Search Space Reduction for MRF Stereo

Liang Wang[1], Hailin Jin[2], and Ruigang Yang[1]

[1] Center for Visualization and Virtual Environments, University of Kentucky, USA
[2] Advanced Technology Labs, Adobe Systems Incorporated, San Jose, CA, USA

Abstract. We present an algorithm to reduce per-pixel search ranges for Markov Random Fields-based stereo algorithms. Our algorithm is based on the intuitions that reliably matched pixels need less regularization in the energy minimization and neighboring pixels should have similar disparity search ranges if their pixel values are similar. We propose a novel bi-labeling process to classify reliable and unreliable pixels that incorporate left-right consistency checks. We then propagate the reliable disparities into unreliable regions to form a complete disparity map and construct per-pixel search ranges based on the difference between the disparity map after propagation and the one computed from a winner-take-all method. Experimental results evaluated on the Middlebury stereo benchmark show our proposed algorithm is able to achieve 77% average reduction rate while preserving satisfactory accuracy.

1 Introduction

Stereo reconstruction has been one of the most investigated topics in Computer Vision for decades. Recently there have been great advances which are based on global energy minimizations. According to the Middlebury benchmark [1], most top performers are based on the Markov Random Field (MRF) formulation that solves stereo matching by minimizing certain cost functions. The two most popular algorithms for approximate inference on Markov Random Fields are Belief Propagation (BP) [2] and Graph Cuts (GC) [3,4]. Although being able to produce top-ranking results, all the algorithms based on the MRF formulation suffer significantly from heavy memory requirements and high computational complexities for large images and disparity search ranges. For instance, the memory required by BP to store all the messages scales on the order of $O(|I| \times |S| \times N)$, where I is the image, S is the disparity range and N is the size of the neighborhood system. For a 1-megapixel image with 200 disparities, a BP-based algorithm would need $3.2GB$ RAM just to only store all the single-precision floating-point messages using the standard 4-connected neighborhood system. GC is less memory consuming. However, as suggested in [5], most min-cut/max-flow algorithms behave nonlinearly (closer to quadratic) to the number of disparities. As a result, when the number of disparities goes up, which occurs naturally in high-resolution or wide-baseline stereo images, neither BP nor GC is practical to run on ordinary computers.

To address the above issue, we present in this paper a novel algorithm for reducing disparity ranges for MRF-based stereo. Regular MRF solver can be

D. Forsyth, P. Torr, and A. Zisserman (Eds.): ECCV 2008, Part I, LNCS 5302, pp. 576–588, 2008.

applied on this reduced label space, which leads to less memory usage and lower computational complexity. Our algorithm springs from the following two intuitions: First reliably matched pixels, for instance the *ground control points* (*GCPs*) [6], need less regularization in the global optimization step; Second, neighboring pixels should have similar disparity search ranges if their pixel values are similar. In more details, our algorithm starts from classifying each pixel as either stable or unstable based on its local cost distribution together with a left-right consistency check [7]. We formulate this bi-labeling process as a maximum a posterior MRF (MAP-MRF) problem which can be efficiently solved using either BP or GC. We then propagate the stable matches to the rest of the image following our second intuition to produce a complete disparity map. Inspired by [8], we formulate this propagation problem using a quadratic cost function which are solved using standard techniques. The complete disparity map from the propagation stage and the one computed in the local winner-take-all method guide the selection of the final per-pixel search ranges. We evaluated our algorithm on the Middlebury data sets with ground-truth and found that our algorithm is able to achieve 77% average reduction rate while preserving satisfactory accuracy.

1.1 Relation to Previous Work

Our work is most related to [9], which aims at reducing search space for GC based stereo algorithms. The authors proposed a reduction strategy based on window matching. The basic idea is to collect disparities from other pixels within a spacial window to construct the search range for the center pixel. They did not take reliable matching into consideration and the reduced search ranges may contain redundancies, especially for pixels near depth discontinuities. There was no explicit number on the reduced label space and an average 2.8 time speed-up was reported. Given the typical nonlinear complexity of GC, our search space reduction rate around 75% would have resulted in a speedup factor over 4.

