# Robust Stereo Matching Method for Radiometric Distortion between Images

Phuc Nguyen Hong*, Musrrat Ali, Chang Wook Ahn
Computer Engineering Department, Sungkyunkwan University, Suwon, Republic of Korea.

* Corresponding author. Email: hongphuc@skku.edu
Manuscript submitted November 3, 2014; accepted March 12, 2015.
doi: 10.17706/ijcee.2015.v7.874

**Abstract:** Stereo images are unavoidably distorted by radiometric distortion, even though the stereo images are taken in the controlled conditions. However, the majority of the existing stereo matching methods assume that corresponding pixels in a stereo pair have same intensity values, so their performances degrade significantly using stereo images with radiometric variations. In this paper, we propose a novel local stereo matching method that can operate robustly with stereo images captured with different illuminations and exposures between two cameras. We exploit Kemeny and Snell's distance to compute matching cost values and use segmentation-based plane fitting to locally smooth the matching cost values. We conduct experiments using the proposed method and compare it with belief propagation, census transform-based, and adaptive normalized cross correlation stereo matching methods. Experimental results show that our proposed matching method outperforms the test stereo matching methods under radiometric differences.

**Key words:** Stereo matching, plane fitting, radiometric distortion, images.

## 1. Introduction and Related Work

Stereo matching technologies are highly applies in machine vision, robotics and image analysis [1]. Basically stereo matching algorithms can also be categorized into local algorithms and global algorithms**.** The goal of two approaches is to find homologous points in the stereo pair from an image by images principle of the same scene. Local stereo matching methods find a density solution, relying over support windows, whereas global stereo matching methods minimize a certain cost function over the whole stereo pair.

Stereo matching methods can divided into four steps matching cost computation, matching cost aggregation, disparity optimization, and disparity refinement processes. Commonly used matching costs include absolute differences (AD) and squared differences (SD); their sum version can be derived by sum matching values over a support window, including sum of absolute differences (SAD) and sum of squared differences (SSD). AD, SD, SAD, and SSD degrade their performances significantly when operating with radiometrically distorted stereo images. Normalized cross correlation (NCC) is invariant to linear intensity changes between stereo images. However, NCC fails to recover disparities for stereo images with nonlinear intensity transformation.

Some popular cost aggregation methods include shiftable window [2], multiple window [3], Adaptive window [4], and cost volume filtering [5]. Shiftable window [2] method uses a fix window the anchor pixel in all position in the window; the position with the best score is selected. The multiple window method [3] computes the optimal window from a number of windows that has the best cost. Support windows in multiple window method have various shapes, however the size of window is not changed. Cost aggregation methods in [5]-[7] integrate spatial weights to improve smooth in object boundaries. Cost volume filtering [8] exploits guided filtering [9] to locally smooth matching values in a matching cost space **C**.

There are two common used non-parametric local transform in stereo matching which are ranks and census transforms. Rank transform is a non-parametric measure of area-based intensity [10]. Rank transform is known to be susceptible to noise in textureless areas which sets a pixel of the matching window as a center, taking of pixels in some square neighborhood transformation window. The two non-parametric transforms base on the relative order between value between the center pixel and its neighborhood. If the center pixel value less than neighborhood pixel, it returns 0 and 1 otherwise. Thus, the set of pixels in reference image and target image become the set of ordered pairs [2].

$$E(P) = \bigcup_{P' \in N(P)} \left(P', \xi\left(P, P'\right)\right). \tag{1}$$

We let $P$ be a pixel, $I(P)$ is an intensity pixel and $N(P)$ is some square neighborhood of diameter $d$ surrounding $P$. The differ in terms of their exact reliance on $E$. The Rank transformation $R(P)$ which is not intensity, but rather an integer in the range {0, $d^2$-1} is as:

$$R(P) = \left\| \left\{ P' \in N(P) | I(P') < I(P) \right\} \right\|. \tag{2}$$

Census transforms maps the local neighborhood surrounding of pixel $p$ to a bit string representing the thresholds values and compare every pixels. If pixel intensity less than pixel intensity center the bit value is 1, otherwise the bit value returns 0.

