Census Transform based Stereo Matching with Gradient Stereo Images

Jaeryun Ko and Yo-Sung Ho
Gwangju Institute of Science and Technology (GIST)
123 Cheomdan-gwagiro, Buk-gu, Gwangju, 500-712, Republic of Korea
Telephone +82-62-715-2258, Fax: +82-62-714-3164
Email:{jrko,hoyo}@gist.ac.kr

***Abstract*—In this paper, we propose global stereo matching method with census transform as the matching cost, not only on the stereo intensity images. In order to get more accurate disparity values on the area of depth discontinuity, we applied census transform on the gradient stereo images. Determining the matching cost, we applied predetermined weight value to both operations of census transform. Finally we can get better results of disparity map than the same method without gradient images from our experiments.**

Keywords—stereo matching; disparity map; census transform; belief propagation; gradient image;

I. INTRODUCTION

Disparity map generation from stereo images is one of the fundamental process for three-dimensional reconstruction and has been researched into various methods. One of these methods is called as stereo matching, which is estimating the depth of projected view from stereo images of different views. Generally, it is divided into two classes; one is local method and another is global method.

Local methods on stereo matching are estimating the disparity with comparison of the matching cost from views on the right and the left. It is simple to implement and fast to be suitable for real-time applications. For calculating matching costs, the size of the matching window must be determined at first, which has an effect on the quality of disparity map. The disparity map from using local methods usually shows low accuracy than global-based methods.

Global methods on stereo matching are based on Markov random fields and estimating proper disparity values by minimizing predefined energy functions. Energy function can be minimized by using optimization algorithms such as belief propagation, graph cut and dynamic programming technique. Energy optimization process takes the relationship between the entire pixels in the image into account, so it is slower than local methods.

In this paper, we applied gradient stereo images with census transform to improve the accuracy of disparity values on the area of depth discontinuities. Gradient images contain the edge information of original stereo images and it is related to the depth discontinuity regions. Basically we exploited belief propagation with intensity-based census transform to calculate the matching cost on stereo matching. In our experiments we can get the better results than without using gradient images.

II. GLOBAL METHOD ON STEREO MATCHING

A. Belief Propagation (BP)

Belief propagation on stereo matching is going along Markov network which is like Fig. 1. Each node on the Markov network, that is equal to each pixel, send the value called as *message* to neighboring pixels [1]. Disparity values are assigned to corresponding pixels with calculating messages from neighboring pixels.

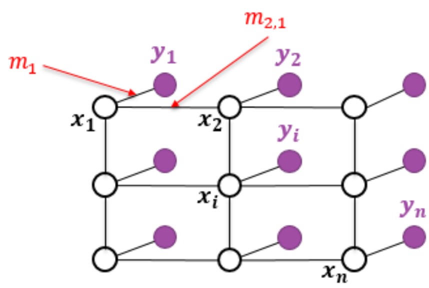


Fig. 1. Markov network in the image

In Fig. 1, disparity candidate $y_i, i = 1, \ldots, n$ has the matching cost function. m_i represents message from node y_i to node x_i. m_{ij} is message from node x_i to node x_j. These nodes must be initialized into uniform distribution as the first step of belief propagation.

For the next step, the updating process iterates with two costs for matching until the messages from neighboring nodes converge through Eq. (1). $D_i(y_i)$ is the matching cost of assigning y_i to node i. We used census transform as the matching cost $D_i(y_i)$. $V(y_i, y_j)$ represents the smoothness cost, which does not change smoothly on large discontinuities such as object boundaries. For the smoothness term, we used the truncated linear model like Eq. (2), where c is the constant that controls the rate of increase and k is the upper threshold. After the updating process, each belief $b_i(x_i)$ of the corresponding pixel i is calculated by using Eq. (3). Finally the disparity value with corresponding pixel is determined to x_i which minimizes the belief of corresponding pixel.

$$\mathrm{m}_{\mathrm{ij}}^{t} = \min_{y_i}[V(y_i, y_j) + D_i(y_i) + \textstyle\sum_{s \in N(i)\backslash j} m_{s,i}^{t-1}] \quad (1)$$

$$V(y_i, y_j) = \min(c|y_i - y_j|, k) \quad (2)$$

$$\mathrm{b_i}(y_i) = D_i(y_i) + \textstyle\sum_{i \in N(j)} m_{j,i}^{t} \quad (3)$$

B. Census Transform with Gradient Images

Census Transform uses relative intensity of input images which performs robust under different absolute intensities of input images and noises [2]. Function C(P) is census transform of a pixel P in Eq. (4), where $\otimes$ denotes concatenation, D means non-parametric window around P and ξ is defined by Eq. (5). This function maps the local neighborhood of surrounding a pixel P to a bit string representing the set of neighboring pixels, whose intensity is less than that of P.

$$\mathrm{C(P)} = \otimes_{[\mathrm{i,j}] \in D}\, \xi(P, P + [i, j]) \quad (4)$$

$$\xi(P, P + [i, j]) = \begin{cases} 1, if\ P > P + [i, j] \\ 0, otherwise \end{cases} \quad (5)$$

Census transform converts relative intensity difference to 0 or 1 into one-dimensional vector. Two pixels of census transformed images are compared for similarity using Hamming distance, i.e. the number of bits that differ in the two bit strings as shown in Fig. 2. To compute correspondence, the Hamming distance is minimized after applying the census transform.

