Adaptive Cross Image Filters for Underwater Image Enhancement

Huimin Lu[1,2*], Yujie Li[1], Shiyuan Yang[1] and Seiichi Serikawa[1*]

Abstract

This paper describes a novel method to enhance images by adaptive cross image filters. Scattering and color change are two major problems of distortion for imaging, especially in underwater environment. Scattering is caused by large suspended particles, like fog or turbid water. Color change corresponds to the varying degrees of attenuation encountered by light traveling in the water with different wavelengths, rendering ambient underwater environments dominated by a bluish tone. Our key contribution is to propose a self-referenced image enhancement algorithm, to compensate the attenuation discrepancy along the propagation path, and to take the influence of the possible presence of an artificial lighting source into consideration. A simple colorization method is also utilized for restoring color balance. We demonstrated that the enhanced images are characterized by reduced noised level, better exposedness of the dark regions, improved global contrast while the finest details and edges are enhance significantly. In addition, our enhancement method is comparable to higher quality than the state of the art methods.

Keywords: de-scattering, colorization, image enhancement, image restoration

1. Introduction

As we all known, underwater vision is an important issue in ocean engineering. Different from the common images, underwater image suffers from poor visibility due to the medium scattering and light distortion. Because of the challenging environmental conditions and the differential light dissemination, most of traditional computer vision methods cannot be applied directly in underwater images. However, it is important to notice that the underwater images play a pivotal role in scientific and military missions such as mine counter measure, inspection of underwater things and assessing biological environments [1]. So, how to improve the visibly of underwater image or video is an important issue in recent years.

Capturing images underwater is difficult, mostly due to attenuation caused by light that is reflected from a surface and is deflected and scattered by particles, and absorption substantially reduces the light energy. The random attenuation of the light is mainly cause of the haze appearance while the fraction of the light scattered back from the water along the sight considerable degrades the scene contrast. In particular, the objects at a distance of more than 10 meters are almost indistinguishable while the colors are faded due to the characteristic wavelengths are cut according to the water depth [2].

There have been many techniques to restore and enhance the underwater images. Y.Y. Schechner et al. exploited the polarization filter method to compensate for visibility degradation [3], using image fusion method in turbid medium for reconstruct a clear image [4], and combining point spread function and a modulation transfer function to reduce the blurring effect [5]. Although the aforementioned approaches can enhance the image contrast, these methods have demonstrated several drawbacks that reduce their practical applicability. First, the equipment of imaging is difficult in practice (e.g. range-gated laser imaging system). Second, multiple input images are required (e.g. two illumination images [4], white balanced image and color corrected image [2]).

In this paper, we introduce a novel approach that is able to enhance underwater images based on single image, as well as colorization. We propose a new cross image filter instead of the matting Laplacian [6] to solve the alpha mattes more efficiently. In short summary, our technical contributions are in threefold: first, the proposed neighbor guided filter can perform as an edge-preserving smoothing operator like the popular bilateral filter, but has better behavior near the edges. Second, the neighbor guided filter has a fast and non-approximate linear-time algorithm, whose computational complexity is independent of the filtering kernel size. Third, automatically colorization can perform well by white balancing.

*Corresponding Author: Huimin Lu
(E-mail: luhuimin@boss.ecs.kyutech.ac.jp).
[1] Department of Electrical Engineering and Electronics
Kyushu Institute of Technology, Japan

2. Underwater imaging Model

In the optical model, the acquired image can be modeled as being composed of two components. One is the direct transmission of light from the object, and the other is the transmission due to scattering by the particles of the medium (e.g. air light). Mathematically, it can be written as

$$I(x) = J(X)t(X) + (1 - t(X))A \quad (1)$$

where I is the achieved image. J is the scene radiance or haze-free image, t is the transmission along the cone of vision, and t(x) = exp(-βd(x)), β is the attenuation coefficient of the medium, d(x) is the distance between the camera and the object, A is the veiling color constant and x = (x, y) is a pixel. The optical model assumes linear correlation between the reflected light and the distance between the object and observer.

The light propagation model is slightly different underwater environment. In the underwater model, absorption plays an important role in image degrading. Furthermore, unlike scattering, the absorption coefficient is different for each color channel, being the highest for red and lowest for blue in seawater. These leads to achieve the following simplified hazy image formation model:

$$I(X) = J(X)e^{-(\beta_s + \beta_a)d(X)} + (1 - e^{-\beta_s d(X)})A \quad (2)$$

where βs is the scattering coefficient and βa is the absorption coefficient of light. The effects of haze are highly correlated with the range of the underwater scene. In this paper, we simplify the situation as at a certain water depth, the transmission t is defined only by the distance between camera and scene.