$$C(P) = \underset{[i,j] \in D}{\otimes} \xi\left(P, P + [i,j]\right). \tag{3}$$

Note that $C(P)$ present Census transform of one point $P$, where $\otimes$ is the Minkowski sum, $D$ is non-parametric window around. And $\xi$ defined as

$$\xi\left(P, P + [i,j]\right) = \begin{cases} 1, & P > P + [i,j] \\ 0, & otherwise \end{cases} \tag{4}$$

Fig. 1 shows an example of a window image. Set the point of the pixel as the center and take 3 ×3 Census transformation rectangular window. Census converts relative intensity difference to 0 or 1 in 1 dimensional vector form.

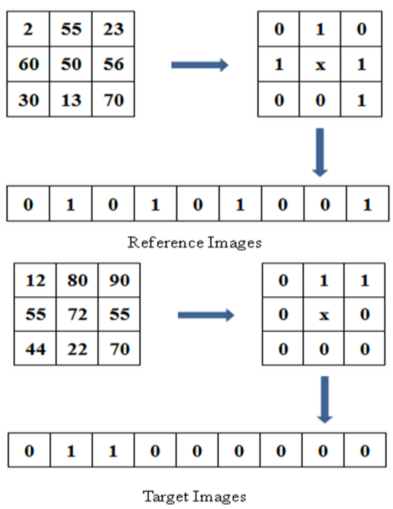



Fig. 1. Census transform example.

With two pixels of census transformed data are compared by using Hamming distance technique which is the number of differences between two vectors. See Fig. 2.

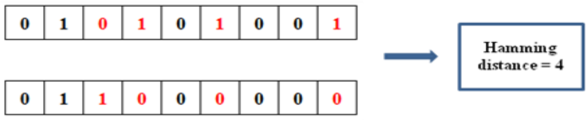



Fig. 2. Compute of the Hanming distance.

The benefits of using Rank transformation in stereo matching can reduce effects of various gains and bias camera, insensible to image noise and brightness differences, and improve in robustness to outliers near depth discontinuities. However, it losses of information related with the pixel, and Rank transforms cannot distinguish between rotations and reflections since non-parametric costs rely on the relative ordering of pixel values, they are invariant under all radiometric changes that preserve this order and it is weak with respect to local radiometric variations. Census is very robust to illumination distortion and commonly use in outdoor stereo matching. The Census problem is based on the order between pixel intensities does not vary after radiometric variations. Rank and Census transform base on the principle that the order between pixel intensities does not vary after radiometric variations.

Another commonly used in stereo matching is called Belief Propagation (BP) [11]. BP produces accurate disparity map with normal condition between stereo images. However, it degrades significantly when stereo images are distorted by illumination because BP is based on AD matching cost which operates well with stereo images of the same lighting conditions.

Adaptive Normalized Cross-Correlation (ANCC) [12] is improved of Normalized Cross-Correlation. The advantage is robustness to lighting geometry, and illuminant color between stereo images. ANCC exploits the log-chromaticity color model. The weak of ANCC fails in case of serve illumination changes, non-Lambertian reflectance object. Another drawback of ANCC is that it only operates with color stereo images.

In this paper, we present a new local approach to the correspondence problem, based on combined between Kemeny and Snell's distance and segmentation-based plane fitting (KSP). By using Kemeny and Snell's distance, the proposed stereo method can tolerate monotonic intensity transformation between stereo images. We also apply plane fitting over matching values in the matching cost space **C**, determined by image segmented regions. We exploit an assumption that pixels having similar intensities are likely on the same object structure, and thus have the similar disparities.

