$) were used, they could easily be incorporated though [9, 12]. To prevent influence of the post-processing steps on the evaluation, we performed this evaluation on the unmodified outputs of the winner-take-all stage of all methods. In the supplementary material are included results after post-processing the outputs of SGM and MGM with a $3 \times 3$ median filter [13]. The conclusion is that, even after filtering the results of SGM present more errors than MGM, specially in poorly textured areas.

**Discussion.** The accuracy of the three methods is shown in figure 4. It can be seen that our MGM method is the most accurate on both datasets. This improvement is attained with negligible computation overhead. Our current implementation of SGM runs in 130s on an 8-core Xeon@2.60GHz computer for a $2964 \times 1988$ image with 226 disparity levels. On the same image the equivalent implementation of MGM runs in 137s.

A qualitative analysis of the disparity maps (figure 5) shows that MGM produces results that are denser and cleaner, with less streaking artifacts (more results in the supplementary material). To facilitate the qualitative evaluation these results were filtered with the left-to-right consistency check [8] with threshold set to 1. Moreover, for all these experiments MGM yielded (with respect to SGM) a systematic reduction of about 40% in the energy minima for the 8-connected energy (see supplementary material).

# 5 Conclusion

We proposed MGM a new method for stereo matching. Our method is a variant of SGM where messages are propagated on a quadrant of the whole graph instead of a line subgraph. This elaborates on a recent interpretation of SGM in terms of belief propagation. With a very small overhead MGM improves up to a factor of five the energy gap of SGM with respect to the best global algorithms. We validated experimentally that MGM produces better results than the baseline SGM, denser and without streaking artifacts. In summary, MGM produces better results than the baseline SGM with practically no computation overhead.

**Acknowledgements.** We thank Jean-Michel Morel for his support and fruitful discussions. This work partly founded by Centre National d'Etudes Spatiales (CNES, MISS Project), European Research Council (advanced grant Twelve Labours), Office of Naval research (ONR grant N00014-14-1-0023), DGA Stéréo project, ANR-DGA (project ANR-12-ASTR-0035), and Institut Universitaire de France.

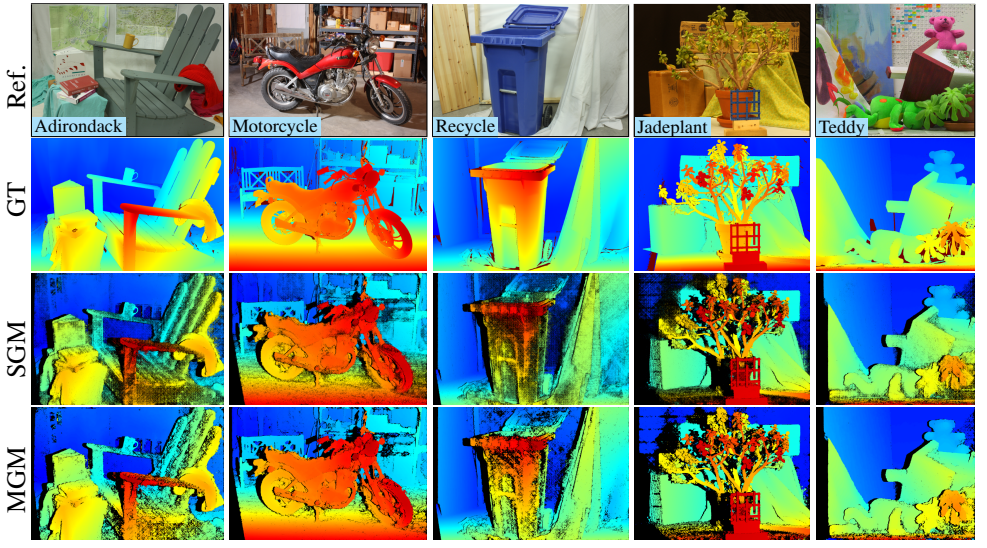



Figure 5: Results on some of the test images with ground truth.

## References

[1] F. Besse, C. Rother, A. Fitzgibbon, and J. Kautz. PMBP: PatchMatch Belief Propagation for Correspondence Field Estimation. In *Proc. BMVC*, pages 132.1–132.11, 2012.

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

[3] M. Bleyer and M. Gelautz. Simple but effective tree structures for dynamic programming-based stereo matching. In *Proc. VISAPP*, pages 415–422, 2008.

