View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by DSpace@MIT

Except for this watermark, it is identical to the version available on IEEE Xplore.

Depth Map Estimation and Colorization of Anaglyph Images Using Local Color Prior and Reverse Intensity Distribution

Williem
Inha University
Incheon 402-751, Korea
williem.pao@gmail.com

Ramesh Raskar
MIT Media Lab
Cambridge, MA 02139, USA
raskar@media.mit.edu

In Kyu Park
Inha University
Incheon 402-751, Korea
pik@inha.ac.kr

Abstract

In this paper, we present a joint iterative anaglyph stereo matching and colorization framework for obtaining a set of disparity maps and colorized images. Conventional stereo matching algorithms fail when addressing anaglyph images that do not have similar intensities on their two respective view images. To resolve this problem, we propose two novel data costs using local color prior and reverse intensity distribution factor for obtaining accurate depth maps. To colorize an anaglyph image, each pixel in one view is warped to another view using the obtained disparity values of non-occluded regions. A colorization algorithm using optimization is then employed with additional constraint to colorize the remaining occluded regions. Experimental results confirm that the proposed unified framework is robust and produces accurate depth maps and colorized stereo images.

1. Introduction

During the last couple of decades, 3D technology has become popular in both research activities and consumer applications. In consumer electronics, polarized filter glass and shutter glass have been widely used to deliver the stereoscopic experience to 3D TV users. A more primitive method involves anaglyph imagery, which has been used to implement stereoscopic visualization in a cost effective way using color filtered glasses. Anaglyph is often used to distribute 3D videos or images in the online video/image database. Besides for image size compression, it is a useful method for general consumers who want to enjoy 3D experience without expensive devices.

An anaglyph image packs partial information of stereo images in a single color image, which typically consists of red (only the red channel) from the left image and cyan (blue and green channels) from the right image. Consequently, it suffers from missing color information and it is difficult to process the anaglyph image using conventional computer vision algorithms. In early research, several studies have aimed to colorize the pixels in the missing channels [10, 13]. However, the state-of-the-art technique in anaglyph colorization [10] is still inaccurate due to the high dependency on the initial correspondence map. Note that the dense matching algorithm [14] is erroneous because of the nature of anaglyph image. Thus, it is necessary to develop a reliable anaglyph matching method to restore the original color of stereo images accurately.

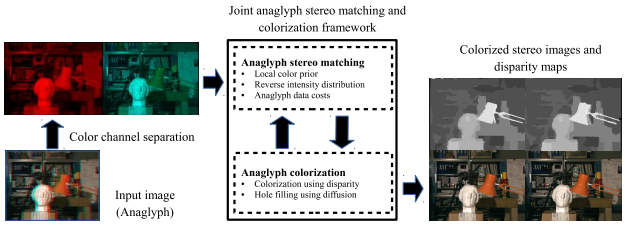


Figure 1: Pipeline of the proposed algorithm

Typical stereo matching approaches have been built on the photometric consistency assumption that requires the corresponding points of a stereo pair to have similar intensity values [18]. For that reason, anaglyph stereo matching becomes a difficult problem because half of the color channels are completely missing. To our best knowledge, there has been no previous anaglyph stereo matching algorithm that can obtain accurate disparity map.

In this paper, we propose an iterative joint method to not only compute the disparity map but also simultaneously colorize the missing color information of the anaglyph image. The overview of the proposed framework is shown in Figure 1. An accurate disparity map is estimated by employing two novel anaglyph data costs which are based on *local color prior* (LC) and *reverse intensity distribution* (RID). To restore the missing color, we first transfer the known color information from one image to another using the obtained disparity. We then colorize the remaining pixels in occluded regions using the diffusion-based colorization. A novel weight kernel function based on the color similarity is introduced to achieve accurate colorization. In summary, this paper makes the following contributions.

- Provide an iterative framework for obtaining an accurate disparity map and colorized stereo images from a single anaglyph image.
- Develop the new idea of (i) local color prior for pseudo color reconstruction of missing channels and (ii) reverse intensity distribution for cross color channel correlation.
- Provide a pair of robust anaglyph data costs to compute the dense stereo correspondence and their integration.

2. Related Works

This paper relates to previous studies in stereo matching with radiometric differences and colorization. Hirschmüller and Scharstein [8] evaluated a few stereo matching costs under radiometric differences. It is concluded that the census filter [21] and mutual information (MI) [11] are the most robust techniques that outperform the others. Heo *et al.* [6, 7] proposed the adaptive normalized cross correlation (ANCC) and modified mutual information to solve the stereo matching problem with illumination changes. Although their method obtains an accurate disparity with the existence of significant photometric differences, it cannot be directly applied to anaglyph stereo matching because their method utilizes all information of three color channels which are assumed to be correlated with each other. Bando *et al.* [2] proposed an algorithm to extract a depth map from the coded aperture image. However, this method is not accurate, especially when the scene is complex and the anaglyph pair has large intensity variance.

During the last decade, several algorithms have been introduced to colorize grayscale images or to edit the color layout of images. Levin *et al.* [13] model the colorization problem as an energy minimization function so that sparse user scribbles are propagated to the whole image. The energy function is based on the sparse affinity matrix which encodes the color similarity of each pixel and its neighbors. Yatziv and Sapiro [19] proposed a colorization method based on the geodesic distance between neighboring pixels. The user input is iteratively propagated to the neighboring pixel with the minimum distance. Gastal and Oliveira [5] considered Levin's colorization method [12] as the application of edge-aware filtering, in which grayscale images are colorized using a domain transform to achieve fast performance. Levin's colorization algorithm was additionally extended by Chen *et al.* [4]. Instead of using neighborhood pixels in geometry space, they utilize the neighborhood pixels in the feature domain. They compute the k-nearest neighbors in the feature space and employ locally linear embedding (LLE) to compute the weight of each neighbor.