This work also relates to a recent paper by Yu et al. [10], who studied the feasibility of applying compression techniques to the messages in the BP algorithm to improve the memory efficiency. Instead of treating the messages as generic data, we investigate the problem from a more domain-specific perspective by finding a plausible subset of search space derived from the image data.

By using reliably matched pixels as a starting point, our work also falls in the category of semi-dense stereo matching [11,12,13,14,15]. We emphasize that our reliable matching extraction algorithm is designed to address a different problem from these methods. Instead of finding the "most likely" disparity assignment for a given pixel, we decide whether its disparity given by the winner-take-all approach is reliable or not. Hence what we try to solve is a binary segmentation problem whose complexity does not increase with respect to disparity search ranges. In addition our algorithm has many distinct features and advantages. Firstly, our reliability measure is based on probabilistic models studied from a set of training stereo images with ground-truth disparities and thereby the most sensitive parameters are found automatically. Secondly, with a well-designed

energy function the found $GCPs$ are almost free of outliers. What is more, the global optimization formulation allows existing MRF stereo solvers to be employed directly to the stable matching extraction.

2 Search Space Reduction

The general framework for MRF-based stereo can be defined as follows: Let I be the reference image and S be a finite set of disparity candidates. A disparity function f assigns each pixel $p \in I$ a disparity value $f_p \in S$. The quality of the labeling function is given by an energy function of the form:

$$E(f) = \sum_{p \in I} C_p(f_p) + \sum_{(p,q) \in \xi} P(f_p, f_q). \tag{1}$$

where the first term, the data term, measures how well the labeling function agrees with the image data, while the second term, the smoothness term, encourages neighboring pixels to have similar disparity assignments based on the assumption that the scene is locally smooth. The optimal labeling function for the energy $E(f)$ can be approximately solved using either BP or GC. For more details, we refer the reader to [16].

This paper addresses the following problem within the framework: How to construct a non-empty subset S_p from S for each pixel such that the optimal labeling function based on $\{S_p\}$, where $\{S_p\} \doteq \cup S_p$, is as close as possible to the optimal labeling function based on S. We intuitively define the reduction rate as:

$$\Re = 1 - \frac{\sum_{p \in I} |S_p|}{|I| \cdot |S|}. \tag{2}$$

2.1 Matching Cost Computation

We first compute the per-pixel matching cost using the algorithm proposed by Birchfield and Tomasi [17]. We then aggregate the per-pixel matching cost over a spacial window. To avoid the "fattening" artifacts and be efficient, we adopt the two-pass adaptive algorithm proposed by [18] which is an approximation of the bilateral filtering of [19] which has shown to be remarkably effective for obtaining high quality disparity maps [18,20]. This approximation reduces the per pixel complexity from $O(\ell^2)$ to $O(\ell)$, where ℓ is the side length of the square window. Our implementation shows, although the disparity map from winner-take-all is slightly noisier compared to the result from full 2D aggregation, the processing time for Tsukuba image is only 1.2 seconds with approximation.

We emphasize that the adaptive aggregation is introduced to disambiguate the inaccurate pixel-wise matching cost for the search space reduction purpose. In Fig. 1 we show the winner-take-all disparity maps and our approximate two-pass aggregation in the first two columns. The error percentage in non-occluded areas are 69.4% and 16.6% respectively based on the Middlebury evaluation system [1]. We further demonstrate the disparity maps from GC using the corresponding

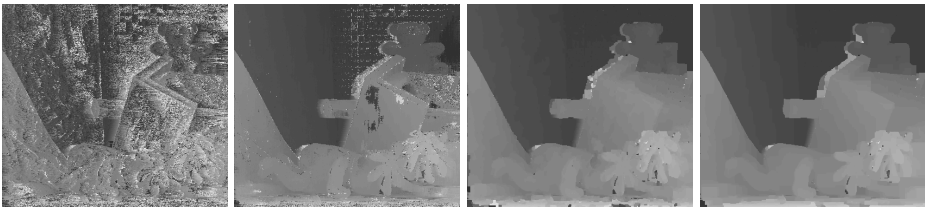

Fig. 1. From left to right: disparity map computed before ($e\% = 69.4$) and after ($e\% = 16.6$) adaptive cost aggregation using winner-take-all; the disparity map computed using Graph Cuts from the pixel-wise matching cost ($e\% = 9.25$); the disparity map computed using Graph Cuts after cost aggregation ($e\% = 6.81$)

matching cost as the data term in the third and fourth columns. While the results are visually similar, our aggregation step not only reduces the error rate from 9.25% to 6.81% but also brings GC's converging time from 27.8 seconds down to 24.2 seconds. It can be seen that MRF stereo can benefit from more reliable matching cost in terms of both quality and efficiency.