[4] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE Trans. Pattern Anal. Machine Intell.*, 23(11):1222–1239, 2001.

[5] Q. Chen and V. Koltun. Fast mrf optimization with application to depth reconstruction. In *Proc. CVPR*, pages 3914–3921, 2014.

[6] A. Drory, C. Haubold, S. Avidan, and F. A. Hamprecht. Semi-global matching: a principled derivation in terms of message passing. In *Pattern Recognition*, pages 43–53. Springer, 2014.

[7] P. Felzenszwalb and D. Huttenlocher. Efficient Belief Propagation for Early Vision. *Int. J .Comput. Vis.*, 70(1):41–54, October 2006.

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

[9] A. Geiger, M. Roser, and R. Urtasun. Efficient large-scale stereo matching. In *Lecture Notes in Comput. Sci.*, volume 6492 LNCS, pages 25–38, 2011.

[10] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521623049, 2000.

[11] S. Hermann and R. Klette. Iterative semi-global matching for robust driver assistance systems. In *Lecture Notes in Comput. Sci.*, volume 7726 LNCS, pages 465–478, 2013.

[12] H. Hirschmüller. Stereo processing by semiglobal matching and mutual information. *IEEE Trans. Pattern Anal. Machine Intell.*, 30(2):328–41, February 2008.

[13] H. Hirschmüller and D. Scharstein. Evaluation of stereo matching costs on images with radiometric differences. *IEEE Trans. Pattern Anal. Machine Intell.*, 31(9):1582–1599, September 2009.

[14] V. Kolmogorov. Convergent tree-reweighted message passing for energy minimization. *IEEE Trans. Pattern Anal. Machine Intell.*, 28(10):1568–1583, 2006.

[15] V. Kolmogorov and R. Zabih. Computing visual correspondence with occlusions using graph cuts. In *Proc. ICCV*, volume 2, pages 508–515. IEEE Comput. Soc, 2001.

[16] N. Komodakis, G. Tziritas, and N. Paragios. Fast, approximately optimal solutions for single and dynamic MRFs. In *Proc. CVPR*, 2007.

[17] Y. Ohta and T. Kanade. Stereo by Two-level Dynamic Programming. In *International Joint Conference on Artificial Intelligence*, pages 1120–1126, 1985.

[18] J. Pearl. *Probabilistic Reasoning in Intelligent Systems*, volume 88. 1988. ISBN 1558604790.

[19] C. Rhemann, A. Hosni, M. Bleyer, C. Rother, and M. Gelautz. Fast cost-volume filtering for visual correspondence and beyond. In *Proc. CVPR*, pages 3017–3024, 2011.

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

[21] D. Scharstein, H. Hirschmüller, Y. Kitajima, K. Krathwohl, N. Nešić, X. Wang, and P. Westling. High-resolution stereo datasets with subpixel-accurate ground truth. In *Pattern Recognition*, pages 31–42. Springer, 2014.

[22] S. Sinha, D. Scharstein, and R. Szeliski. Efficient High-Resolution Stereo Matching using Local Plane Sweeps. In *Proc. CVPR*, pages 1582–1589, 2013.

[23] C. Strecha, W. Von Hansen, L. Van Gool, P. Fua, and U. Thoennessen. On benchmarking camera calibration and multi-view stereo for high resolution imagery. In *Proc. CVPR*, pages 1–8, 2008.

[24] 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 with smoothness-based priors. *IEEE Trans. Pattern Anal. Machine Intell.*, 30(6):1068–1080, 2008.

[25] M.F. Tappen and W.T. Freeman. Comparison of graph cuts with belief propagation for stereo, using identical MRF parameters. *Proc. ICCV*, 2003.

[26] J. Čech and R. Šára. Efficient sampling of disparity space for fast and accurate matching. In *Proc. CVPR*, pages 1–8, 2007.

[27] O. Veksler. Stereo correspondence by dynamic programming on a tree. In *Proc. CVPR*, volume 2, pages 384–390, 2005.

[28] R. Zabih and J. Woodfill. Non-parametric local transforms for computing visual correspondence. *Proc. ECCV*, 2:151–158, 1994.

[29] K. Zhu, P. D'Angelo, and M. Butenuth. Evaluation of stereo matching costs on close range, aerial and satellite images. In *Proc. ICPRAM*, pages 379–385. SciTePress, 2012.