# Supplementary note for "Discrete MRF Inference of Marginal Densities for Non-uniformly Discretized Variable Space"

Masaki Saito Takayuki Okatani Koichiro Deguchi
Tohoku University, Japan
{msaito, okatani}@vision.is.tohoku.ac.jp

## Detailed derivation of the MF and BP algorithms

We show here detailed derivation of the new MF and BP algorithms. In our main paper, we present simplified derivation which needs an assumption that the number $S_i$ of mixtures is the same for all sites, i.e., $S_i = S$ for any $i$. In what follows, we present complete derivation of the two algorithms which do not require this assumption; several equations that are omitted in the main paper are also given. Our derivation follows that of conventional MF and BP algorithms in [2].

### Derivation of the new MF algorithm

As mentioned in our main paper, MF and BP algorithms find $P$ that minimizes the following free energy:

$$F[P] = \langle E \rangle_P - S[P], \tag{30}$$

where the first term is the expectation defined as $\langle E \rangle_P = \int P(\boldsymbol{x})E(\boldsymbol{x})d\boldsymbol{x}$, and the second term is the entropy of $P$, i.e., $S[P] = -\int P(\boldsymbol{x}) \ln P(\boldsymbol{x})d\boldsymbol{x}$.

The derivation of MF algorithms start with assuming that the variable of each site $i$ is independent of that of any other site:

$$P(\boldsymbol{x}) \equiv \prod_i p_i(x_i). \tag{31}$$

Our approach is to represent $p_i(x_i)$ by a mixture of $S_i$ rectangular densities as

$$p_i(x_i) = \sum_{s=1}^{S_i} \alpha_i^s h_i^s(x_i), \tag{32}$$

where $\alpha_i^s$ is a mixing weight and $h_i^S(x_i)$ is a component rectangular density; the former is defined such that for any $i$

$$\sum_{s=1}^{S_i} \alpha_i^s = 1, \tag{33}$$

and the latter is defined such that the supports of any pair of the component densities do not overlap in the variable space, which is expressed as $\mathcal{X}_i^s \cap \mathcal{X}_i^t = \emptyset$ using the notations in our main paper.

Using Eqs.(31) and (32), $\langle E \rangle_P$ reduces as follows:

$$\begin{aligned}
\langle E \rangle_P &= \sum_i \int p_i(x_i) f_i(x_i) dx_i \\
&\quad + \sum_{(i,j)\in\mathcal{E}} \iint p_i(x_i) p_j(x_j) f_{ij}(x_i, x_j) dx_i dx_j \\
&= \sum_i \sum_s \alpha_i^s \int f_i(x_i) h_i^s(x_i) dx_i \\
&\quad + \sum_{(i,j)\in\mathcal{E}} \sum_{s,t} \alpha_i^s \alpha_j^t \iint f_{ij}(x_i, x_j) h_i^s(x_i) h_j^t(x_j) dx_i dx_j \\
&= \sum_i \sum_s \alpha_i^s f_i^s + \sum_{(i,j)\in\mathcal{E}} \sum_{s,t} \alpha_i^s \alpha_j^t f_{ij}^{st}.
\end{aligned} \tag{34}$$

Although the calculation of the entropy of a mixture density is in general intractable, owing to the introduced constraint $\mathcal{X}_i^s \cap \mathcal{X}_i^t = \emptyset$, we can reduce $S[P]$ as follows:

$$\begin{aligned}
S[P] &= -\sum_i \sum_s \alpha_i^s \int h_i^s(x_i) \ln \left( \sum_{s'} \alpha_i^{s'} h_i^{s'}(x_i) \right) dx_i \\
&= -\sum_i \sum_s \alpha_i^s \int h_i^s(x_i) \ln \left( \alpha_i^s h_i^s(x_i) \right) dx_i \\
&= -\sum_i \sum_s \alpha_i^s \ln \alpha_i^s - \sum_i \sum_s \alpha_i^s \int h_i^s(x_i) \ln h_i^s(x_i) dx_i \\
&= -\sum_i \sum_s \alpha_i^s \ln \alpha_i^s + \sum_i \sum_s \alpha_i^s B_i^s.
\end{aligned} \tag{35}$$