2.2 Stable Matching Extraction

Inspired by the fact that a majority of correct disparities can be estimated from their local costs after aggregation, we propose to reduce search space by locating those stably matched pixels and limit their disparity ranges substantially. There are existing algorithms designed to find unambiguous disparity assignments and derive semi-dense disparity maps. For instance, [11] uses dynamic programming to selectively assign disparities to pixels whose reliability exceed a given threshold; [12] utilizes an ordering constraint and uniqueness constraint together with some confidence measure to detect unambiguous component. Since these approaches need to explore the entire 3D cost volume, their computational complexities increase with respect to the search range.

To efficiently find the unambiguously matched pixels, we formulate a stable matching extraction as a binary segmentation problem. More specifically, given a disparity map D computed from winner-take-all, a function β assigns each pixel p with a label β_p: $\beta_p = 1$ if p's disparity estimate $D(p)$ is close enough to its true disparity and $\beta_p = 0$ otherwise. Our objective is to find the optimal labeling function that minimizes the following energy:

$$E(\beta) = \sum_{p \in I} U_p(\beta_p) + \sum_{(p,q) \in \xi} V(\beta_p, \beta_q). \tag{3}$$

Equation 3 is similar in form to equation 1 where the first term measures how well the configuration from β fits given matching data and the second term encodes a smoothness assumption on the binary labels.

Data Term. Binocular stereo matching is ill-posed inherently. Matching ambiguities can originate from many factors, among which insufficient signal-to-noise

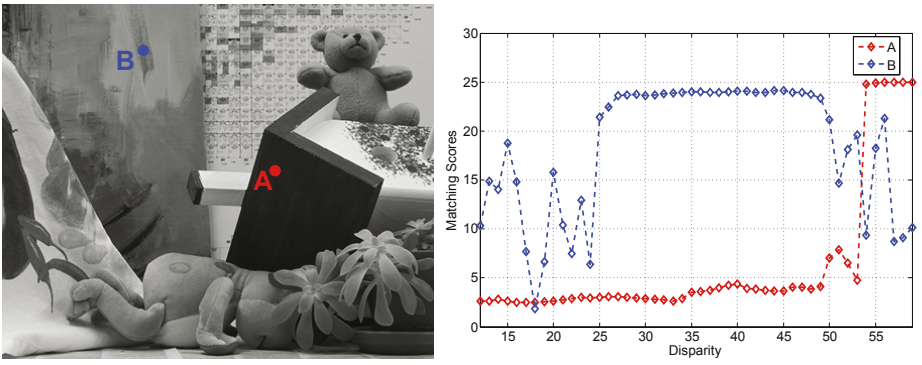

Fig. 2. Left: an image from the Teddy data set. Two points, A and B, are labeled in color. Point A belongs to a weakly texture region and point B belongs to an area with rich texture. Right: the cost profile of points A and B. Note that B's cost distribution has a clear global minimum while A's cost distribution is relatively flat.

ratio (SNR) in weakly textured regions and missing data in occlusion areas are the two most dominant ones [12]. To provide some insight into the behavior of matching costs, we select two scene points from the Teddy data set (shown in Fig. 2). The point A belongs to a textureless region while point B lies in an area with rich texture. On the right we plot their cost distributions respectively. We can see that B has a distinct global minimum at disparity 18. A's cost profile, however, is flat from disparity 0 to 35 and there is no distinctive global minimum. By referring to the ground-truth disparity map, B's true disparity is exactly 18 which is consistent with our estimate. A's disparity is incorrectly estimated at disparity 17 since $C_A(17) = 2.498$, which is slightly better than the cost 2.623 given by A's true disparity value 33.