According to some researches, we found that the red color channel is attenuated at a much higher rate than the green or blue channel. We further assume that the transmission in the water is constant. We denote the patch's transmission as $\tilde{t}(X)$. Take the maximum intensity of the red color channel to compare with the maximum intensity of the green and blue color channels. We define the dark channel Jdark(x) for the underwater image J(x) as,

$$J_{dark}(x) = \min_{c \in \{r,g,b\}} (\min_{y \in \Omega(x)} J_c(y)) \quad (3)$$

where Jc(x) refers to a pixel x in color channel c ∈{r,g,b} in the observed image, and Ω refers to a patch in the image. The dark channel is mainly caused by three factors, shadows, colorful objects or surfaces and dark objects or surfaces.

Here, take the min operation in the local patch on the haze imaging function (1), we assume the transmission as,

$$\min(I_c(y)) = \tilde{t}_c(x) \min_{y \in \Omega(x)} (J_c(y)) + (1 - \tilde{t}(x))A_c \quad (4)$$

Since Ac is the homogeneous background light and the above equation and the above equation perform one more min operation among all three color channels as follows:

$$\min_c \min_{y \in \Omega(x)} (\frac{I_c(y)}{A_c}) = \tilde{t}_c(x) \min_c (\min_{y \in \Omega(x)} (\frac{J_c(y)}{A_c})) + (1 - \tilde{t}(x)) \quad (5)$$

According to the dark channel prior, the first term in the right equation of (5) is very close to zero. Therefore, Equation (5) can be written as,

$$\tilde{t}_c(x) = 1 - \min_c (\min_{y \in \Omega(x)} (\frac{I_c(y)}{A_c})) \quad (6)$$

Among all color channels, red channel possesses the lowest residual value. This is mean that, $\tilde{t}_c(x)$, c∈ {r, g, b}. The background Ac is usually assumed to be the pixel intensity with the highest brightness value in an image. However, in practice, this simple assumption often renders erroneous results due to the presence of self-luminous organisms. So, in this paper, we compute the brightest pixel value among all local min corresponds to the background light Ac as follows:

$$A_c = \max_{x \in I} \min_{y \in \Omega(x)} I_c(y) \quad (7)$$

Figure 1 is the estimated transmission depth map from an underwater image using the patch size 15×15. It is roughly good but the edges are not clearly. In the next subsection, we show the transmission depth map after applying guided alpha matting to remove mosaic distortion.

(a)

(b)

Figure 1. Visible mosaic artifacts are observed due to the block-based operation of dark channel prior, (a) Input underwater image (image size 276×208) , (b) Transmission depth map by 15×15 patch

3. Adaptive Cross Image Filter

He et al. [7] proposed guided image filter (GIF) to overcome the gradient reversal artifacts occurring. The filtering process of GIF is firstly done under the guidance of the image G that can be another or the input image I itself. It is similar to the Joint Bilateral Filter [8] which is used to process the no-flash image I by using the flash image G. Let I_p and G_p be the intensity value at pixel p of the input and guided image, w_k be the kernel window centered at pixel k, to be consistent with bilateral filter. GIF is then formulated by,

$$GIF(I)_p = \frac{1}{\sum_{q \in w_k} W_{GIF_{pq}}(G)} \sum_{q \in w_k} W_{GIF_{pq}}(G) I_q \tag{8}$$

where the kernel weights function $W_{GIF_{pq}}(G)$ can be written by

$$W_{GIF_{pq}}(G) = \frac{1}{|w|^2} \sum_{k:(p,q) \in w_k} \left(1 + \frac{(G_p - \mu_k)(G_q - \mu_k)}{\sigma_k^2 + \varepsilon} \right) \tag{9}$$

where μ_k and σ_k^2 are the mean and variance of guided image G in local window w_k, $|w|$ is the number of pixels in this window. When both G_p and G_q are concurrently on the same side of an edge, the weight assigned to pixel q is large. When G_p and G_q are on different sides, a small weight will be assigned to pixel q. The GIF also can be shortening as follows:

$$GIF(I)_p = \sum_{q \in w_k} W_{GIF_{pq}}(G) I_q \tag{10}$$

The degree of smoothing GIF is adjusted by parameter ε. The larger the value of ε is, the smoother the filtered image will be.

With the transmission depth map, we can recover the scene radiance according to Equation (1). We restrict the transmission $t(\boldsymbol{x})$ to a lower bound t_0, which means that a small certain amount of haze are preserved in very dense haze regions. The final scene radiance $J(\boldsymbol{x})$ is written as,

$$J(X) = \frac{I_c(X) - A_c}{\max(t(X), t_0)} + A_c \tag{11}$$

Typically, t_0=0.1. Figure 2 is our recovered images.

