IEICE Electronics Express, Vol.8, No.2, 57–63

Improving stereo matching with symmetric cost functions

Kuk-Jin Yoon[1a)] **and Sung-Kee Park**[2]
[1] *Computer Vision Laboratory, GIST*
261 Cheomdan-gwagiro, Buk-gu, Gwangju 500-712 Republic of Korea
[2] *Cognitive Robotics Center, KIST*
39-1 Haweolgok-dong, Sungbuk-ku, Seoul 136-791, Republic of Korea
a) *kjyoon@gist.ac.kr*

Abstract: In this paper, we propose new symmetric cost functions for global stereo methods. We first present a symmetric data cost function for the likelihood and then propose a symmetric discontinuity cost function for the prior in the MRF model for stereo. In defining cost function, both the reference image and the target image are taken into account to improve performance without modeling half-occluded pixels explicitly. The performance improvement of stereo matching due to the proposed symmetric cost functions is verified by applying the proposed symmetric cost functions to the belief propagation (BP) based stereo method.

Keywords: stereo vision, stereo matching
Classification: Science and engineering for electronics

References

[1] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *IJCV*, vol. 47, no. 1–3, pp. 7–42, 2002.
[2] M. F. Tappen and W. T. Freeman, "Comparison of graph cuts with belief propagation for stereo, using identical MRF parameters," *ICCV*, vol. 2, pp. 900–906, 2003.
[3] K.-J. Yoon and I. S. Kweon, "Adaptive support-weight approach for correspondence search," *PAMI*, vol. 28, no. 4, pp. 650–656, 2006.
[4] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *PAMI*, vol. 23, no. 11, pp. 1222–1239, 2001.
[5] J. Sun, N.-N. Zheng, and H.-Y. Shum, "Stereo matching using belief propagation," *PAMI*, vol. 25, no. 7, pp. 787–800, 2003.
[6] J. Sun, Y. Li, S. B. Kang, and H.-Y. Shum, "Symmetric stereo matching for occlusion handling," *CVPR*, vol. 2, pp. 399–406, 2005.

1 Introduction

Many global stereo methods have recently achieved good results by model-

DOI: 10.1587/elex.8.57
Received November 08, 2010
Accepted December 03, 2010
Published January 25, 2011

IEICE Electronics Express, Vol.8, No.2, 57–63

ing a disparity surface as a Markov random field (MRF) and by solving an optimization problem [1]. They mainly focus on how to minimize conventional cost functions efficiently to improve performance. However, lower cost solutions do not always correspond to better performance as pointed out in [2]. Therefore, it is more important to define cost functions to be minimized than to improve optimization techniques for improving performance. Nevertheless, there is a relatively small amount of work on defining cost functions well.

In this paper, we propose new symmetric cost functions for improving the performance of global stereo methods. We first present a symmetric data cost function for the likelihood and then propose a symmetric discontinuity cost function for the prior in the MRF model for stereo. In defining cost functions, we take both the reference image and the target image into account aiming at improving performance without modeling half-occluded pixels explicitly and without using color segmentation, which are also difficult problems.

2 MRF model for stereo matching

Although global stereo methods formulate the stereo problem in various ways, the MRF formulation is most general. Bayesian stereo matching can be formulated as a maximum a posteriori MRF (MAP–MRF) problem. Given a rectified stereo pair of images, the stereo problem can be modeled as

$$P(\mathbf{D}|I) = \frac{P(I|\mathbf{D})P(\mathbf{D})}{P(I)} \tag{1}$$

where $\mathbf{D}$ is the smooth disparity field of the reference image and I is a pair of input stereo images (i.e., $I = (I_L, I_R)$ where I_L is the reference image and I_R is the target image). The goal of the stereo problem is to find the disparity field $\mathbf{D}$ that maximizes Eq. (1) for given I as

$$\mathbf{D}_{opt} = \arg\max_{\mathbf{D}} P(\mathbf{D}|I) = \arg\max_{\mathbf{D}} \frac{P(I|\mathbf{D})P(\mathbf{D})}{P(I)} \tag{2}$$

Here, $P(I|\mathbf{D})$ is referred to as the likelihood and $P(\mathbf{D})$ is referred to as the prior. When assuming that the observation follows an independent identical distribution (i.i.d.), the likelihood $P(I|\mathbf{D})$ in Eq. (1) can be expressed as

$$P(I|\mathbf{D}) \propto \prod_{p} \exp(-\phi(p, d_p, I)) \tag{3}$$

where $\phi(p, d_p, I)$ is the cost function of pixel p with disparity d_p given observation I. Therefore, the likelihood $P(I|\mathbf{D})$ is related to the data cost when pixel p has the disparity d_p with given images.