We intuitively define the matching confidence γ_p for a pixel to be how distinctive the global minimum in the matching cost profile is:

$$\gamma_p = \begin{cases} 1 - \frac{c_p^{1st}}{c_p^{2nd}} & c_p^{2nd} > T_c \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

where c_p^{1st} and c_p^{2nd} are the best (lowest) and the second best matching cost of p, respectively. T_c is a small threshold value (set to $1e-3$ in our implementation) to avoid division by zero. $\gamma_p \in [0, 1]$ can be used to measure ambiguous matching caused by poor SNR. We experimentally found the confidence value itself is insufficient to model complex matching errors especially those caused by occlusions. To that end, we apply the left-right consistency check [7] to disambiguate occlusion pixels. For each pixel $p \in I$, if p violates the equality $D(p) = D'(p - D(p))$ then p is declared as occluded and added to the occlusion set O. We also observed that pixels with large disparity variations in a small neighborhood are prone to erroneously matched. Thus, if pixel p fails to satisfy the following inequality

$$|D(p) - \frac{\sum_{q \in \Phi_p} D(q)}{|\Phi_p|}| \leq 1, \tag{5}$$

it is treated as suspicious and added to the questionable set Q. Here Φ is a 3×3 neighborhood centered at pixel p. Note that both O and Q are $\emptyset$ at the beginning. Finally we define a binary map, denoted by M, on the reference image I. $M(p)$ is set to 1 if $p \in O \cup Q$ and 0 otherwise.

We define the reliability measure ρ_p for a pixel p using the conditional probability that p's estimated disparity is correct given γ_p and $M(p)$, as

$$\rho_p \propto P(\beta_p = 1|\gamma_p, M(p)). \tag{6}$$

By further assuming γ_p and $M(p)$ are independent variables we arrive at

$$\rho_p \propto P(\beta_p = 1|\gamma_p)P(\beta_p = 1|M(p)). \tag{7}$$

Likewise, p's unreliability measure $\overline{\rho_p}$ can be derived from equations 6 and 7 as,

$$\begin{aligned} \overline{\rho_p} &\propto P(\beta_p = 0|\gamma_p, M(p)) = P(\beta_p = 0|\gamma_p)P(\beta_p = 0|M(p)) \\ &= (1 - P(\beta_p = 1|\gamma_p))(1 - P(\beta_p = 1|M(p))). \end{aligned} \tag{8}$$

Our cost function U is constructed from equations 7 and 8 using the negative log-likelihood as

$$U_p(\beta_p) = \begin{cases} -\ln(\kappa_p \cdot \rho_p) & \beta_p = 1 \\ \\ -\ln(\kappa_p \cdot \overline{\rho_p}) & \beta_p = 0, \end{cases} \tag{9}$$

where κ_p in equation 9 is a normalization constant such that $\kappa_p(\rho_p + \overline{\rho_p}) = 1$.

Instead of using hard-coded $P(\beta_p = 1|\gamma_p)$ and $P(\beta_p = 1|M(p))$ we estimate the corresponding conditional probabilities based on stereo data sets with ground-truth disparities provided by [21,22]. In our experiment a total number of 13 stereo data sets are used to learn the parameters. We carefully select the training set so the matching difficulties of these stereo images span a wide range.

Since $M(p) \in \{0, 1\}$, $P(\beta_p = 1|M(p))$ can be easily estimated. After computing winner-take-all disparity maps $\{D, D'\}$ and the mask M, the conditional probability mass function is obtained by comparing masked and unmasked disparities with their ground-truth values and recording the correct rate. Our learned conditional probability mass function for $P(\beta_p = 1|M(p))$ is given in by:

$$P(\beta_p = 1|M(p)) = \begin{cases} 0.19 & M(p) = 1 \\ \\ 0.58 & M(p) = 0. \end{cases} \tag{10}$$