(a)

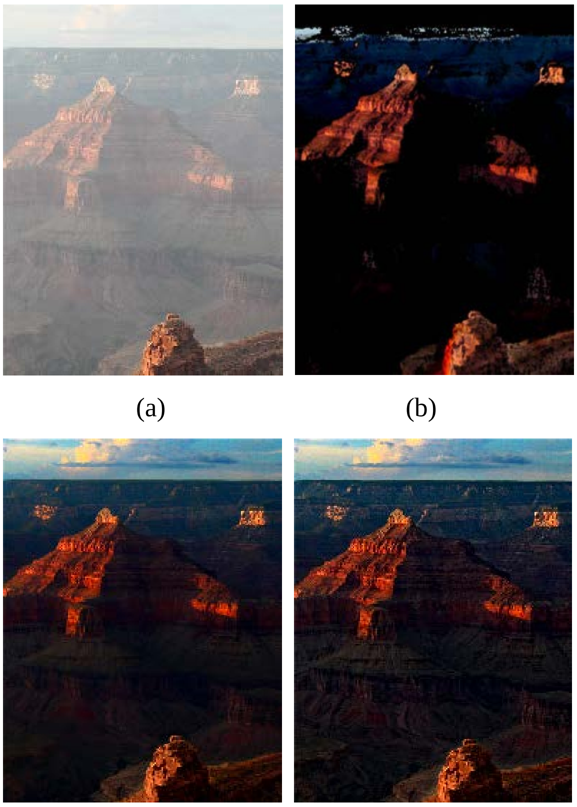
(b)

Figure 2. Results of our proposed method for dehazing. (a) Transmission depth map obtain after guided alpha matting, (b) Enhanced image.

4. Experimentand Analysis

The performance of the proposed algorithm is evaluated both objectively and subjectively by utilizing ground-truth color patches. We also compare the proposed method with the state of the art methods. Both results demonstrate superior haze removing and color balancing capabilities of the proposed method over the others.

In the first experiment, we take two groups of image from alpha matting website [9]. In test 1 (natural image), we compare our method with Fattal and He's work. Here, we select patch radius r = 4, ε = 0.1×0.1, in Windows XP, Intel Core 2 (2.0GHz) with 1 GB RAM. In Figure 3, we show the results of different methods. The drawback of Fattal's method is elaborated on the Ref. [7]. To compare with He's method, our approach performs better. In He's approach, because of using soft matting, the visible mosaic artifacts are observed. Some of the regions are too dark (e.g. the center of the mountains image) and hazes are not removed (e.g. the sky of the image). Our approach not only works well in haze removal, but also cost little computational complex. We also compare the results in test 2 (urban image). We choose the patch radius r = 8, ε = 0.2×0.2 for computing. The results are shown in Figure 4. The results also demonstrate that our proposed method is the best.



(a) (b)

(c) (d)

Figure 3. Comparisons with different methods (Mountain), (a) Input image. (image size 225×300) , (b) Fattal's method, (c) He's method., (d) Our proposed method.

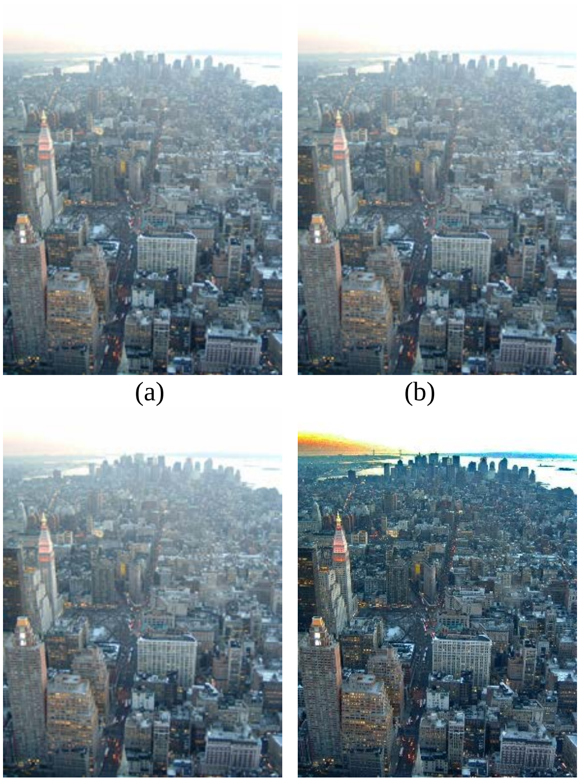
In order to prove the robustness of our method, we compare the proposed method with the state of the art methods. In Figure 5, we compare the Y.Y. Schechner's result, Fattal's result, and He's result in detail. Through visual analysis, we found the details of the boundary of the objects are smoothness, without gradient reversal artifacts. The visual assessment demonstrates that our proposed method performs well. In addition to the visual analysis of these figures, we conducted quantitative analysis, mainly from the perspective of mathematical statistics and the statistical parameters of the images (see Table 1). These include JPEG Quality Index [10] (A quality score between 1 and 10, 10 represents the best quality and 1 is the worst.), High-Dynamic Range Visual Difference Predictor2 (HDR-VDP2) [11], and CPU time. Table 1 displays the average ration (%) of the pixels that have been filtered by applying HDR-VDP2-IQA, CPU computing time and JPEG-QI measure on several images.