The Markov property asserts that the probability of each site in a field depends only on its neighboring sites. By specifying the first order neighborhood of pixel p, $N(p)$, the prior can be expressed as

$$P(\mathbf{D}) \propto \prod_{p} \prod_{q \in N(p)} \exp(-\psi_c(d_p, d_q)) \tag{4}$$

DOI: 10.1587/elex.8.57
Received November 08, 2010
Accepted December 03, 2010
Published January 25, 2011

IEICE Electronics Express, Vol.8, No.2, 57–63

where q is the neighboring pixel of p in $N(p)$ and $\psi_c(d_p, d_q)$ is the joint clique potential function of two disparities d_p and d_q.

By combining Eq. (3) and Eq. (4), we can simply get the following equation.

$$-\ln(P(D|I)) \propto \sum_p \phi(p, d_p, I) + \sum_p \sum_{q \in N(p)} \psi_c(d_p, d_q) \tag{5}$$

Here, Eq. (5) can be rewritten in terms of cost functions as

$$E(\mathbf{D}|I) = \sum_p D(p, d_p, I) + \sum_p \sum_{q \in N(p)} V(d_p, d_q) \tag{6}$$

where

$$E(\mathbf{D}|I) \propto -\ln(P(\mathbf{D}|I)) \tag{7}$$

$$D(p, d_p, I) = \phi(p, d_p, I) \tag{8}$$

$$V(d_p, d_q) = \psi_c(d_p, d_q) \tag{9}$$

$E(\mathbf{D}|I)$ is a global cost to be minimized to obtain a disparity map. $D(p, d_p, I)$ is referred to as the data cost that measures the cost of assigning disparity d_p to pixel p with given I. On the other hand, $V(d_p, d_q)$ is referred to as the discontinuity cost that measures the cost of assigning disparities d_p and d_q to two neighboring pixels p and q. Our goal is then to define $D(p, d_p, I)$ and $V(d_p, d_q)$ in consideration of both the reference image and the target image together to improve the performance of global methods without modeling half-occluded pixels explicitly and without using color segmentation.

3 Symmetric data cost function for the likelihood

Most global methods compute the data cost using an individual pixel intensity (or color) and then try to solve the image ambiguity using global reasoning with a smoothness constraint, which results from the ambiguous local appearances of image pixels owing to image noise and insufficient/repetitive texture. However, it is still difficult to get an accurate disparity map when there are severe errors in the data cost.

To get the reliable data cost, it may be useful to use local support windows as in local methods. However, local support windows cause the foreground fattening phenomenon resulting in severe errors at depth discontinuities. In this work, we use the symmetric data cost function that we have proposed in [3]. This method provides the reliable data cost in consideration of the reference image and the target image even near depth discontinuities even when using large local support windows.

4 Symmetric discontinuity cost function for the prior

Among the two cost functions in the MRF formulation, the discontinuity cost function between nodes, $V(d_p, d_q)$, determines how support is aggregated from neighboring nodes. This cost function is directly related to the smoothness constraint. In most global methods, it is generally computed by

DOI: 10.1587/elex.8.57
Received November 08, 2010
Accepted December 03, 2010
Published January 25, 2011

IEICE Electronics Express, Vol.8, No.2, 57–63

using the truncated linear model or the Potts model [4] assuming piecewise constant disparities. The typical Potts model can be expressed as

$$V(d_p, d_q) = \begin{cases} 0 & \text{if } d_p = d_q \\ \rho(\Delta C) & \text{otherwise} \end{cases} \tag{10}$$

The function $\rho(\Delta C)$ is defined in terms of the magnitude of image gradient between p and q, ΔC, as

$$\rho(\Delta C) = \begin{cases} P \times s & \text{if} \Delta C < T \\ s & \text{otherwise} \end{cases} \tag{11}$$

where T is a magnitude threshold and s is a penalty term for violating the smoothness constraint. P is a penalty term that increases the penalty when the gradient magnitude is small. This form of the smoothness constraint makes depth discontinuities coincide with color or intensity discontinuities.