Only two previous works exist in the field of anaglyph colorization. Lin *et al.* [13] extended Bando's algorithm by applying a color prior to colorize an anaglyph image. However, this method is likely to fail when dealing with complex scenes. In the notable work of Joulin and Kang [10], an iterative framework was proposed using modified scale-invariant feature transform (SIFT) flow [14], which is called anaglyph SIFT (ASIFT) flow, and diffusion-based colorization. They employ a colorization method similar to that of Levin's [12] but use a different and larger kernel size. However, their method fails when the computed correspondence map is inaccurate. Note that both algorithms [10, 13] highly rely on the performance of correspondence estimation. Furthermore, both works only focus on anaglyph colorization without reconstructing the accurate disparity map.

3. Anaglyph Stereo Matching

In this paper, the proposed stereo matching algorithm is defined as an energy minimization problem in the MAP-MRF framework [3] as follows:

$$E(f) = \sum_{p} D_p(f_p) + \sum_{p} \sum_{q \in N_p} V_{pq}(f_p, f_q) \qquad (1)$$

where f_p and N_p are the label and neighborhood pixels of pixel p, respectively. $D_p(f_p)$ is the data cost which measures how appropriate label f_p is for a given pixel p. $V_{pq}(f_p, f_q)$ is the smoothness cost, which measures how consistent a label f_p is for a given pixel p with its neighbor pixel q having a label f_q. The optimal disparity result can be obtained by minimizing the energy $E(f)$ in Eq. (1) using graph cuts [3].

As referred by Meltzer *et al.* [15], the energy function is more important than the optimization algorithm for obtaining the optimal disparity result. Therefore, the proposed framework designs accurate data costs that are robust for the anaglyph image. Two novel data costs for anaglyph stereo matching are proposed: adaptive data cost using local color prior (D_p^{LC}) and modified census data cost based on reverse intensity distribution (D_p^{RID}). To improve the accuracy, a segmentation-based plane fitting data cost (D_p^{Seg}) [9] is employed additionally. The smoothness cost is modeled using the truncated linear cost. The final data and smoothness costs are defined as follows:

$$D_p(f_p) = D_p^{LC}(f_p) + D_p^{RID}(f_p) + D_p^{Seg}(f_p) \qquad (2)$$

$$V_{pq}(f_p, f_q) = \alpha \min(|f_p - f_q|, V_{max}) \qquad (3)$$

where α is the smoothness weight and V_{max} is the maximum disparity difference. The detail of each cost is described in the following subsections.

3.1. Adaptive Data Cost using Local Color Prior

Conventional adaptive stereo matching [20] forces similar intensity assumption to obtain an accurate disparity map.

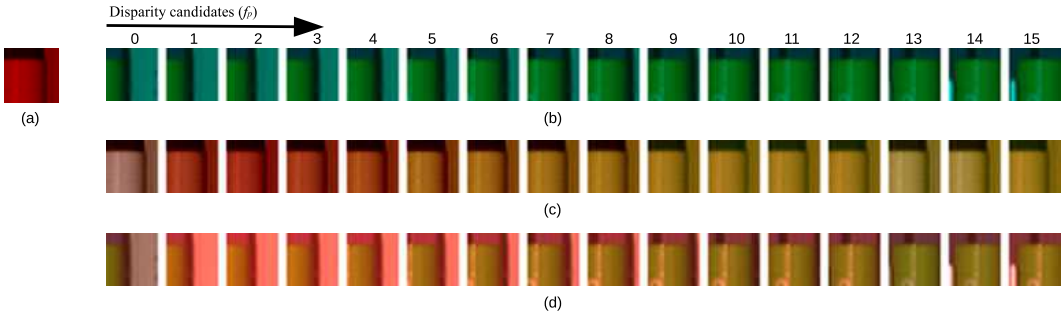


Figure 2: Pseudo color reconstruction of a patch in Tsukuba data (p = (304,97)). (a) Left red patch; (b) Right cyan patches; (c) Left pseudo color patches; (d) Right pseudo color patches.

However, the nature of anaglyph image does not satisfy the assumption. Thus, we introduce a novel local color prior to estimate the missing color channel (*i.e.* to reconstruct pseudo color) so that it can make good use of the assumption. We exploit color transfer method [16] to model the prior. The pseudo color of a patch is reconstructed by locally transferring the known color channel from the corresponding patch. For that reason, each patch has different pseudo color depending on the disparity candidate. The local color prior is the adaptive form of color transfer with weight $w(p,q)$ is defined as follows:

$$w(p,q) = exp(-\frac{\Delta c_{pq}}{\lambda_c} - \frac{\Delta s_{pq}}{\lambda_s}) \tag{4}$$

where Δc_{pq} and Δs_{pq} are the color and spatial differences between pixel p and q, respectively. λ_c and λ_s are the parameters that control how much each difference influences the weight $w(p,q)$. Then, adaptive mean $\hat{\mu}$ and standard deviation $\hat{\sigma}$ are computed as:

$$\hat{\mu}(p) = \frac{\sum_{q \in N_p} w(p,q) I(q)}{\sum_{q \in N_p} w(p,q)} \tag{5}$$

$$\hat{\sigma}(p) = \sqrt{\frac{\sum_{q \in N_p} w(p,q) \| I(q) - \hat{\mu}(p) \|^2}{\sum_{q \in N_p} w(p,q)}} \tag{6}$$

where I is the original intensity and N_p is the local patch centered at p. Finally, the pseudo intensity $\tilde{I}$ is obtained by computing:

$$\tilde{I}_t(q)|_{q \in N_p} = (I_s(q) - \hat{\mu}_s(p)) \frac{\hat{\sigma}_t(p')}{\hat{\sigma}_s(p)} + \hat{\mu}_t(p') \tag{7}$$

where s and t denote the source and target color channels. Given pixel p and label f_p, the corresponding pixel in another view is denoted by p'. For example, if we want to compute the pseudo intensity of left green channel $\tilde{I}_g^L$, we utilize the left red channel I_r^L as the source and right green channel I_g^R as the target patch (s = red, t = green). The pseudo colors are utilized together with the original colors to measure the adaptive data cost. Adaptive data cost using local color prior ($D_p^{LC}(f_p)$) is defined as follows:

$$\begin{aligned} e(q,q') = min\{ & |I_r^L(q) - \tilde{I}_r^R(q')| + \\ & |\tilde{I}_g^L(q) - I_g^R(q')| + |\tilde{I}_b^L(q) - I_b^R(q')|, T\} \end{aligned} \tag{8}$$

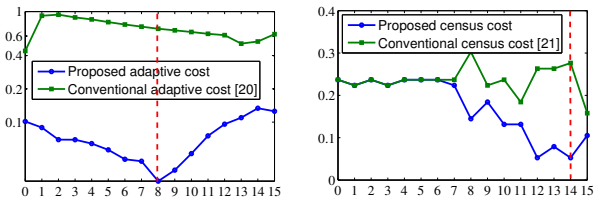
$$D_p^{LC}(f_p) = \frac{\sum_{q \in N_p, q' \in N_{p'}} w(p,q) w(p',q') e(q,q')}{\sum_{q \in N_p, q' \in N_{p'}} w(p,q) w(p',q')} \tag{9}$$

where $\{L, R\}$ and $\{r, g, b\}$ are the set of image positions (left and right) and color channels, consecutively. $e(q,q')$ is the pixel-based matching cost while T is the truncation value of the cost. Figure 2 shows an example of pseudo color reconstruction (local color prior generation) of a patch. It is revealed that the pseudo color of both left and right patches are different depending on the disparity candidate (*e.g.* 16 disparity candidates for Tsukuba data). In summary, we estimate the missing color of a patch with the color information of corresponding patch. Thus, the color similarity is preserved for each channel. Since we apply adaptive color transfer, minimum cost is obtained when the local structure is similar between corresponding patches.

Figure 3 (a) shows the data cost curve comparison of the corresponding patches in Figure 2. It is shown that the proposed adaptive data cost obtains the correct disparity as its ground truth. Figure 4 (c) and (d) display the disparity map comparison with the conventional adaptive stereo matching [20]. The proposed data cost achieves more pleasing result because the local color prior satisfies the intensity similarity assumption.

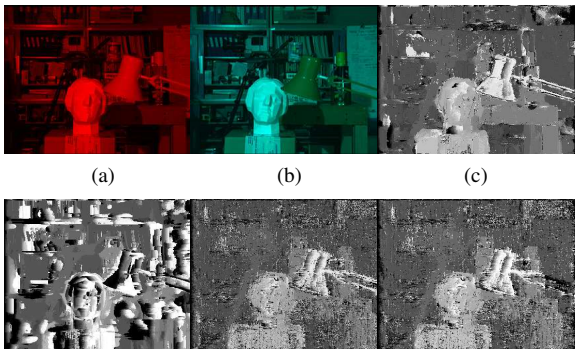
3.2. Reverse Intensity Distribution and Modified Census Data Cost

An anaglyph image is comprised of color channels from different views. Therefore, it is worth searching for a useful factor for computing the correlation across different color channels. For each individual color channel, modified census filtering is performed to capture the relative intensity distribution. Given pixel p and local patch N_p centered at p in single channel image I, the filter output $F(p)$ is computed by counting and comparing the number of brighter



(a) (b)

Figure 3: Data cost curve comparison of a patch in Tsukuba data (red line denotes the ground truth). (a) Adaptive data cost (p = (304,97)); (b) Modified census data cost (p = (336,170)).



(a) (b) (c)

(d) (e) (f)

Figure 4: Disparity maps of each data cost (Tsukuba data). (a) Left red image; (b) Right cyan image; (c) Result of proposed adaptive data cost; (d) Result of conventional adaptive data cost [20]; (e) Result of proposed census data cost; (f) Result of conventional census data cost [21].

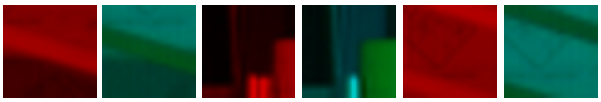
and darker pixels in N_p compared with the intensity of p. The mathematical formulation of the filter output is given by:

$$C(q)|_{q\in N_p} = \begin{cases} 1 & \text{if } I(q) < I(p) \\ -1 & \text{otherwise} \end{cases} \tag{10}$$

$$F(p) = \begin{cases} 1 & \text{if } \sum_q C(q) > 0 \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

where $I(p)$ is the intensity of center pixel p. Consequently, filter output $F(p)$ encodes the intensity distribution at p, thereby indicating whether p belongs to the brighter group of pixels (*i.e.* $F(p) = 1$) or darker group of pixels (*i.e.* $F(p) = 0$).

After performing filtering on three individual color channels, the filtered pixel value between a pair of color channels (*e.g.* red to green and red to blue) is compared. If the values are the same at the given pixel location, two channels have coherent intensity distribution at the pixel. On the contrary, if they are different, they have the reverse relative intensity, which means that the local bright/dark relation between pixels is reversed. In this paper, this is called *reverse intensity distribution*. Figure 5 (a) and (b) show the patch pairs with reverse intensity distribution.



(a) (b) (c) (d) (e) (f)

Figure 5: Corresponding left and right patches (using ground truth disparity). (a),(c),(e) Left red patches; (b),(d),(f) Right cyan patches.