This paper is organized as follows: we present our approach in Section 2. After that, in Section 3, we present experimental results of our method and compare it with BP, Census-based and ANCC-based stereo methods. Finally, we provide conclusions in Section 4.

## 2. Propose Method

In this paper, an alternative approach based on a non-parametric transform is proposed. The flow chart of KSP algorithm is shown in Fig. 3. Taking into account radiometric changes, both reference image and target image first are pre-processed by Kemeny and Snell's in matching cost computation step. Then using plane fitting based on segmentation, after that winner-takes-all strategy is obtain density disparity maps. The KSP algorithm is illustrated in Fig. 3.

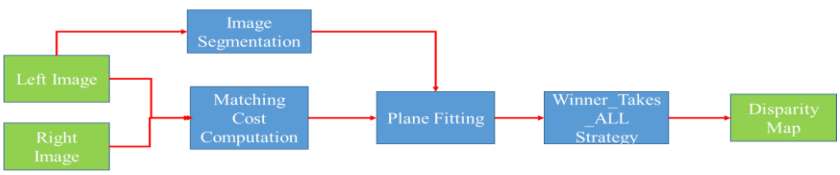



Fig. 3. The flow chart of KSP algorithm.

### 2.1. Image Segmentation

In this paper, we use the mean-shift algorithm [13] to segment the reference image in terms of color. The result of the mean-shift algorithm is accurate enough at acceptable speeds to achieve our goal in the proposed method. Other segmentation methods, however, can be used as long as they generate regions of sufficiently high quality. Fig. 4 shows an example of mean-shift algorithm using the Baby image.

Fig. 4. Mean shift example.

### 2.2. Matching Cost Computation

In this subsection, we present Kemeny and Snell's distance ($d_{KS}$) that is used to compute a matching cost space. ($d_{KS}$) can measure the correlation between two image windows of a same size, and it can tolerate monotonically nonlinear intensity changes between stereo images. The algorithm compares the relative ranking of each pair of pixels extracted from an image window with its relative ranking in the other. If two pixels have a same intensity, ($d_{KS}$) return ½; If the two pixels have different intensities, ($d_{KS}$) return 1 or 0 at all possible ordered pairs of pixel locations. We can follow from

$$b(v_1, v_2) = \begin{cases} 1, & v_1 > v_2 \\ 1/2 & v_1 = v_2 \\ 0, & v_1 < v_2 \end{cases} \tag{5}$$

where $a^I_{p_1,p_2} = b(I(p_1), I(p_2))$ and $A^I = (a^I_{p_1,p_2})$ is constructed for each image

$$d_{KS}(I_1, I_2) = \sum_{p_1 \in P} \sum_{p_2 \in P} \left| a^{I_1}_{p_1,p_2} - a^{I_2}_{p_1,p_2} \right| \tag{6}$$

Kemeny and Snell's distance can be direct computation for any image of reasonable size, contribute at most one to the summation for each ordered pair of pixel locations ($p_1$, $p_2$), and presence of a large number of pixels of equal value image to reduce $d_{KS}$.

### 2.3. Plane Fitting

Once the matching cost image space is computed, plane fitting based on segmentation is applied to locally smooth the values at pixel $p=(x, y)$ in the space. Let $z$ is the matching cost value, so a tuple $(x, y, z)$ represent three components. The aim of using plane fitting is to smooth values of $z$ over a segmented region. Given a set of samples $\{(x_i, y_i, z_i)\}_{i-1}^{m}$, determine $A$, $B$, and $C$ so that the plane $z = Ax + By + C$ best fits the samples in the sense that the sum of the squared errors between $z_i$ and the plane values $Ax_i + By_i + C$ is minimized.