Using Eqs.(34) and (35), Eq.(30) is rewritten as

$$\begin{aligned}
F_{\mathrm{MF}}[\boldsymbol{\alpha}] &= \sum_i \sum_s \alpha_i^s (f_i^s - B_i^s) \\
&\quad + \sum_{(i,j)\in\mathcal{E}} \sum_{s,t} \alpha_i^s \alpha_j^t f_{ij}^{st} + \sum_i \sum_s \alpha_i^s \ln \alpha_i^s.
\end{aligned}$$

We then wish to find $P$ minimizing this free energy. The minimization is performed with respect to $\alpha_i^s$'s under the constraint of Eq.(33). Thus, introducing Lagrange multipliers $\gamma_i$, we consider minimizing

$$J_{\mathrm{MF}} = F_{\mathrm{MF}}[\boldsymbol{\alpha}] + \sum_i \gamma_i \left(1 - \sum_s \alpha_i^s\right). \tag{36}$$

The calculation of $\partial J_{\mathrm{MF}}/\partial \alpha_i^s = 0$ followed by elimination of $\gamma_i$'s yields the following fixed point equation:

$$\alpha_i^s \propto \exp\left[-\left((f_i^s - B_i^s) + \sum_{j\in\mathcal{N}_i}\sum_t f_{ij}^{st}\alpha_j^t\right)\right]. \tag{37}$$

Our MF algorithm iteratively updates $\alpha_i^s$'s according to this equation.

## Derivation of the new BP algorithm

The derivation of BP algorithms starts with assuming

$$P(\boldsymbol{x}) = \frac{\prod_{ij} p_{ij}(x_i, x_j)}{\prod_i p_i(x_i)^{z_i - 1}}. \tag{38}$$

Using this, $\langle E\rangle_P$ reduces to

$$\begin{aligned}\langle E\rangle_P = &\sum_i \int p_i(x_i) f_i(x_i) dx_i \\ &+ \sum_{(i,j)\in\mathcal{E}} \iint p_{ij}(x_i, x_j) f_{ij}(x_i, x_j) dx_i dx_j.\end{aligned} \tag{39}$$

Similarly, $S[P]$ reduces to

$$\begin{aligned}S[P] = &\sum_i (z_i - 1) \int p_i(x_i) \ln p_i(x_i) dx_i \\ &- \sum_{(i,j)\in\mathcal{E}} \iint p_{ij}(x_i, x_j) \ln p_{ij}(x_i, x_j) dx_i dx_j.\end{aligned} \tag{40}$$

(The free energy consisting of Eqs.(39) and (40) is called the Bethe Free Energy.)

Our approach is to represent $p_i$ and $p_{ij}$ as

$$p_i(x_i) = \sum_{s=1}^{S_i} \alpha_i^s h_i^s(x_i), \tag{41a}$$

$$p_{ij}(x_i, x_j) = \sum_{s=1}^{S_i}\sum_{t=1}^{S_j} \alpha_{ij}^{st} h_i^s(x_i) h_j^t(x_j), \tag{41b}$$

where $\alpha_i^s$ and $\alpha_{ij}^{st}$ are the parameters to be determined such that they satisfy the following three constraints:

$$\sum_s \alpha_i^s = 1, \tag{42a}$$

$$\sum_{s,t} \alpha_{ij}^{st} = 1, \tag{42b}$$

$$\sum_s \alpha_{ij}^{st} = \alpha_j^t. \tag{42c}$$

By substituting Eqs.(41) into Eq.(39), we have

$$\begin{aligned}\langle E\rangle_P &= \sum_i \sum_s \alpha_i^s \int f_i(x_i) h_i^s(x_i) dx_i \\ &\quad + \sum_{(i,j)\in\mathcal{E}} \sum_{s,t} \alpha_{ij}^{st} \iint f_{ij}(x_i, x_j) h_i^s(x_i) h_j^t(x_j) dx_i dx_j \\ &= \sum_i \sum_s \alpha_i^s f_i^s + \sum_{(i,j)\in\mathcal{E}} \sum_{s,t} \alpha_{ij}^{st} f_{ij}^{st}.\end{aligned} \tag{43}$$