Learning $P(\beta_p = 1|\gamma_p)$ needs slightly more effort. Since $\gamma_p \in [0, 1]$, we discretize the continuous interval $[0, 1]$ into 200 bins. Pixel p is thus assigned to $bin_{\tilde{p}}$, whose corresponding value $val_{\tilde{p}}$ is most close to γ_p. In $bin_{\tilde{p}}$ the percentage of pixels that have correct estimated disparities is an approximation to $P(\beta_p = 1|val_{\tilde{p}})$ statistically. In Fig. 3 we plot the sampled conditional probability in red. There is a heavy tail in the plot because much fewer pixels are found with high matching confidence. Another interesting observation is that the probability actually

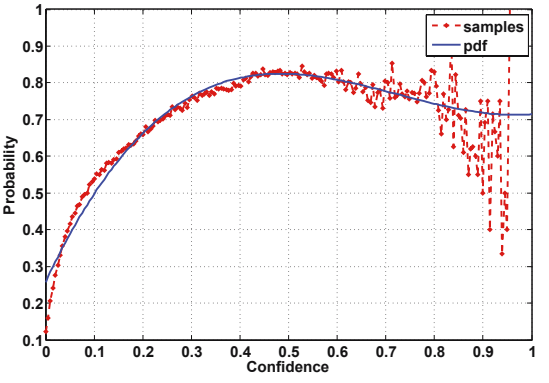

Fig. 3. The sampled conditional probabilities (red) and the cubic function computed via least square fitting to approximate the continuous pdf are given (blue)

starts decreasing when the confidence exceeds a certain value around 0.5. This phenomenon implies high confidences may partially result from occlusion in the image. To obtain a continuous probability density function (pdf) we apply a curve fitting using the sampled probability values given by bins that contain enough number of pixels. In this paper, bins whose number of pixels are less than 100 are discarded. Through the least square fitting we found the cubic function in 11 experimentally approximates the pdf quite well. The plot of this function is shown in Fig. 3 in blue.

$$P(\beta = 1|x) \approx 2.02x^3 - 4.38x^2 + 2.82x + 0.257 \tag{11}$$

Smoothness Term. The second term in equation 3 encourages spatial coherence and is helpful to remove matching outliers caused by noise and occluded pixels that survived the left-right check. The binary function V is defined as

$$V(\beta_p, \beta_q) = \lambda|\beta_p - \beta_q|. \tag{12}$$

The parameter λ controls the strength of the smoothness and is set to 0.5 throughout our experiments. Note that it is the only parameter designed experimentally in equation 3.

By treating the data term as an evidence function and the smoothness term as a compatibility function, equation 3 is the standard Gibbs function defined on a Markov network. As a two-label problem, the global optimal configuration which minimizes equation 3 can be exactly solved by GC. Although BP cannot guarantee the global optimal solution, our experiments show BP works equally successful on this problem in practice. In Fig. 4 we demonstrate the winner-take-all disparity map and the computed $GCPs$ of the Teddy data set using BP in the first two columns. The outliers in the candidate $GCPs$ are labeled in red. The matching density is 38% and the corresponding outlier percentage is 0.23%. It can be seen $GCPs$ from our approach are almost free of outliers.

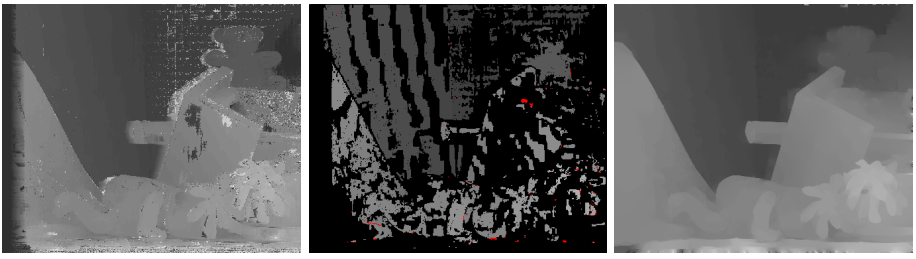

Fig. 4. From left to right: disparity map from winner-take-all method with cost aggregation; computed $GCPs$ of the Teddy data set and the outliers in the candidate $GCPs$ are labeled in red; $\overline{D}$ of the Teddy data set computed from the candidate $GCPs$ using adaptive propagation

2.3 Adaptive Propagation

The motivation of adaptive propagation is to allow propagation into regions where local matching costs are treated as unreliable by the proposed stable matching extraction algorithm. Our adaptive propagation algorithm is given as input the reference image I together with a semi-dense disparity map D_S on which the unreliable disparities D_S^U are left to be determined by those reliable disparities D_S^R. The objective is to infer D_S^U based on D_S^R and output a dense disparity map $\overline{D}$.