Define $E(A,B,C) = 2\sum_{i=1}^{m}\left[(Ax_i + By_i + C) - z_i\right]^2$. This function is nonnegative and its graph is a hyper-parabolic whose vertex occurs when the gradient satisfies $\nabla E = (0,0,0)$. This leads to a system of three linear equations in $A$, $B$, and $C$ which can be easily solved. Precisely,

$$(0,0,0) = \nabla E = 2\sum_{i=1}^{m}\left[\left(Ax_i + By_i + C\right) - z_i\right]\left(x_i, y_i, 1\right)$$

and so

$$\begin{bmatrix} \sum_{i=1}^{m} x_i^2 & \sum_{i=1}^{m} x_i y_i & \sum_{i=1}^{m} x_i \\ \sum_{i=1}^{m} x_i y_i & \sum_{i=1}^{m} y_i^2 & \sum_{i=1}^{m} y_i \\ \sum_{i=1}^{m} x_i & \sum_{i=1}^{m} y_i & \sum_{i=1}^{m} 1 \end{bmatrix} \begin{bmatrix} A \\ B \\ C \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{m} x_i z_i \\ \sum_{i=1}^{m} y_i z_i \\ \sum_{i=1}^{m} z_i \end{bmatrix} \tag{7}$$

The solution provides the least squares solution $z = Ax + By + C$.

## 2.4. Winner-Takes-All Strategy

A winner–take-all strategy is a local method for computation of a disparity map which is the simplest way to obtain a dense disparity map.

$$D(p) = \arg\min_{d}\left(C_d(p)\right) \tag{8}$$

where $C_d(p)$ is the matching cost of the pixel *p* in the left image (reference image), *d* is the disparity hypothesis, and *D* is the disparity map.

# 3. Experiment

We use the Middlebury datasets to measure the performance of the matching costs including ANCC, BP, Proposed and Census. We use the available source code of BP and ANCC at [14], [15]. For fair comparison, all of the testing data costs using the default values of parameters described in original papers. We use the nine Middlebury sub-datasets of rectified stereo images Aloe, Cloth4, Flowerpots, Lampshade1, Lampshade2, Midd1, Midd2, Monopoly, and Plastic We evaluate the effects of illumination or/and exposures changes by fixed the left image and changed among the three remaining conditions: I2E1_I2E0, I2E1_I3E1, and I2E1_I3E2.

Fig. 5 shows the results of the test data with I2E1_I2E0 illuminations or/and exposures. Fig. 5a), Fig. 5b) show left and right images are with three illuminations and exposures. Fig. 5c)-Fig. 5e) show the disparity maps of ANCC, BP, and Census. Fig. 5f) shows disparity maps of our proposed method. Fig. 5g), describes show the ground truth of thee left images. As we see from the Fig. 7, the result of KPS is better than other algorithms. Next to, Census gives better result than BP and ANCC. The BP is the worst result in our test.

Fig. 6 shows the results of the test data with I2E1_I3E1 illuminations or/and exposures. Fig. 6a) and Fig. 6b) show left and right images are with three illuminations and exposures. Fig. 6c), Fig. 6d), and Fig. 6e) show the disparity maps of ANCC, BP, and Census.

Fig. 6f) shows disparity maps of our proposed method. Fig. 6g), describes show the ground truth of thee left images. As we see from the Fig. 6, the result of KPS is better than other algorithms. Next to, Census gives better result than BP and ANCC. The BP is the worst result in our test.

Fig. 7 shows the results of the test data with I2E1_I3E2 illuminations or/and exposures. Fig. 7a), Fig. 7b) show left and right images are with three illuminations and exposures. Fig. 7c), Fig. 7d), and Fig. 7e) show the disparity maps of ANCC, BP, and Census. Fig. 7f) shows disparity maps of our proposed method. Fig. 7g), describes show the ground truth of thee left images. As we see from the Fig. 7, the result of KPS is better than other algorithms. Next to, Census gives better result than BP and ANCC. The BP is the worst result in our test.