To derive the entropy $S[P]$, we calculate the entropies of $p_i$ and $p_{ij}$, respectively. Using Eq.(41a), the (negative) entropy of $p_i$ is written as

$$\int p_i(x_i) \ln p_i(x_i) dx_i = \sum_s \alpha_i^s \ln \alpha_i^s - \sum_s \alpha_i^s B_i^s, \tag{44}$$

where we use $B_i^s$ defined in Eq.(16) in our main paper. Similarly, Using Eq.(41b), that of $p_{ij}$ is written as

$$\begin{aligned}&p_{ij}(x_i, x_j) \ln p_{ij}(x_i, x_j) dx_i dx_j \\ &= \sum_{s,t} \alpha_{ij}^{st} \iint h_i^s(x_i) h_j^t(x_j) \ln\left(\sum_{s',t'} \alpha_{ij}^{s't'} h_i^{s'}(x_i) h_j^{t'}(x_j)\right) dx_i dx_j \\ &= \sum_{s,t} \alpha_{ij}^{st} \iint h_i^s(x_i) h_j^t(x_j) \ln\left(\alpha_{ij}^{st} h_i^s(x_i) h_j^t(x_j)\right) dx_i dx_j \\ &= \sum_{s,t} \alpha_{ij}^{st} \ln \alpha_{ij}^{st} - \sum_{s,t} \alpha_{ij}^{st} (B_i^s + B_j^t).\end{aligned} \tag{45}$$

Using these, the entire entropy $S[P]$ is given as

$$\begin{aligned}S[\boldsymbol{\alpha}] = &\sum_i (z_i - 1) \sum_s \alpha_i^s \ln \alpha_i^s - \sum_i (z_i - 1) \sum_s \alpha_i^s B_i^s \\ &- \sum_{(i,j)\in\mathcal{E}} \sum_{s,t} \alpha_{st}^{ij} \ln \alpha_{st}^{ij} + \sum_{(i,j)\in\mathcal{E}} \sum_{s,t} \alpha_{ij}^{st} (B_i^s + B_j^t).\end{aligned} \tag{46}$$

Then, using Eq.(43) and Eq.(46), $F[P]$ is rewritten as

$$\begin{aligned}F_{\mathrm{BP}}[\boldsymbol{\alpha}] = &\sum_i \sum_s \alpha_i^s (f_i^s + (z_i - 1) B_i^s) + \sum_{(i,j)\in\mathcal{E}} \sum_{s,t} \alpha_{ij}^{st} (f_{ij}^{st} - B_i^s - B_j^t) \\ &- \sum_i (z_i - 1) \sum_s \alpha_i^s \ln \alpha_i^s + \sum_{(i,j)\in\mathcal{E}} \sum_{s,t} \alpha_{ij}^{st} \ln \alpha_{ij}^{st}.\end{aligned} \tag{47}$$

To simplify the expression, we define

$$\begin{aligned}\hat{E}_{ij}^{st} \equiv &(f_{ij}^{st} - B_i^s - B_j^t) \\ &+ (f_i^s + (z_i - 1) B_i^s) + (f_j^t + (z_j - 1) B_j^t)\end{aligned} \tag{48a}$$

$$\hat{E}_i^s \equiv f_i^s + (z_i - 1) B_i^s. \tag{48b}$$

The substitution of them into Eq.(47) yields

$$\begin{aligned}F_{\mathrm{BP}}[\boldsymbol{\alpha}] = &\sum_{(i,j)\in\mathcal{E}} \sum_{s,t} \alpha_{ij}^{st} (\hat{E}_{ij}^{st} + \ln \alpha_{ij}^{st}) \\ &- \sum_i (z_i - 1) \sum_s \alpha_i^s (\hat{E}_i^s + \ln \alpha_i^s).\end{aligned} \tag{49}$$