To obtain accurate cross channel correlation value, we introduce a modified census data cost based on the reverse intensity distribution. Given each pair of color channels, two data costs are measured for corresponding patches: reverse data cost ($D_p^R(f_p)$) and non-reverse data cost ($D_p^{NR}(f_p)$). Then, the final modified census data cost ($D_p^{RID}(f_p)$) is obtained by selecting the minimum data cost as follows:

$$D_p^{NR}(f_p) = \sum_{q\in N_p, q'\in N_{p'}} C(q) = C(q') \tag{12}$$

$$D_p^{R}(f_p) = \sum_{q\in N_p, q'\in N_{p'}} C(q) \neq C(q') \tag{13}$$

$$D_p^{RID}(f_p) = min\{D_p^{NR}(f_p)|_{RG}, D_p^{R}(f_p)|_{RG}, D_p^{NR}(f_p)|_{RB}, D_p^{R}(f_p)|_{RB}\} \tag{14}$$

where $\{RG, RB\}$ is the set of cross color channel pairs. p' is the corresponding pixel of pixel p with disparity label f_p. As shown in Figure 3 (b), the proposed census data cost obtains smaller value at the ground truth disparity. Subjective comparison is shown in Figure 4 (e) and (f). Modified census data cost gains better results in the regions that are under reverse intensity distribution.

3.3. Segmentation-based Plane Fitting Data Cost

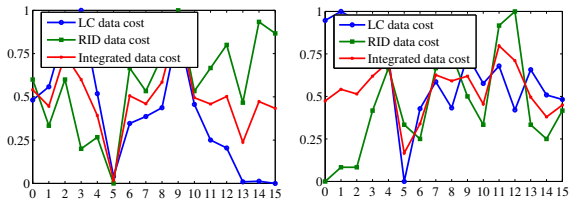
The segmentation-based plane fitting data cost is used as a soft constraint to produce more accurate results [9]. The data cost using segmentation-based plane fitting ($D_p^{Seg}(f_p)$) is formulated as:

$$D_p^{Seg}(f_p) = |a_s x_p + b_s y_p + c_s - f_p| \tag{15}$$

where a_s, b_s, and c_s are the estimated 3D plane coefficients of segment s to which pixel p belongs, and (x_p, y_p) are the coordinates of pixel p in x and y directions. We refer to [9] for details of the segmentation-based plane fitting data cost.

3.4. Data Cost Integration

The integration of two novel data costs is required because each data cost may have cost ambiguity depends on the patch. The adaptive cost meets the ambiguity when there is similar spatial structure with high weight value. On



(a) (b)

Figure 6: Cost curve comparison of patches in (a) Figure 5 (c) and (d); and (b) Figure 5 (e) and (f).

(a) (b) (c)

Figure 7: Intermediate disparity results of Tsukuba image. (a) Result of integrated data cost; (b) Result of 1^{st} iteration; (c) Result of final (5^{th}) iteration.

the contrary, census data cost faces ambiguity when similar relative ordering of patch intensity occurrs. Figure 5 (c) $\sim$ (f) shows two patch pairs example. To mitigate these ambiguities, integration of both data costs is performed by computing the average value. The cost curve of each patch pair is shown in Figure 6. It verifies that the cost ambiguity is removed by integrating both costs together. Finally, the cost is optimized to obtain smooth disparity maps. Figure 7 (a) displays the result of the integrated data cost which can be compared with the results of each data cost in Figure 4 (c) and (e).

At the first iteration, we utilize only the proposed data costs to obtain accurate dense correspondence map from an anaglyph image and then reconstruct the missing colors. From the second iteration, the proposed data costs are cooperated by the conventional adaptive and census data costs. This is due to the possibility of having inaccurately reconstructed color at the first iteration which may lead to inaccurate dense matching if we utilize the conventional data costs only. Figure 7 (b) and (c) illustrate the intermediate disparity results of the first and last iterations.

4. Anaglyph Colorization

Anaglyph colorization is performed using the obtained disparity values of non-occluded regions. First, a left-right consistency check is performed to find the occluded regions in both images. Each pixel from the known color channel is warped to another image using the disparity map. Note that the pixels in occluded regions are not warped because they do not exist in the corresponding image.

To colorize the remaining pixels in occluded regions, we utilize a novel optimization function that improves Levin's colorization algorithm [12]. To improve the accuracy, we propose an additional constraint based on the novel similarity weight kernel for the occluded pixels around image border. For each pixel in occluded regions in border region B, we compute the most similar patch (5×5) in a designated window ($15 \times W$) using template matching where W is the image width. Then, the diffusion energy function is defined as:

$$J(R) = \sum_{p} \left(\begin{array}{c} R(p) - \sum_{q \in N_p} w_c(p,q) R(q) - \\ \sum_{q \in M_p} w_c(p,q) R(q) \delta(p \in B) \end{array} \right)^2 \quad (16)$$

$$w_c(p,q) = exp(-\frac{\Delta c_{pq}}{\lambda_c}) \delta(\Delta c_{pq} < T_c) \quad (17)$$

where $R(p)$ is the value of pixel p in the color channel that will be colorized, $w_c(p,q)$ is the colorization weight between pixel p and q, and T_c is the color threshold. δ is membership function which results in 0 or 1. N_p and M_p are the neighborhood pixels (= 9×9) of pixel p and its similar pixel. To minimize $J(R)$, we use a least-square solver for sparse linear systems.

The conventional weight kernel function could accurately colorize the pixel in occluded regions around the image center. However, it might give inaccurate colorization to the region near the image border that appears only in one image. The additional constraint is well performed when the obtained disparity maps have good quality. These results can be obtained because we consider pixels with similar patches in the energy optimization.