To evaluate the overall performances of the test matching costs, we compute the error in unconcluded area for each sub-datasets including ANCC, Census and proposed over three different illuminations and exposures. As we saw the Fig. 3 result, BP result is the worst so that we do not include BP in our computation. Fig. 8 shows quantitative comparisons of the matching costs over each sub-dataset.

Table 1 shows the average error, to compare with other methods, KSP error is the smallest percentage, then the following is Census, ANCC and the largest is KSP.

To evaluate the computational cost of the proposed stereo method, we used the Aloe stereo images with a resolution of 427×370 and a disparity range of 80. The experimental PC platform had a configuration including of an Intel Core2 Duo 3.00 GHz CPU and 2.00 GB of memory. The proposed method took about 42 seconds.

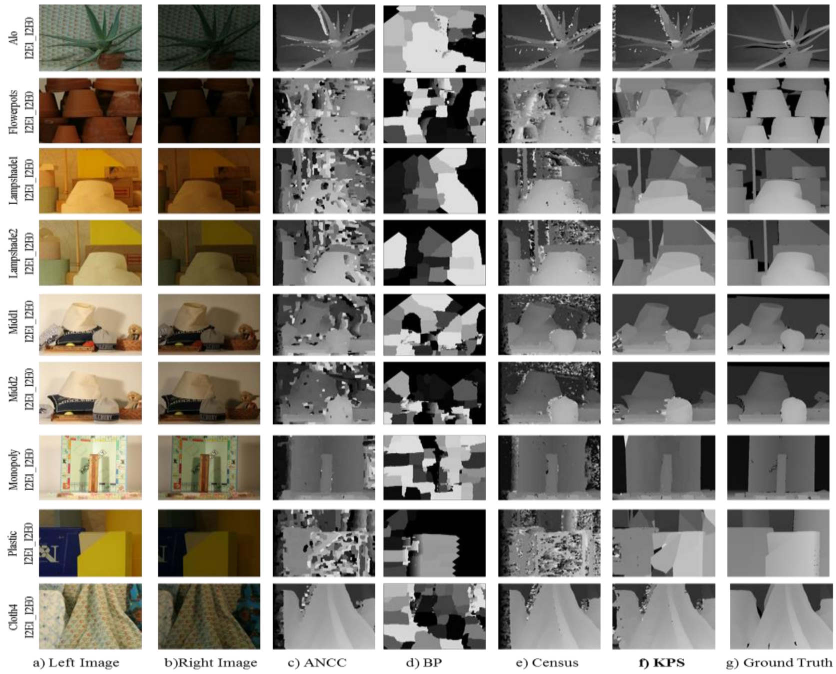



Fig. 5. Result of test data with I2E1_I2E0. a) Left images are with I2E1. b) Right image with I2E0. c) Disparity maps of ANCC. d) Disparity maps of Census. f) Disparity maps of KPS. j) Ground truth.