We then minimize Eq.(49) with respect to $\alpha_i^s$'s and $\alpha_{ij}^{st}$'s under the three constraints of Eqs.(42). By introducing Lagrange multipliers $\lambda_{ij}^t$, $\gamma_i$, and $\gamma_{ij}$, we consider minimizing

$$\begin{aligned}J_{\mathrm{BP}} &\equiv F_{\mathrm{BP}}[\boldsymbol{\alpha}] \\ &+ \sum_{(i,j)\in\mathcal{E}} \sum_t \lambda_{ij}^t \left(\alpha_j^t - \sum_s \alpha_{ij}^{st}\right) + \sum_{(i,j)\in\mathcal{E}} \sum_s \lambda_{ji}^s \left(\alpha_i^s - \sum_t \alpha_{ij}^{st}\right) \\ &+ \sum_i \gamma_i \left(1 - \sum_s \alpha_i^s\right) + \sum_{(i,j)\in\mathcal{E}} \gamma_{ij} \left(1 - \sum_{s,t} \alpha_{ij}^{st}\right). \end{aligned} \tag{50}$$

The calculation of $\partial J_{\mathrm{BP}} / \partial \alpha_{ij}^{st} = 0$ yields

$$\ln \alpha_{ij}^{st} = -\hat{E}_{ij}^{st} + \lambda_{ij}^t + \lambda_{ji}^s + \gamma_{ij} - 1. \tag{51}$$

Similarly, the calculation of $\partial J_{\mathrm{BP}} / \partial \alpha_i^s = 0$ yields

$$(z_i - 1)(\ln \alpha_i^s + 1) = -(z_i - 1)\hat{E}_i^s + \sum_{j\in\mathcal{N}_i} \lambda_{ji}^s + \gamma_i. \tag{52}$$

Introducing *messages* $m_{kj}^t$, are redefined the Lagrange multiplier $\lambda_{ij}^t$ as

$$\lambda_{ij}^t = \ln \prod_{k\in\mathcal{N}_j\backslash i} m_{kj}^t. \tag{53}$$

Then, by substituting Eq.(53) into Eqs.(51) and (52), we have

$$\alpha_{ij}^{st} \propto \psi_{ij}^{st} \phi_i^s \phi_j^t \prod_{k\in\mathcal{N}_i\backslash j} m_{ki}^s \prod_{l\in\mathcal{N}_j\backslash i} m_{lj}^t, \tag{54a}$$

$$\alpha_i^s \propto \phi_i^s \prod_{k\in\mathcal{N}_i} m_{ki}^s. \tag{54b}$$

Substituting Eqs.(54a) and (54b) into Eq.(42c), we have the following updating rule of BP:

$$m_{ij}^t \leftarrow \sum_s \phi_i^s \psi_{ij}^{st} \prod_{k\in\mathcal{N}_i\backslash j} m_{ki}^s.$$

## Additional experimental results

We show here additional results of stereo matching obtained by the proposed method. Choosing three images from the Middlebury MRF library, *Cloth1*, *Rocks1*, and *Flowerpots*, we applied the proposed method to them in a similar manner to *Aloe* shown in the main paper. Fig.6 shows their input images (including *Aloe*) along with their ground truths; the results obtained by the MAP inference (the $\alpha$-expansion algorithm [1]) are also shown for comparison with our results.

Figs.7–12 show the results formatted in the same way as *Aloe* in our main paper. Similar to *Aloe*, it is seen that the mixture densities depict the marginal densities in a finer way with the increasing number of blocks; as compared with the results of the fixed discretization, those of the dynamic discretization draw much finer details even for smaller number of blocks. It is seen from the estimated disparity maps that those of the dynamic discretization tend to be smoother than the fixed discretization, which well agrees with the observation on the estimated marginal densities.

## References


[1] R. Szeliski, R. Zabih, D. Scharstein, O. Veksler, A. Agarwala, and C. Rother. A comparative study of energy minimization methods for Markov random fields. In *ECCV*, pages 16–29, 2006. 3

[2] J. S. Yedidia, W. T. Freeman, and Y. Weiss. Understanding Belief Propagation and Its Generalizations, 2003. 1