5. Experimental Results

The proposed algorithm is implemented on an Intel i7 4770 @ 3.4GHz with 8GB RAM. We generate anaglyph images by extracting the red channel from the left image and green/blue channels from the right image. Several stereo dataset are used for qualitative as well as quantitative evaluation including Middlebury dataset [17], FhG-HHI 3D video database [1], and stereo frames captured by 3D TV broadcasting in service. Note that we utilize Middlebury dataset especially for quantitative evaluation since it has ground truth of both disparity and color.

In our experiment, the window size used for computing the data costs is 19×19. We set $\lambda_c = 5$, $\lambda_s = 5$, $T = 75$, $T_c = 10$, and $V_{max} = 5$. The number of iterations is 5. The parameters for mean-shift segmentation are fixed as $(C_{Seg}, S_{Seg}, R_{Seg}) = (5, 5, 20)$, where C_{Seg}, S_{Seg}, and R_{Seg} represent the color bandwidth, spatial bandwidth, and the size of the minimum region, respectively. The algorithm is implemented using C++ without optimization although a few computationally complex functions are parallelized on the GPU.

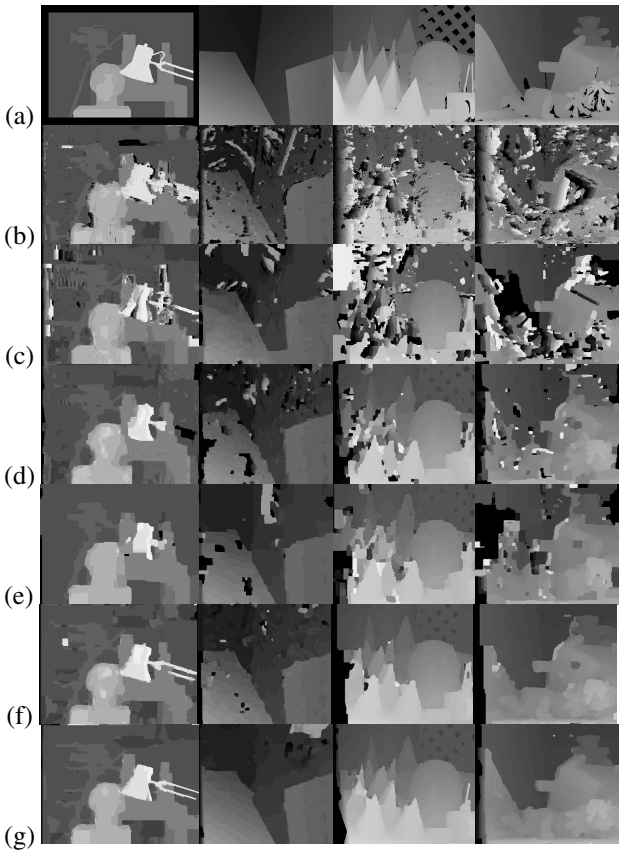


Figure 8: Stereo matching comparison between the proposed method and the conventional approaches. (a) Ground truth; (b) Results of Lin [13]; (c) Results of MI [11]; (d) Results of Census+GC [21]; (e) Results of ASIFT flow [10]; (f) Results of the proposed method (without plane fitting); (g) Results of the proposed method (with plane fitting).

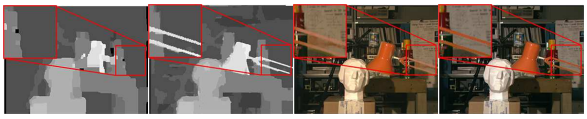
5.1. Anaglyph Stereo Matching Evaluation

To evaluate the accuracy of anaglyph stereo matching, the proposed framework is first compared with the popular data costs that are invariant with illumination changes, such as MI [11] and census data cost [21]. The original optimization method for each data cost is implemented to perform fair comparison. The proposed framework is further compared with Lin's algorithm [13], in which the result is provided by the authors of [13]. In addition, ASIFT flow map [10] is computed since it is the correspondence algorithm used in the state-of-the-art of anaglyph colorization [10]. To prove the strength of the proposed data costs, we also evaluate the proposed method without segmentation-based plane fitting data cost.

Figure 8 shows the qualitative comparison of disparity result. It is evident that the proposed algorithm produces significantly better disparity compared to other approaches. The bad pixel percentage is computed for each image to measure disparity error quantitatively, as summarized in Table 1. The illumination invariant data costs (MI and census) do not work well on the anaglyph image because they depend on similar intensity distribution heavily. Lin's algorithm [13] performs better than the illumination-invariant data costs, but it also fails when the cross color channels have reverse intensity distribution. Similarly, ASIFT flow results in inaccurate correspondence map. To the best of the authors' knowledge, the proposed method is the first anaglyph stereo matching algorithm that produces accurate disparity results. Figure 9 (a) and (b) show the zoomed version of correspondence map comparison between ASIFT flow and the proposed method.

Table 1: Comparison of the bad pixel percentage

Test Images	**Tsukuba**		**Venus**		**Cones**		**Teddy**	
	Left	Right	Left	Right	Left	Right	Left	Right
MI [11]	11.58	N/A	20.35	41.86	47.50	71.60	48.60	65.00
Census + GC	6.52	N/A	15.81	12.91	16.50	16.08	23.12	20.03
Lin's [13]	10.50	N/A	20.44	16.71	34.38	27.50	51.42	39.82
ASIFT Flow [10]	6.69	N/A	12.77	13.00	22.24	22.65	35.38	34.11
Proposed (w/o Plane Fitting)	4.29	N/A	9.15	9.31	6.94	7.83	10.89	15.40
Proposed (w/ Plane Fitting)	**2.55**	N/A	**6.37**	**7.43**	**5.49**	**5.50**	**7.24**	**7.07**



(a) (b) (c) (d)

Figure 9: Comparison of correspondence map and reconstructed color images between Joulin's algorithm [10] and the proposed method with zoomed region. (a) Disparity result of ASIFT flow [10]; (b) Disparity result of the proposed method (with plane fitting); (c) Colorization result of Joulin [10]; (d) Colorization result of the proposed method.