The problem of the conventional Potts model is that it is based on only the reference image. This may cause the erroneous discontinuity cost and result in errors at depth discontinuities. To reduce the errors in the discontinuity cost owing to half-occluded pixels, it may be useful to consider half-occluded pixels formally in the MRF formulation. However, modeling the occlusion field and detecting half-occluded pixels are also difficult problems. To solve this problem, we propose the new symmetric Potts model by redefining Eq. (10) and Eq. (11) while taking the reference and the target images into account together as

$$V_s(d_p, d_q) = \begin{cases} 0 & \text{if } d_p = d_q \\ \rho_{d_p}(\Delta C_r, \Delta C_t) & \text{otherwise} \end{cases} \tag{12}$$

$$\rho_{d_p}(\Delta C_r, \Delta C_t) = \begin{cases} P_r \times P_t \times s & \text{if} \Delta C_r < T, \Delta C_t < T \\ P_r \times s & \text{if} \Delta C_r < T, \Delta C_t \geq T \\ P_t \times s & \text{if} \Delta C_r \geq T, \Delta C_t < T \\ s & \text{otherwise} \end{cases} \tag{13}$$

Here, ΔC_r is the magnitude of color gradient between p and q in the reference image and ΔC_t is the magnitude of color gradient between $\bar{p}_{d_p}$ and $\bar{q}_{d_p}$ in the target image when the disparity of p is d_p. $\bar{p}_{d_p}$ and $\bar{q}_{d_p}$ are the corresponding pixels in the target image when p and q in the reference image have the disparity d_p. P_r and P_t are penalty terms that increase the penalty when the gradient magnitude is small. We can see that, as in the data cost function, Eq. (13) is also symmetric — when the reference image and the target image are switched, the form of this discontinuity cost function does not change. In addition, it is worth of note that $V_s(d_p, d_q)$ is dependent on actual d_p and d_q values while $V(d_p, d_q)$ in the conventional Potts model is not. In addition, it is possible to redefine the linear model in the symmetric form as the Potts model. The main advantage of the proposed symmetric discontinuity cost function is that we can improve the performance of global stereo methods at depth discontinuities without modeling half-occluded pixels explicitly and without using color segmentation. This is because the effect of half-occluded

DOI: 10.1587/elex.8.57
Received November 08, 2010
Accepted December 03, 2010
Published January 25, 2011

IEICE Electronics Express, Vol.8, No.2, 57–63

pixels can be reduced by considering both images together when computing the discontinuity cost.

5 Experiments

The proposed symmetric cost functions are simple and, therefore, can be easily applied to global stereo methods without much modification. To verify the performance improvement by the proposed symmetric cost functions, we applied the proposed symmetric cost functions to the BP-based method implemented by Tappen in [2] and we used well-known testbed images with ground truth [1]. The stereo method is run with a constant parameter setting across all images ($T = 8$, $P_r = P_t = 2$, $s = 1.1$, $\gamma_c = 4.0$, $\gamma_p = 15.5$, window size $= 31 \times 31$).

We first applied the proposed symmetric data cost function and the proposed symmetric discontinuity cost function separately to the BP-based method to check the performance improvement due to each cost function. Fig. 1 shows the matching results for the 'Tsukuba' data set according to the applied symmetric cost function and Fig. 2 shows the matching results for testbed images with both the proposed symmetric cost functions. The performance according to the applied cost functions is summarized in Table I. The numbers in Table I represent the percentage of bad pixels. We can see that both the proposed symmetric cost functions really improve the performance of the BP-based method for all testbed images.

We then compared the performance of the BP-based method using the proposed symmetric cost functions with the performance of other state-of-the-art BP-based methods as shown in Table I, although the run parameters in our experiments are not optimal. The performance of the BP-based method with the proposed symmetric cost functions is comparable to the performance of the state-of-the-art methods even without modeling half-occluded pixels explicitly and without using color segmentation. However, the result for the 'Map' data set is worse than other methods. This is because the 'Map' images are highly textured while the proposed symmetric cost functions are dependent on the color (or intensity) and disparity gradient in both images.

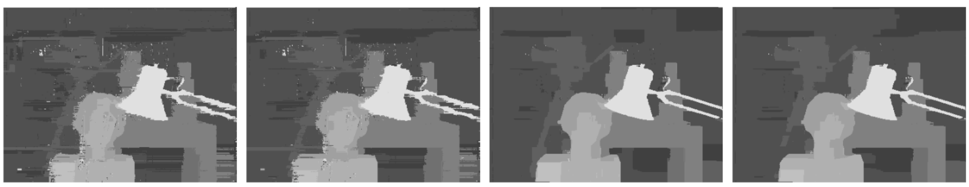

Fig. 1. Matching results according to the applied cost functions. From left to right, conventional, symmetric disc. cost, symmetric data cost, and symmetric disc.&data cost

DOI: 10.1587/elex.8.57
Received November 08, 2010
Accepted December 03, 2010
Published January 25, 2011