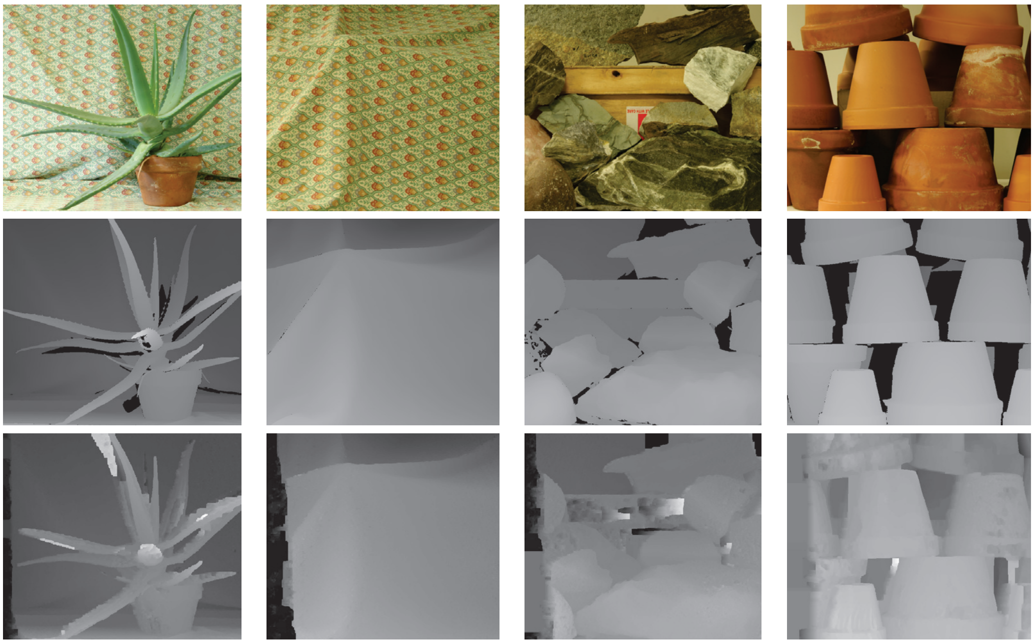

Figure 6. Four datasets used in our experiments: *Aloe*, *Cloth1*, *Rocks1*, and *Flowerpots*. Upper row: Input left images. Middle row: Ground truths. Lower row: Disparity maps estimated by the $\alpha$-expansion algorithm.

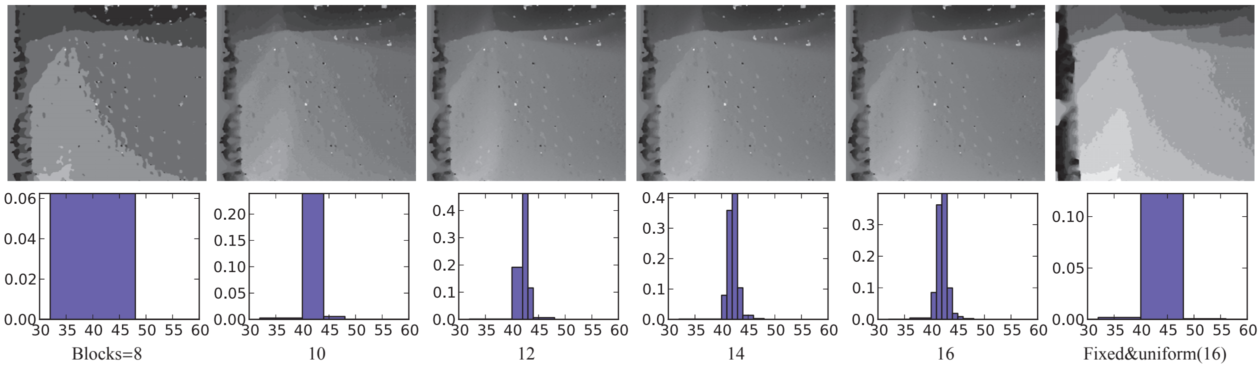


Figure 7. Results for *Cloth1* of the MF algorithm with the dynamic discretization. Upper row: Disparity maps. Lower row: The mixture of rectangular densities approximating the marginal density at the site of the image pixel (100, 100).