5.2. Anaglyph Colorization Evaluation

To evaluate the proposed colorization algorithm, the peak signal-to-noise-ratio (PSNR) is calculated to determine how similar the reconstructed image is to its ground truth. We compare the performance to the state-of-the-art anaglyph colorization algorithm [10, 13]. In our experiment, Joulin's colorization consists of two different implementations depending on the initial dense correspondence: ASIFT flow (ASIFT-Joulin) and our disparity map (P-Joulin). Table 2 summarizes the PSNR of each method, which shows that the proposed method outperforms others in all cases. A subjective comparison is additionally performed as shown in Figure 10. It shows the difference map (shown in double scale) between each result and its ground truth, which confirms that the proposed method has the smallest difference.

Lin's algorithm [13] and ASIFT-Joulin [10] obtain comparable colorization for stereo images with small disparity values (*e.g.* Tsukuba and Venus). However, they fail on the stereo images with large disparity values (Cones and Teddy). This is due to the cost ambiguity for large dis-

Table 2: PSNR comparison of the colorization results

Test Images	Tsukuba		Venus		Cones		Teddy		Supporters		Javelin1		Javelin2		Javelin3		Book Arrival	
	Left	Right	Left	Right	Left	Right	Left	Right	Left	Right	Left	Right	Left	Right	Left	Right	Left	Right
Lin's [13]	32.35	33.15	27.74	28.62	18.12	21.33	18.49	23.54	26.56	30.43	24.58	29.03	28.10	31.81	27.26	33.93	27.45	32.27
ASIFT-Joulin [10]	30.83	32.88	29.66	31.97	21.52	24.54	21.16	24.59	28.68	30.83	27.61	26.83	29.39	32.10	32.16	35.67	32.56	37.45
P-Joulin [10]	32.47	33.87	30.33	33.20	21.79	24.02	26.69	31.57	29.41	31.56	28.98	27.05	31.21	33.93	**34.50**	38.30	33.09	34.21
Proposed	**33.38**	**35.11**	**31.47**	**33.90**	**25.44**	**27.06**	**31.47**	**33.90**	**30.06**	**32.86**	**30.66**	**33.72**	**31.55**	**34.76**	34.38	**39.15**	**36.20**	**38.64**

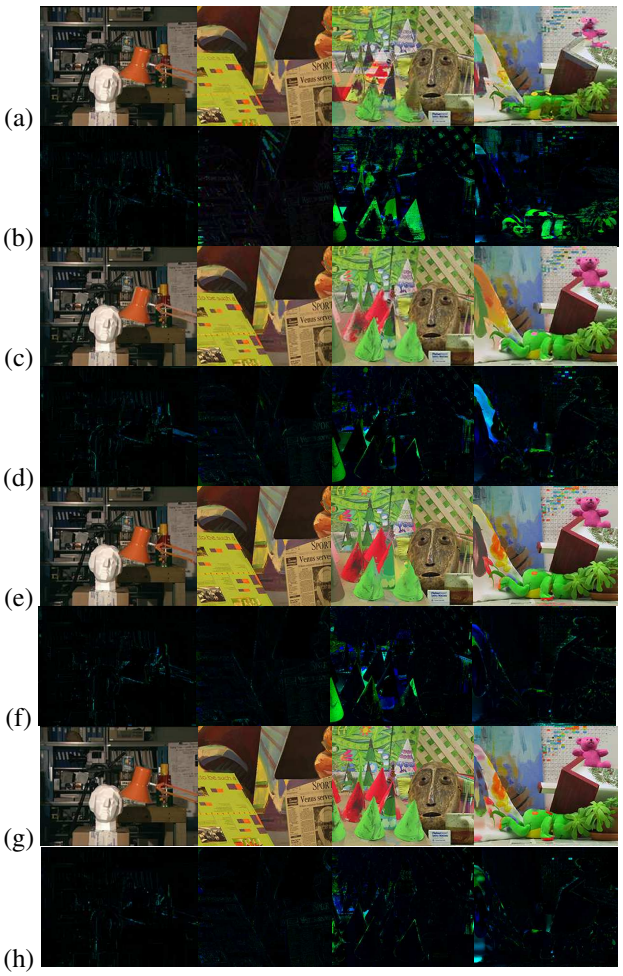


Figure 10: Colorization results comparison between the proposed framework and the conventional methods. (a) Results of Lin [13]; (b) Difference maps of Lin [13]; (c) Results of ASIFT-Joulin [10]; (d) Difference maps of ASIFT-Joulin [10]; (e) Results of P-Joulin [10]; (f) Difference maps of P-Joulin [10]; (g) Results of the proposed method; (h) Difference maps of the proposed method.

parity candidates. Both algorithms also fail on the regions with reverse intensity distribution because they highly rely on their correspondence estimation. On the other hand, the proposed framework and P-Joulin obtain better colorization results in general stereo images. However, P-Joulin fails on colorizing boundary region while the proposed method achieves better PSNR value due to the additional constraint.

Figure 9 (c) and (d) show the zoomed version of colorization comparison. The state-of-the-art ASIFT-Joulin [10] fails to reconstruct the original color because of the inaccurately computed correspondence map. On the contrary, the proposed method performs more robustly than others. Additional results and comparisons on real dataset are shown in Figure 11 and Figure 12. It shows that the proposed method still shows stable and outstanding performance on a variety of real data. This also proves that the proposed local color prior and reverse intensity distribution are practical and effective for real anaglyph applications.