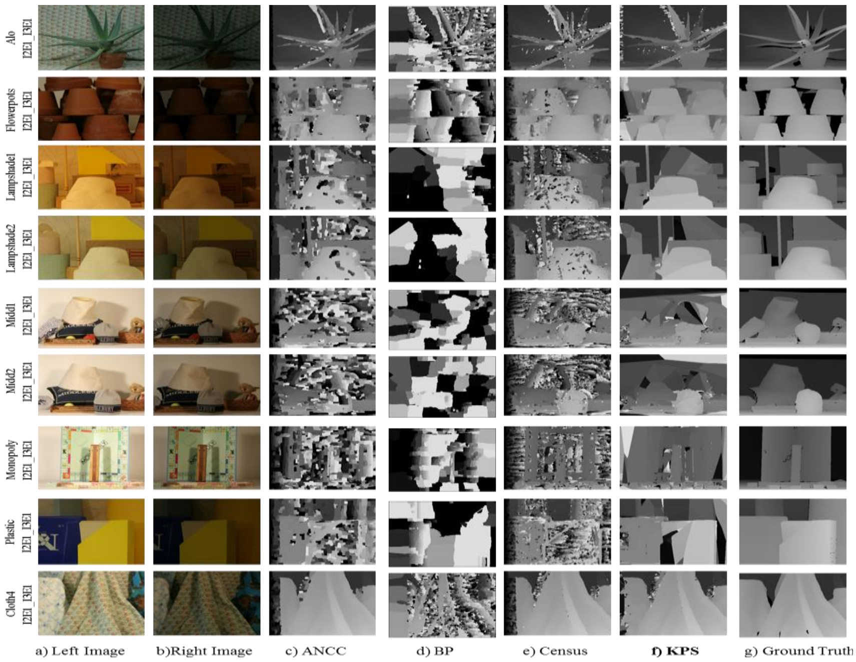



Fig. 6. Result of test data with I2E1_I3E1. a) Left images are with I2E1. b) Right image with I3E1. c) Disparity maps of ANCC.

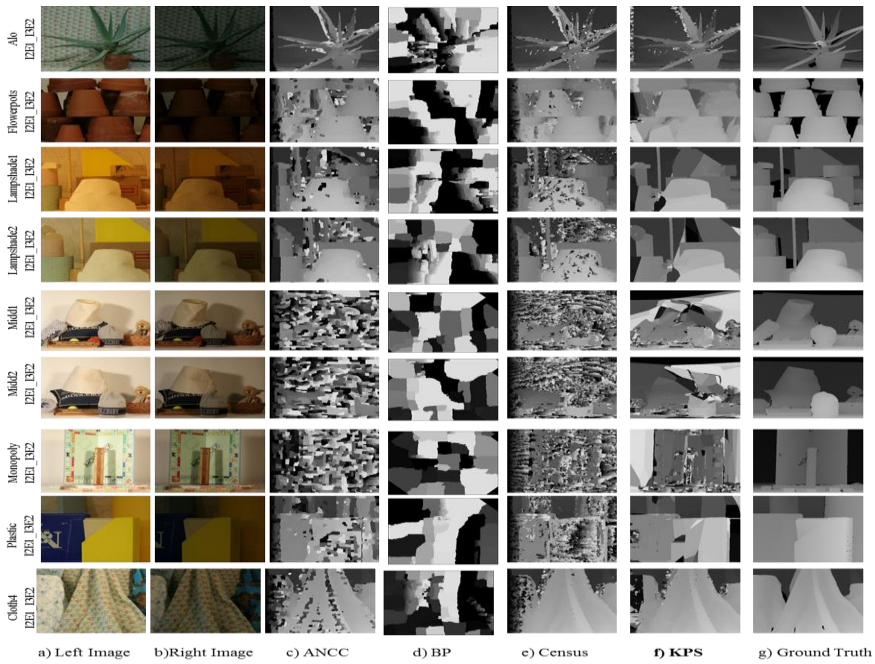
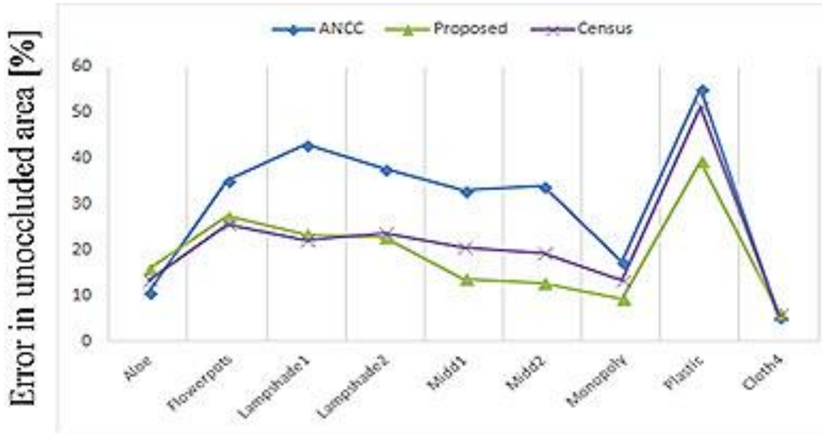
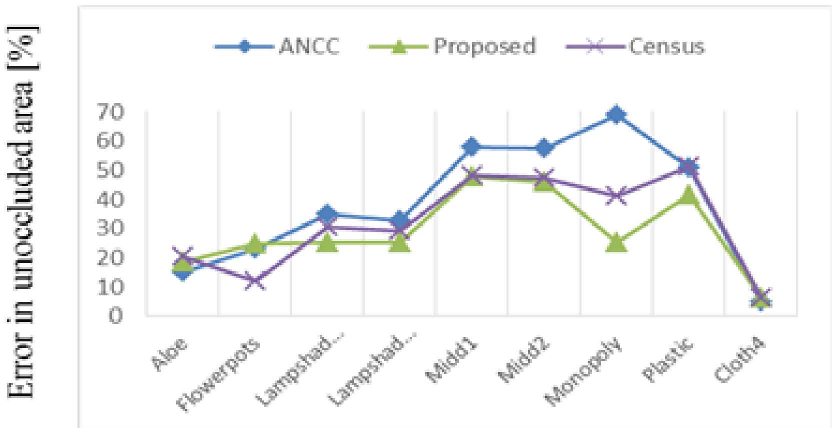