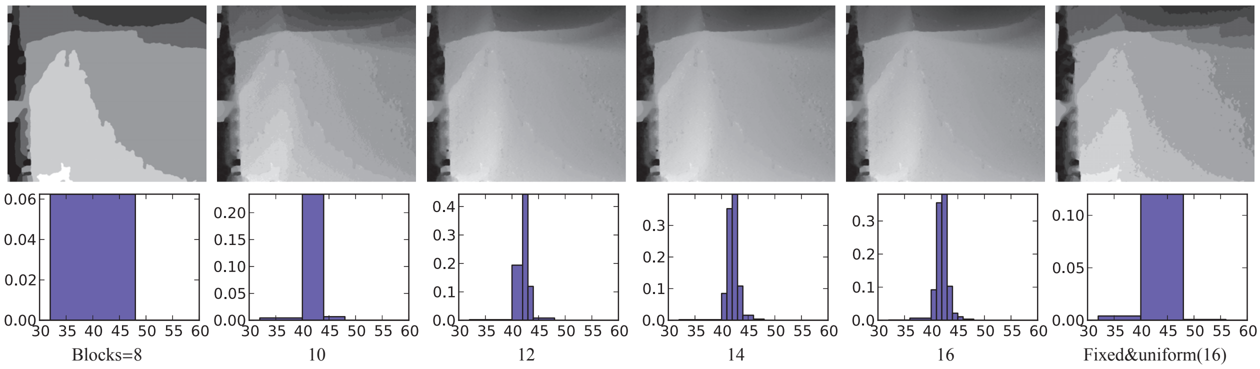


Figure 8. Results for *Cloth1* of the BP algorithm with the dynamic discretization. Upper row: Disparity maps. Lower row: The mixture of rectangular densities approximating the marginal density at the site of the image pixel (100, 100).

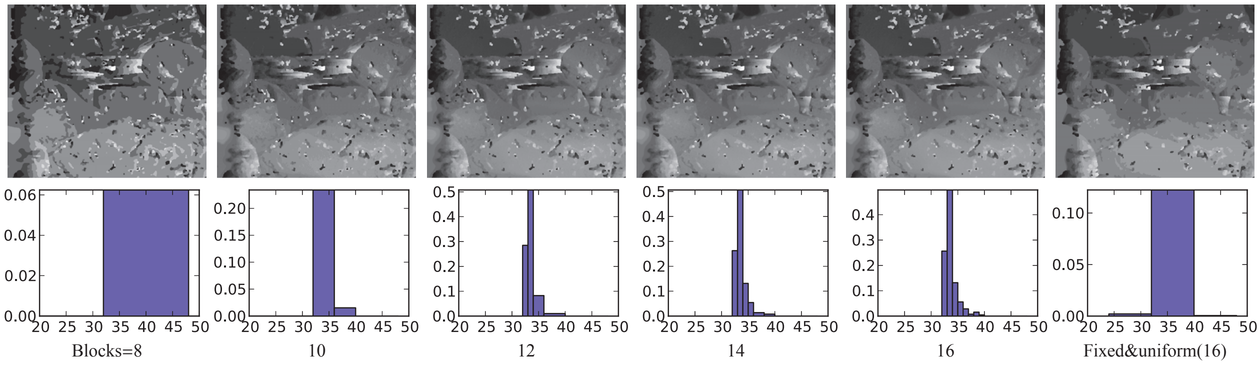

Figure 9. Results for *Rocks1* of the MF algorithm with the dynamic discretization. Upper row: Disparity maps. Lower row: The mixture of rectangular densities approximating the marginal density at the site of the image pixel (100, 100).