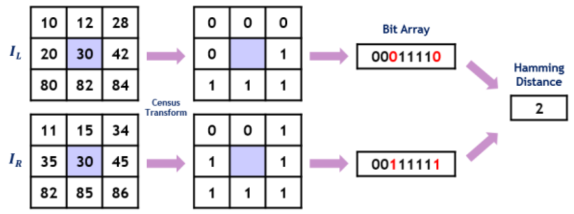


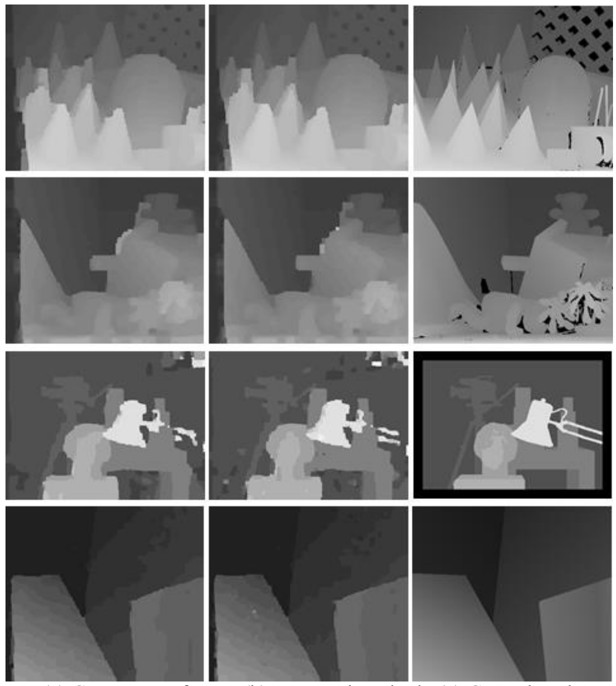
Fig. 2. Census transform from windows of stereo image

For improving the possibility of assigning more accurate disparity into the pixels near the depth discontinuities, we applied census transform to gradient images generated from original stereo images and exploit it to calculate the matching cost [3]. First we obtain gradient images $\mathrm{G_x}$ from input stereo images by applying horizontal Sobel derivative operator. Then same as census transform to intensity images, we get the Hamming distance of $\mathrm{G_x}$ to define the matching cost with weights like Eq. (6).

$$\mathrm{Cost}(\mathrm{x_i}, d) = \alpha Hamming_I + (1 - \alpha) Hamming_{G_x} \quad (6)$$

III. Experiment Results

In order to compare proposed improvement method with using census transform based stereo matching without gradient images, we use for stereo images, Cones, Teddy, Tsukuba, and Venus which are used in [4]. Window size of census transform is fixed to 5 × 5 and the weight α in the matching cost is determined to 0.6. We use hierarchical belief propagation as one type of belief propagation implements. Fig. 3 shows how well the proposed method with gradient images makes just a little better result than the ordinary one without them. In the last row of the result images, proposed method shows better depth discontinuity smoothness with the object at the right bottom. We can get more accurate disparity values on the depth discontinuity area like object boundaries as well.



(a) Census transform (b) Proposed method (c) Ground truth

Fig. 3. Census transform from windows of stereo image

IV. Conclusion

Census transform as a matching cost function usually shows good performances for global method stereo matching with belief propagation. In order to get more accurate disparity values on the area of depth discontinuity, we proposed global stereo matching with census transform on gradient stereo images. We modified the matching cost with added Hamming distance of census-transformed gradient images. Eventually we can get more accurate disparity values near the depth discontinuities

Acknowledgment

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Science, ICT & Future Planning(2011-0030079).

References

[1] S. Yedidia, W. T. Freeman, and Y. Weiss, "Understanding belief propagation and its generalizations," Exploring Artificial Intelligence in the New Millenium, pp. 239-269, 2003.

[2] S. Patil, J. S. Nadar, J. Gada, S. Motghare and S. S. Nair, "Comparison of Various Stereo Vision Cost Aggregation Methods," International Journal of Engineering and Innovative Technology, vol. 2, pp. 222-226, 2013.

[3] C. Rhemann, A. Hosni, M. Bleyer, C. Rother, and M. Gelautz, "Fast Cost-Volume Filtering for Visual Correspondence and Beyond," Computer Vision and Pattern Recognition, pp. 3017-3024, 2011.

[4] D. Scharstein, and R. Szeliski "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," International Journal of Computer Vision, vol 47, no. 1-3, pp. 7-42, 2002