We wish to minimize the difference between the disparity of pixel p and the weighted average of the disparities at p's neighbors. A target cost function can be defined as

$$E(\overline{D}) = \sum_{p \in I} (\overline{D}(p) - \sum_{q \in N_p} \alpha_{pq} \overline{D}(q))^2, \tag{13}$$

where N_p is p's second order neighborhood system and α_{pq} is a weighting function that satisfies

$$\sum_{q \in N_p} \alpha_{pq} = 1. \tag{14}$$

The adaptation is based on the computation of the weighting function α. α_{pq} should be large when p and q have similar intensities and small otherwise. α controls the propagation strength so propagation stops near intensity edges, which are very likely to be depth discontinuities. In our experiments we adopt the weighting function defined in [8] as

$$\alpha_{pq} \propto 1 + \frac{(I(p) - \overline{I}_p)(I(q) - \overline{I}_p)}{\sigma_p^2}, \tag{15}$$

where $\overline{I}_p$ and σ_p^2 are the mean and variance of the intensities in a 5×5 window around p. By taking those stabled matched disparities D_S^R as constraints we can compute $\overline{D}$ by solving a system of sparse linear equations from a number of existing techniques [23]. When handling large stereo images, the resulting linear

system cannot be efficiently solved due to the large problem size. We therefore downsample both I and D_S to a reasonable size and upsample the solution to get $\overline{D}$. For example given a 1200×1000 image a downsampling factor 4 is used in our experiments. The corresponding $\overline{D}$ of the Teddy data set is shown in the third column of Fig. 4.

The cost function defined by equation 13 is not new. Levin et al. [8] use the same cost function to propagate color for colorization purpose. [24] minimizes similar function in segmentation. But based on our knowledge we believe it has not been used for stereo matching before. Different from tasks like colorization that needs only visually plausible result, stereo requires high accuracy estimate while such disparity propagation strategy clearly cannot guarantee enough accuracy. However, since our objective is not to locate the true disparity, this simple approach suits our problem pretty well by offering us a rough idea where the correct disparity is.

2.4 Search Range Selection

Given the disparity map D from winner-take-all together with the disparity map $\overline{D}$ from adaptive propagation our final per-pixel disparity search range is designed as follows: For each pixel p, first we compute a uncertainty radius Υ_p as

$$\Upsilon_p = \max(\frac{|D(p) - \overline{D}(p)|}{2}, 1). \tag{16}$$

Υ_p measures how well the disparity estimated from matching costs agrees with the disparity acquired from stably matched pixels. If Υ_p is large, we intuitively think the uncertainty is high and assign a large search range to p. Otherwise, we limit the search range for p since its matching cost is quite stable. The final disparity candidate set S_p is set to

$$\begin{aligned} S_p = &\{\max(d^{min}, D(p) - \Upsilon_p), \ldots, \min(D(p) + \Upsilon_p, d^{max})\} \bigcup \\ &\{\max(d^{min}, \overline{D}(p) - \Upsilon_p), \ldots, \min(\overline{D}(p) + \Upsilon_p, d^{max})\}. \end{aligned} \tag{17}$$

3 Experimental Results

In this section we report the results of applying our method on various stereo images and comparing against some existing algorithms that are related to our work. The side width of the window for adaptive cost aggregation is set to 33. Two constant aggregation parameters γ_c and γ_g defined in [18,19] are set to 12 and 40 respectively throughout.