IEICE Electronics Express, Vol.8, No.2, 57–63

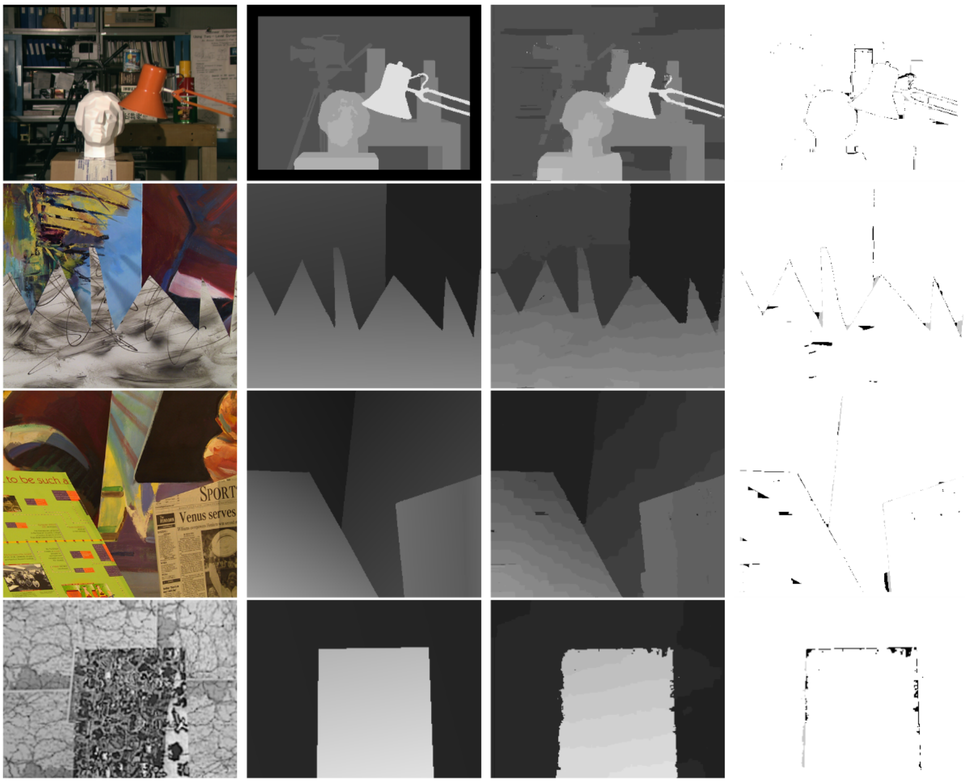

Fig. 2. Dense disparity maps for testbed images. From left to right, left image, ground truth, our result, and bad pixels (error > 1)

6 Conclusion

In this paper, we have proposed symmetric cost functions for both the likelihood and the prior in the MRF model for stereo, aiming at improving performance. In defining cost functions, we took both the reference image and the target image into account. We finally verified the performance improvement of stereo matching due to the proposed symmetric cost functions by applying the proposed symmetric cost functions to the BP-based stereo method. Experimental results for standard testbed images show that the performance of the BP based stereo method is greatly improved by the proposed symmetric cost functions.

DOI: 10.1587/elex.8.57
Received November 08, 2010
Accepted December 03, 2010
Published January 25, 2011

IEICE Electronics Express, Vol.8, No.2, 57–63

Table I. Performance according to the applied cost function

	Tsukuba			Sawtooth			Venus			Map	
	nonocc	untex	disc	nonocc	untex	disc	nonocc	untex	disc	nonocc	disc
BP with symm. cost	**1.07**	**0.35**	**6.05**	**0.69**	**0.00**	**4.17**	**0.64**	**0.62**	**3.05**	**1.06**	**13.20**
symm. data only	1.15	0.41	6.49	0.75	0.16	4.18	0.78	0.87	2.91	1.10	13.58
symm. disc. only	2.52	1.45	9.69	5.24	0.49	18.15	4.72	5.37	11.63	43.21	54.89
conventional	2.76	1.62	9.97	7.06	2.51	20.09	9.23	14.07	14.51	46.41	57.18
BP [5]	1.61	0.66	9.17	0.85	0.37	7.92	1.17	1.00	12.87	0.67	3.42
BP+segm. [5]	1.15	0.42	6.31	0.98	0.30	4.83	1.00	0.76	9.13	0.84	5.27
symm. BP (no segm.) [6]	1.01	0.28	5.79	0.57	0.05	3.46	0.66	0.71	8.72	0.14	1.97
symm. BP+segm. [6]	0.97	0.28	5.45	0.19	0.00	2.09	0.16	0.02	2.77	0.16	2.20

DOI: 10.1587/elex.8.57
Received November 08, 2010
Accepted December 03, 2010
Published January 25, 2011