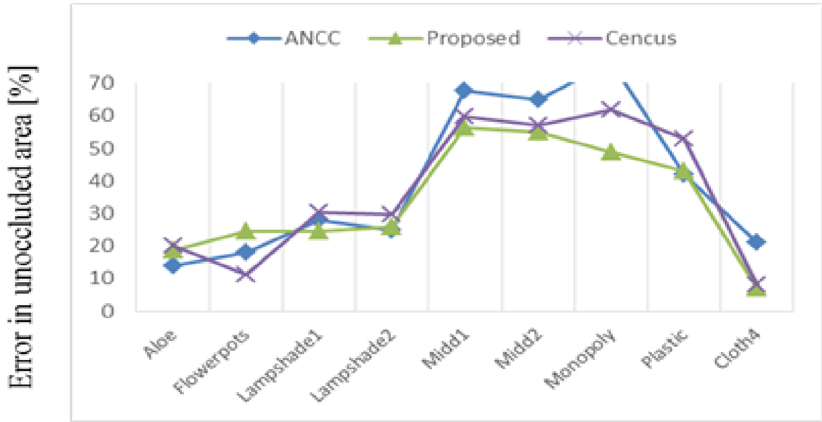

Fig. 7. Result of test data with I2E1_I3E2. a) Left images are with I2E1. b) Right image with I3E2. c) Disparity maps of ANCC. d) Disparity maps of Census. f) Disparity maps of KPS. j) Ground truth.




I2E1_I2E0




I2E1_I3E1




I2E1_I3E2

Fig. 8. Quantitative comparisons of the matching costs over eachsubdataset (stereo pairs from Middlebury dataset).

Table 1. Error of the Test Algorithms in the Middlebury Dataset

| Average Error % | ANCC | BP | KSP | Census |
|---|---|---|---|---|
| I2E1_I2E0 | 30.236 | 74.859 | 19.050 | 21.711 |
| I2E1_I3E1 | 38.413 | 72.058 | 28.992 | 31.779 |
| I2E1_I3E2 | 39.634 | 74.743 | 33.733 | 36.740 |

## 4. Conclusion

In this paper, we propose a novel stereo matching method that can performs robustly with stereo image captured under different illumination and exposure conditions. We conducted experiments using Middlebury dataset to investigate the proposed method, as well as compare it with Census transform, belief propagation, and adaptive normalized cross correlation. Experimental results show that our proposed stereo matching method is superior to the test methods with stereo images captured under varying illuminations and exposures.