We first evaluate the effectiveness of our stable matching extraction algorithm using ground-truth disparities and compare it with several semi-dense stereo algorithms. Note that there is no intersection between the training set and data sets we used to evaluate our algorithm. Table 1 shows the comparisons of different approaches including the current state-of-the-art [11]. Density D is the

Table 1. Performance comparison of the proposed stable matching detection method with other semi-dense stereo approaches

Algorithm	Tsukuba		Sawtooth		Venus		Map	
	$D(\%)$	e(%)	$D(\%)$	e(%)	$D(\%)$	e(%)	$D(\%)$	e(%)
Ours	**72**	**0.22**	**66**	**0.07**	**53**	**0.08**	**37**	**0.04**
RDP [11]	76	0.32	89	0.07	73	0.18	86	0.07
Veksler [13]	75	0.36	87	0.54	73	0.16	87	0.01
Sara [12]	45	1.40	52	1.60	40	0.80	74	0.30

Table 2. The performance of the proposed search space reduction algorithm. $\Re$ is the percentage of the reduced problem size. $\hbar$ is the percentage of pixels whose correct disparity values are preserved by our reduced search space.

	Tsukuba	Venus	Teddy	Cones
$\Re(\%)$	70.6	75.1	80.3	83.4
$\hbar(\%)$	99.1	99.7	97.3	97.5

percentage of reliable matching detected by the algorithm and error e is the percentage of wrong estimations (error $> \pm 1$ true disparity in non-occluded areas) in the semi-dense disparity maps. Numbers other than ours are from [11]. Note that all other three algorithms are designed for solving the stereo matching problem. Ours, however, only decides whether the disparity given by the best cost value is correct or not. This evaluation shows that our approach produces the most error-free disparity maps (in all but one case) and can preserve reasonable densities.

Table 2 shows our search space reduction rates. The stereo images we used for the test are the four standard data sets provided by the Middlebury stereo evaluation system [1]. Besides the reduction rate $\Re$ defined in equation (2), we further define the hit rate $\hbar$ as the percentage of pixels whose correct disparity values (within ± 1 true disparity) are preserved by our reduced search space. As can be seen, our algorithm is able to achieve 77% reduction rate while preserve 98.5% correct disparities in the reduced search space on average. The results are evaluated on the whole image domain including occlusion areas.

We also modify the traditional BP implementation such that each node can take different number of labels. By doing so the message length is no longer globally defined and memory requirement is greatly reduced. We use our BP implementation to compare the reconstruction quality given by full search range and the reduced search space generated from our framework. The evaluation results from [1] are demonstrated in Table 3. The second row shows the error percentages in non-occlusion areas using the full search space and the third row provides the corresponding error rates from our reduced space. As shown although the quality is not as good as the full search algorithm the differences are not significant. Considering the size of our search space is only 20% to 30% of the full search range these numbers are actually quite satisfactory. Some disparity maps are given in Fig. 5 for comparison.

Table 3. The comparison of reconstruction quality by using full search space (second row) and our reduced search space (third row). Disparity maps generated from BP are evaluated using Middlebury benchmark.

	Tsukuba	Venus	Teddy	Cones
Full range e(%)	1.06	0.78	7.59	5.26
Reduced space e(%)	1.10	1.02	7.96	6.49

Fig. 5. The first row shows results from full range algorithm and the second row demonstrates results from our reduced search space

At last we test our algorithm using high resolution stereo images. The image resolution is 1200×1000 and the full disparity search range is 120 pixels. Traditional Belief Propagation will require more than $2GB$ memory to only store the floating point messages. Our proposed algorithm, however, is able to reduce over 80% disparity search space so the remaining problem is solvable by ordinary computers[1].

4 Conclusions

In this paper we propose a novel algorithm for reducing search space for MRF-based stereo. The algorithm is simple yet very effective and can be used as a preprocessor for both BP and GC-based solvers. Interesting future work includes: applying our algorithm to facilitate high-resolution, large scale stereo reconstruction; investigating the performance gain of Belief Propagation and Graph Cuts given the reduced search spaces.

[1] The high resolution stereo images and the resulting disparity maps are available at www.vis.uky.edu/∼wangl

Acknowledgements

This work is supported in part by a US National Science Foundation grant IIS-0448185 and a grant from the US Department of Homeland Security.