5.3. Limitation and Failure Case

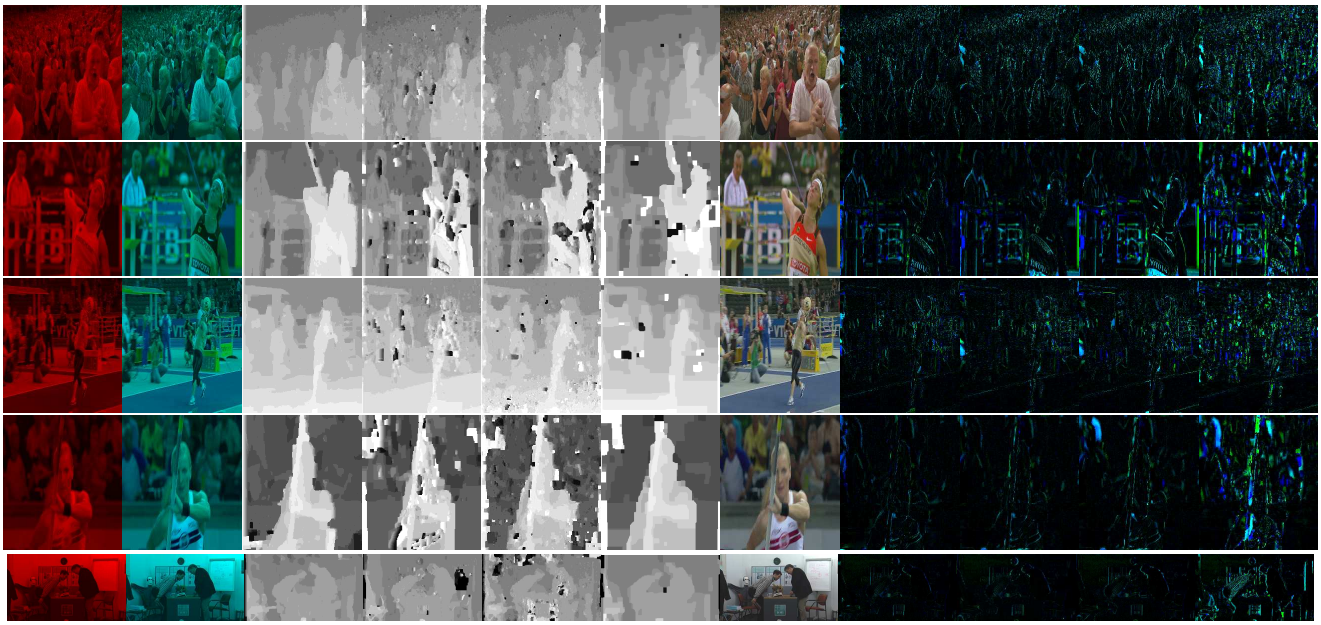
Although accurate depth is obtained by using the proposed data costs, it still cannot achieve the sub-pixel accuracy nor preserve the object edges accurately. As shown in Figure 12, there are some artifact around the object edges with depth discontinuity. Another limitation is that the proposed method is computationally expensive to reconstruct an anaglyph image. We leave some of them as future works to improve the quality of the current work, including efficient implementation of the proposed data costs, edge preserving matching, and colorization optimization.

6. Conclusion

We proposed a joint iterative method to reconstruct the disparity map and missing color information in anaglyph images. First, the local color prior was proposed to compute the pseudo color of missing color channels. Second, the reverse intensity distribution was identified to model the cross color channel correlation. Both were utilized effectively in the cost function design. A novel weight kernel function was proposed to colorize the image boundary region more accurately. It was shown through extensive experiment that both stereo matching and colorization results significantly outperformed the conventional approaches.

