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

Semi-global matching [3] (SGM) is a stereovision algorithm that approximately minimizes a global energy composed of pixel-wise matching cost and pair-wise smoothness terms. The accuracy and speed of SGM are the main reasons for its widespread adoption, even for applications beyond stereovision. In SGM the two-dimensional smoothness constraint is approximated as the average of one-dimensional line optimization problems, which amounts to solving the problem on a star-shaped graph (usually with 8 cardinal directions) centered at each pixel. However, since two adjacent scan lines share little information, this approximation also produces characteristic streaks in the final disparity image (see fig.1).

Based on a recently proposed interpretation of SGM as a min-sum Belief Propagation algorithm [1], we propose in this paper a new algorithm that improves the energy gap of SGM with respect to more comprehensive optimization algorithms. 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 yields higher quality results by removing the streaking artifacts of SGM.

SGM formulates stereo matching 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 the unary terms $C_{\mathbf{p}}(d)$ represent the pixel-wise cost of matching $\mathbf{p}$ for disparity $d \in \mathcal{D}$ ($\mathcal{D} = \{d_{min}, \cdots, d_{max}\}$). The pairwise terms $V(\cdot,\cdot)$ enforce smoothness of the solution by penalizing changes of neighboring disparities on the edge set $\mathcal{E}$ (usually the 8-connected image graph). SGM considers truncated pairwise terms of the form (with $P2 > P1$)

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

In SGM the 2D problem (1) is splitted into 1D sub-problems defined on scan lines that run through the image in the 8 cardinal directions. For each direction $\mathbf{r}$ SGM recursively computes the costs $L_{\mathbf{r}}$ from the edges of the image along the path in the direction $\mathbf{r}$:

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

The form of the smoothness potential (2) permits to compute $L_{\mathbf{r}}(\mathbf{p}, \cdot)$ with just 7 instructions per disparity [2]. The costs $L_{\mathbf{r}}$ computed for all directions $\mathbf{r}$ are then added to obtain the aggregated cost volume from which the final disparity is selected with a Winner-Take-All (WTA) strategy.

The SGM algorithm amounts to the min-sum Belief Propagation algorithm on a star-shaped graph centered at each pixel [1]. That is, the recursive formula (3) is actually computing the state belief of the node $\mathbf{p}$ for each $\mathbf{r}$-oriented path. And the aggregate of state beliefs for the 8 directions ($N_{dir} = 8$)

$$S_{oc}(\mathbf{p}, d) = \sum_{\mathbf{r}} L_{\mathbf{r}}(\mathbf{p}, d) - (N_{dir} - 1) C_{\mathbf{p}}(d), \tag{4}$$

corresponds to the min-marginals for the star-shaped graph centered at $\mathbf{p}$.

Min-sum Belief Propagation (BP) [4] can be used as an approximate energy minimization algorithm on a graph. On a generic graph, BP computes each node's belief by sending messages along the edges of the graph. A message from node $\mathbf{q}$ to node $\mathbf{p}$ is defined recursively as

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

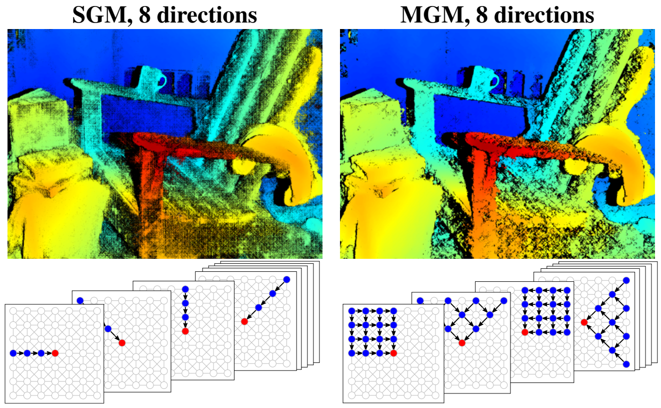

Figure 1: Disparity results of our baseline implementation of SGM and the proposed method MGM for the *Adirondack* pair (Middlebury 2014). The diagrams depict for each method the information used by the recursive update of the red pixel in each of the 8 scans of the algorithms.

The state belief of a node is then computed from the messages as

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

Messages are iteratively updated according to some schedule [5] and upon convergence $\arg\min_d B(\mathbf{p}, d)$ yields the estimated solution. For the star-shaped graph associated with SGM [1] equations (3) and (5) are related by $L_{\mathbf{r}}(\mathbf{p}, d) = C_{\mathbf{p}}(d) + m_{(\mathbf{p}-\mathbf{r})\to(\mathbf{p})}(d)$ and $S_{oc}(\mathbf{p}, d) = B(\mathbf{p}, d)$.

More Global Matching[1] (MGM). Our contribution consists in changing the recursive update formula (3). During the left-to-right pass of SGM the image is traversed in raster order (left-right, top-down), but node $\mathbf{p}$ is updated using only the cost of the node on its left $L_{\mathbf{r}}(\mathbf{p} - \mathbf{r}, \cdot)$. In MGM we propose (in the spirit of BP) to also consider the costs from the node directly above $\mathbf{p}$ (indicated by $\mathbf{r}^{\perp}$). Because of the raster traversal $L_{\mathbf{r}}(\mathbf{p} - \mathbf{r}^{\perp}, \cdot)$ is up-to-date, so the recursion becomes:

$$L_{\mathbf{r}}(\mathbf{p}, d) = C_{\mathbf{p}}(d) + \sum_{\mathbf{x}\in\{\mathbf{r},\mathbf{r}^{\perp}\}} \frac{1}{2} \min_{d'\in\mathcal{D}} \left(L_{\mathbf{r}}(\mathbf{p} - \mathbf{x}, d') + V(d, d')\right). \tag{7}$$

This recursion gathers information from an entire quadrant of the graph, instead of a segment as in SGM (illustrated in fig.1). The costs $L_{\mathbf{r}}$ computed in all traversals are combined using the eq. (4) and the disparity is estimated by WTA. Compared to SGM, MGM only requires a few extra operations per pixel and parallelization is also possible (along diagonals). Since messages from non visited nodes are initialized to 0, each pass of MGM can be seen as the first iteration of a sequential BP algorithm [5].

In summary, MGM produces qualitatively and quantitatively denser results than the baseline SGM with little computation overhead. MGM also yields lower energies than SGM for problem (1).

[1] A. Drory, C. Haubold, S. Avidan, and F. A. Hamprecht. Semi-global matching: a principled derivation in terms of message passing. *Pattern Recogn.*, 2014.

[2] P. Felzenszwalb and D. Huttenlocher. Efficient Belief Propagation for Early Vision. *IJCV*, 2006.

[3] H. Hirschmüller. Stereo processing by semiglobal matching and mutual information. *TPAMI*, 2008.

[4] J. Pearl. *Probabilistic Reasoning in Intelligent Systems*, 1988.

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

[1] **On-line demo of MGM is available at:** `http://dev.ipol.im/~facciolo/mgm`