References

1. Scharstein, D., Szeliski, R.: `vision.middlebury.edu/stereo/`
2. Sun, J., Zheng, N.N., Shum, H.Y.: Stereo matching using belief propagation. IEEE Trans. on Pattern Analysis and Machine Intelligence 25(7), 787–800 (2003)
3. Boykov, Y., Veksler, O., Zabih, R.: Markov random fields with efficient approximations. In: Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, pp. 648–657 (1998)
4. Boykov, Y., Veksler, O., Zabih, R.: Fast approximate energy minimization via graph cuts. IEEE Trans. on Pattern Analysis and Machine Intelligence 23(11), 1222–1239 (2001)
5. Boykov, Y., Kolmogorov, V.: An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. IEEE Trans. on Pattern Analysis and Machine Intelligence 26(9), 1124–1137 (2004)
6. Bobick, A.F., Intille, S.S.: Large occlusion stereo. Int. J. of Computer Vision, 181–200 (November 1999)
7. Egnal, G., Wildes, R.P.: Detecting binocular half-occlusions: empirical comparisons of five approaches. IEEE Trans. on Pattern Analysis and Machine Intelligence 24(8), 1127–1133 (2002)
8. Levin, A., Lischinski, D., Weiss, Y.: Colorization using optimization. In: Proc. of ACM SIGGRAPH, pp. 689–694 (August 2004)
9. Veksler, O.: Reducing search space for stereo correspondence with graph cuts. In: British Machine Vision Conf., vol. 2, pp. 709–718 (2006)
10. Yu, T., Lin, R.S., Super, B., Tang, B.: Efficient message representations for belief propagation. In: Proc. of Intl. Conf. on Computer Vision (2007)
11. Gong, M., Yang, Y.H.: Fast unambiguous stereo matching using reliability-based dynamic programming. IEEE Trans. on Pattern Analysis and Machine Intelligence 27(6), 998–1003 (2005)
12. Sara, R.: Finding the largest unambiguous component of stereo matching. In: Proc. of Europ. Conf. on Computer Vision, pp. 900–914 (2002)
13. Veksler, O.: Extracting dense features for visual correspondence with graph cuts. In: Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, vol. 1, pp. 689–694 (2003)
14. Cech, J., Sara, R.: Efficient sampling of disparity space for fast and accurate matching. In: Intl. Workshop on Benchmarking Automated Calibration, Orientation, and Surface Reconstruction from Images (2007)
15. Zhang, Z., Shan, Y.: A progressive scheme for stereo matching. In: Europ. Workshop on 3D Structure from Multiple Images of Large-Scale Environments (July 2000)
16. Szeliski, R., Zabih, R., Scharstein, D., Veksler, O., Kolmogorov, V., Agarwala, A., Tappen, M.F., Rother, C.: A comparative study of energy minimization methods for markov random fields. In: Proc. of Europ. Conf. on Computer Vision, vol. 2, pp. 19–26 (2006)

17. Birchfield, S., Tomasi, C.: A pixel dissimilarity measure that is insensitive to image sampling. IEEE Trans. on Pattern Analysis and Machine Intelligence 20(4), 401–406 (1998)
18. Wang, L., Liao, M., Gong, M., Yang, R.: High-quality real-time stereo using adaptive cost aggregation and dynamic programming. In: Intl. Symposium on 3D Data Processing, Visualization and Transmission, pp. 798–805 (2006)
19. Yoon, K.J., Kweon, I.S.: Locally adaptive support-weight approach for visual correspondence search. In: Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, pp. 924–931 (2005)
20. Yang, Q., Wang, L., Yang, R., Stewenius, H., Nister, D.: Stereo matching with color-weighted correlation, hierachical belief propagation and occlusion handling. In: Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, pp. 2347–2354 (2006)
21. Hirschmuller, H., Scharstein, D.: Evaluation of cost functions for stereo matching. In: Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (2007)
22. Scharstein, D., Pal, C.: Learning conditional random fields for stereo. In: Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (2007)
23. Press, W., Flannery, B., Teukolsky, S., Vetterling, W.: Numerical recipes inc. (1988)
24. Shi, J., Malik, J.: Normalized cuts and image segmentation. IEEE Trans. on Pattern Analysis and Machine Intelligence 22(8), 888–905 (2000)