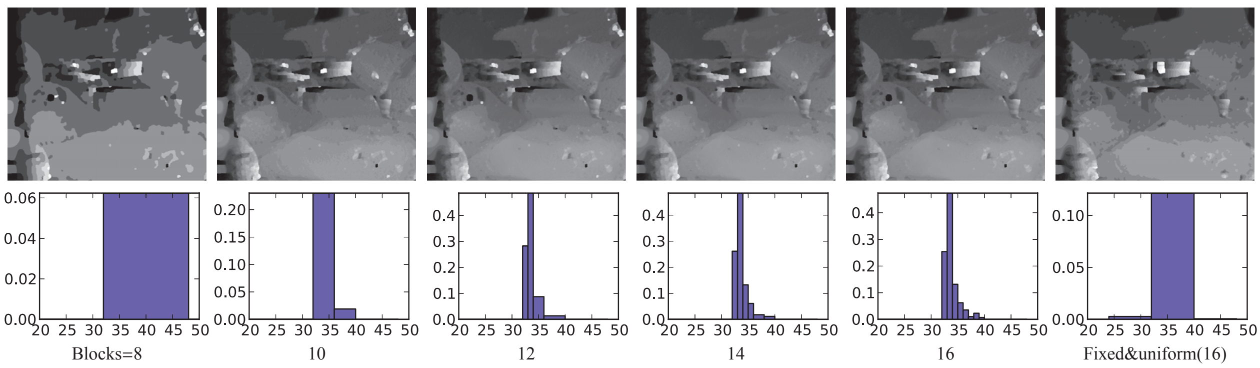

Figure 10. Results for *Rocks1* of the BP algorithm with the dynamic discretization. Upper row: Disparity maps. Lower row: The mixture of rectangular densities approximating the marginal density at the site of the image pixel (100, 100).

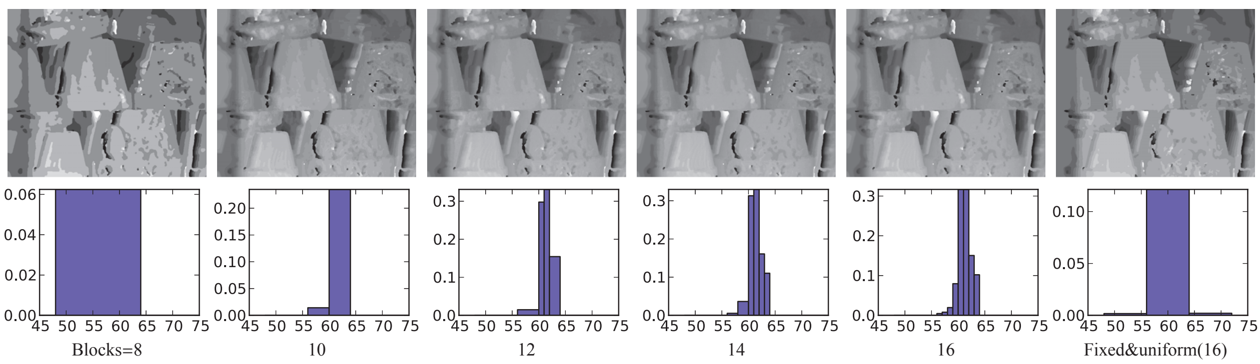

Figure 11. Results for *Flowerpots* of the MF algorithm with the dynamic discretization. Upper row: Disparity maps. Lower row: The mixture of rectangular densities approximating the marginal density at the site of the image pixel (100, 100).

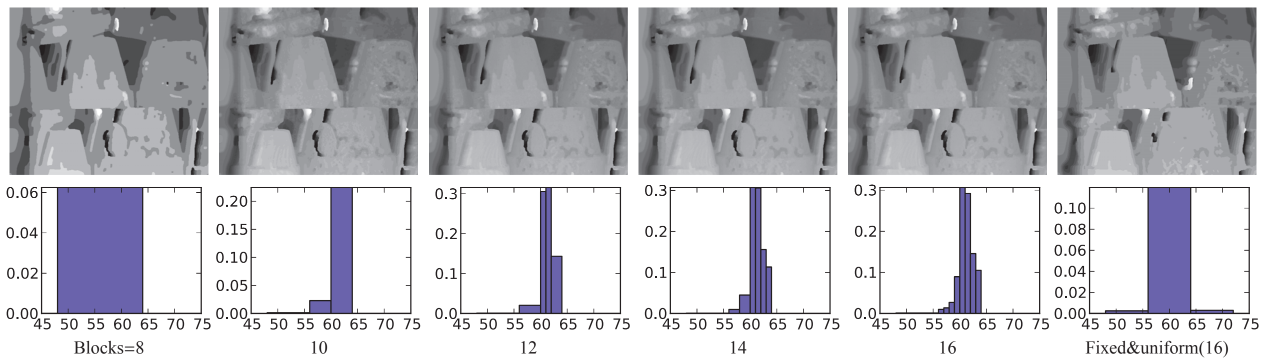


Figure 12. Results for *Flowerpots* of the BP algorithm with the dynamic discretization. Upper row: Disparity maps. Lower row: The mixture of rectangular densities approximating the marginal density at the site of the image pixel (100, 100).