## 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 (2014R1A1A1008177).

## References

[1] Nalpantidis, L., Sirakoulis, G. C., & Gasteratos, A. (2007). Review of stereo matching algorithms for 3D vision. *Proceedings of 16th International Symposium on Measurement and Control in Robotics ISMCR*.

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

[3] Hirschmuller, H., Innocent, P., & Garibaldi, J. (2002). Real-time correlation-based stereo vision with reduced border errors. *International Journal of Computer Vision*, *47*, 229–246.

[4] Yoon, K.-J., & Kweon, I.-S., (2005). Locally adaptive support-weight approach for visual correspondence search. *Proceedings of Conference on Computer Vision and Pattern Recognition* (pp. 924–931).

[5] Phuc, N. H., & Ahn, C. W. (2014). Stereo matching using fusion of spatial weight variable window and adaptive support weight. *International Journal Computer and Electrical Engineering, 6(3)*.

[6] Phuc, N. H., Dinh, T. T., & Vinh, D. Q. (Dec. 2012). Enhancement of local methods by using spatial support weight. *Proceedings of IEEE International Symposium on of Signal Processing and Information Technology* (pp. 12-15).

[7] Phuc, N. H., Dinh, T. T., & Vinh, D. Q. (2012). Local stereo matching by joining shiftable window and non-parametric transform. *Proceedings of Multi-Disciplinary International Workshop on Artificial Intelligence*.

[8] Hosni, A., Rhemann, C., Bleyer, M., Rother, C., & Gelautz, M. (2013). Fast cost-volume filtering for visual correspondence and beyond. *IEEE Trans. Pattern Anal. Mach. Intell. (PAMI)*, *35*, 504-511.

[9] He, K., Sun, J. & Tang, X. (2010). Guided image filtering. Presented at European Conference on Computer Vision.

[10] Zabih, R., & Woodfill, J. (May 1994). Non-parametric local transforms for computing visual correspondence. Presented at Third European Conference on Computer Vision. Stockholm, Sweden.

[11] Felzenszwalb, P. F., & Huttenlocher, D. P. (October 2006). Efficient belief propagation for early vision. *International Journal of Computer Vision*, *70(1)*.

[12] Heo, Y. S., Lee, K. M., & Lee, S. U. (2011). Robust stereo matching using adaptive normalized cross-correlation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *33*, 807-822.

[13] Comaniciu, D., & Meer, P. (2002). Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

[14] Efficient Belief Propagation for Early Vision. from http://cs.brown.edu/~pff/bp/

[15] Softwares. from http://cv.snu.ac.kr/softwares


**Phuc Nguyen Hong** received the BS degree in science service from Auckland University of Technology (AUT), New Zealand in 2013. She is currently working toward the master degree at the Computer Engineering Department, Sungkyunkwan University, Suwon, Republic of Korea.

Her research interests include genetic algorithms and computer vision.

**Musrrat Ali** has been the member of the Sungkyunkwan Evolutionary Algorithms Laboratory (SEAL) in the Department of Computer Science and Engineering at Sungkyunkwan University, Suwon, Republic of Korea since March 2012. He received his MSc and PhD degrees from Indian Institute of Technology (IIT) Roorkee, India in 2005, and 2011 respectively. He was a visiting scholar at the University Paris-Est Créteil (UPEC) in 2010. He has published more than 32 papers in journals and conference proceedings. His research interests include evolutionary computation, multi-objective optimization, digital watermarking and image processing.

**Chang Wook Ahn** received his PhD degree at the Department of Information and Communications, Kwang-Ju Institute of Science and Technology, Kwang-ju, Republic of Korea.

From 2007 to 2008, he was a research professor in Gwangju Institute of Science & Technology (GIST). He is currently an assistant professor in Sungkyunkwan University (SKKU), Republic of Korea.

His research interests include genetic algorithms, multiobjective optimization, neural networks, and the applications of evolutionary techniques to wireless networks.