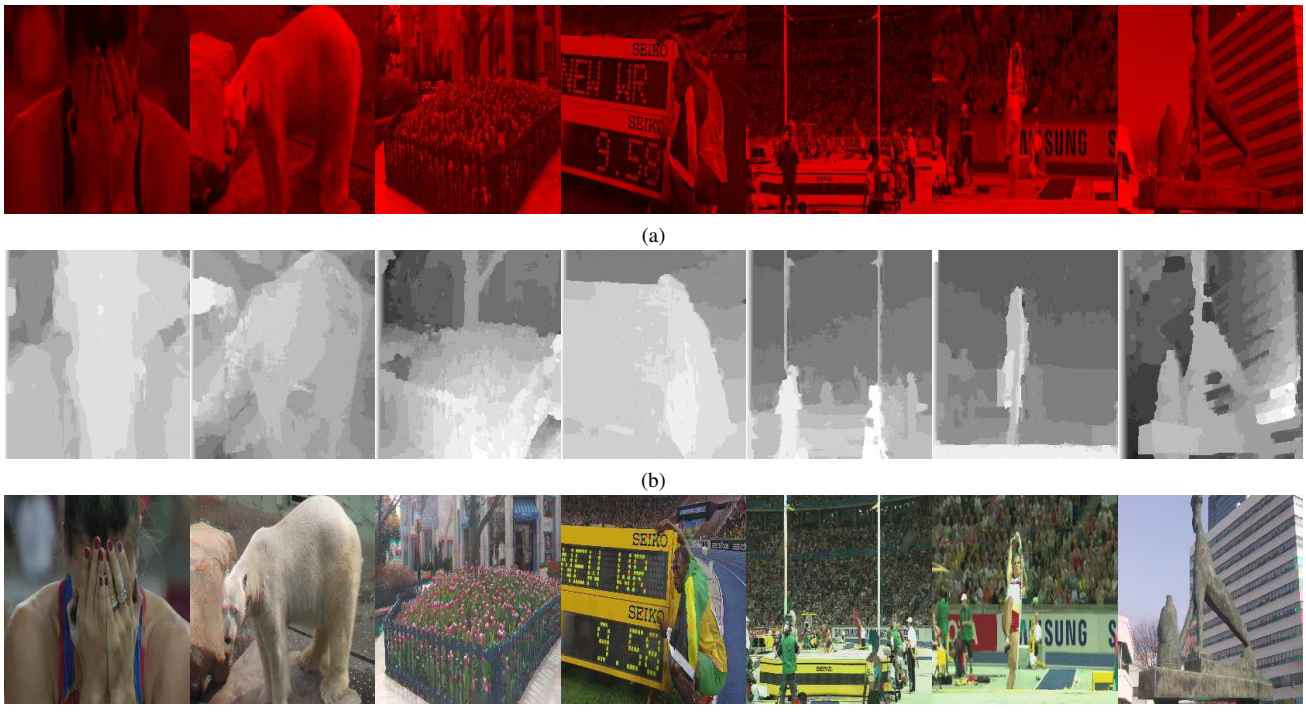
Acknowledgement

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2012R1A1A2009495). This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (No. NRF-2013R1A2A2A01069181).



(a) (b) (c) (d) (e) (f) (g) (h) (i) (j) (k)

Figure 11: Disparity and colorization results comparison of real dataset. (a) Input left images; (b) Input right images; (c) Disparity maps of the proposed method; (d) Disparity maps of MI [11]; (e) Disparity maps of Census+GC [21]; (f) Disparity maps of ASIFT flow [10]; (g) Colorization results of the proposed method; (h) Difference maps of the proposed method; (i) Difference maps of P-Joulin [10]; (j) Difference maps of ASIFT-Joulin [10]; (k) Difference maps of Lin [13]. PSNR values are listed in Table 2.



(a)

(b)

(c)

Figure 12: Additional result of real dataset. (a) Input left images; (b) Disparity maps of the proposed method; (c) Colorization results of the proposed method.

References

[1] Mobile 3DTV content delivery optimization over DVB-H system. http://sp.cs.tut.fi/mobile3dtv/stereo-video/. 5

[2] Y. Bando, B.-Y. Chen, and T. Nishita. Extracting depth and matte using a color-filtered aperture. *ACM Trans. on Graphics*, 27(5):134, Dec. 2008. 2

[3] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 23(11):1222–1239, Nov. 2001. 2

[4] X. Chen, D. Zou, Q. Zhao, and P. Tan. Manifold preserinvg edit propagation. *ACM Trans. on Graphics*, 31(6):132:1–132:7, Nov. 2012. 2

[5] E. S. L. Gastal and M. M. Oliveira. Domain transform for edge-aware image and video processing. *ACM Trans. on Graphics*, 30(4):69:1–69:12, July 2011. 2

[6] Y. S. Heo, K. M. Lee, and S. U. Lee. Robust stereo matching using adaptive normalized cross correlation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 33(4):807–822, Apr. 2011. 2

[7] Y. S. Heo, K. M. Lee, and S. U. Lee. Joint depth map and color consistency estimation for stereo images with different illuminations and cameras. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 35(5):1094–1106, May 2013. 2

[8] H. Hirschmüller and D. Scharstein. Evaluation of stereo matching costs on images with radiometric differences. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 31(9):1582–1599, Sept. 2009. 2

[9] L. Hong and G. Chen. Segment-based stereo matching using graph cuts. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pages I–74–I–81, 2004. 2, 4

[10] A. Joulin and S. B. Kang. Recovering stereo pairs from anaglyphs. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pages 289–296, 2013. 1, 2, 6, 7, 8

[11] J. Kim, V. Kolmogorov, and R. Zabih. Visual correspondence using energy minimization and mutual information. In *Proc. of IEEE International Conference on Computer Vision*, pages 1033–1040, 2003. 2, 6, 8

[12] A. Levin, D. Lischinski, and Y. Weiss. Colorization using optimization. *ACM Trans. on Graphics*, 23(3):689–694, Aug. 2004. 2, 5

[13] H. S. Lin, C. L. Zheng, Y. H. Lin, and M. Ouhyoung. Optimized anaglyph colorization. In *Proc. of SIGGRAPH Asia 2010 Technical Briefs*, 2012. 1, 2, 6, 7, 8

[14] C. Liu, J. Yuen, A. Torralba, J. Sivic, and W. T. Freeman. Sift flow: Dense correspondence across different scenes. In *Proc. of the European Conference on Computer Vision (ECCV): Part III*, pages 28–42, 2008. 1, 2

[15] T. Meltzer, C. Yanover, and Y. Weiss. Globally optimal solutions for energy minimization in stereo vision using reweighted belief propagation. In *Proc. of IEEE International Conference on Computer Vision*, pages 428–435, 2005. 2

[16] E. Reinhard, M. Adhikhmin, B. Gooch, and P. Shirley. Color transfer between images. *IEEE Computer Graphics and Applications*, 21(5):34–41, Sept. 2001. 3

[17] D. Scharstein and R. Szeliski. Middlebury stereo vision page. http://vision.middlebury.edu/stereo/. 5

[18] D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision*, 47(1):7–42, May 2002. 1

[19] L. Yatziv and G. Sapiro. Fast image and video colorization using chrominance blending. *IEEE Trans. on Image Processing*, 15(5):1120–1129, May 2006. 2

[20] K.-J. Yoon and I. S. Kweon. Adaptive support-weight approach for correspondence search. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 28(4):650–656, Apr. 2006. 2, 3, 4

[21] R. Zabih and J. Woodfill. Non-parametric local transforms for computing visual correspondence. In *Proc. of European Conference on Computer Vision*, pages 151–158, 1994. 2, 4, 6, 8