In Ref. [10], the no-reference JPEG quality index is defined as,

$$S = \alpha + \beta B^{\gamma_1} A^{\gamma_2} Z^{\gamma_3} \tag{12}$$

where α, β, γ_1, γ_2, and γ_3 are the model parameters that must be estimated with the subjective test image. B is the blockiness, which is estimated as the average differences across block boundaries. A is average absolute difference between in-block image samples. Z is the horizontal ZC rate. In this paper, the parameters used for test images are α =-245.9, β =261.9, γ_1 =-0.0240, γ_2 =0.0016, and γ_3 =0.0064, respectively.

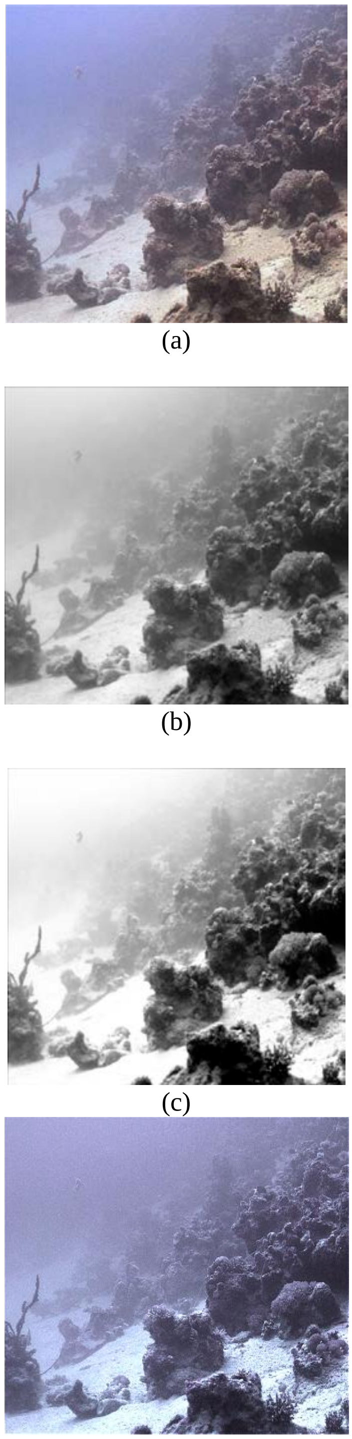
HDR-VDP-2 is a very recent metric that uses a fairly advanced model of human perception to predict both visibility of artifacts and overall quality in images [11]. The visual model used is based on existing experimental data, and accounts for all visible luminance conditions. This metric makes use of a detailed model of the optical and retinal pathway (including intra-ocular light scatter, photoreceptor spectral sensitivities and luminance masking) and takes into account contrast sensitivity for a wide range of luminance, as well as inter- and intra-channel contrast masking. We again refer the reader to the original publication for the details. HDR-VDP-2 can yield different outputs: an estimation of the probability of detecting differences between the two images compared, or an estimation of the quality of the test image with respect to the reference image. In this work we have used the latter, a prediction of the quality degradation with respect to the reference image, expressed as a mean-opinion score (from 0 to 100). We set the color encoding parameter of the metric to luma-display in order to work with the luminance channel of LDR images; the pixels-per-degree parameter, related to the viewing distance and the spatial resolution of the image is set to a standard value of 30. The HDR-VDP-2 results are presented in Figure 6.



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

Figure 4. Comparisons with different methods (New York City), (a) Input image(image size 288×384) , (b) Fattal's method. , (c) He's method, (d) Our proposed method.

Above all, the proposed method is superior to the others. We apply this approach in underwater image dehazing. At first, we take the clear water images, and then add some particles and mud in the experimental tank. We took 30 underwater images and dehazed them. The images size is below 400×400 pixels for avoiding out of memory. In order to analysis the performance of the proposed method for dehazing, we utilize 225 random affine transformations on both the original images and dehazed images. After that, we take SIFT matching [14] to correspondences between the unwarped image and each warped image. We calculate the inlier correspondences using RANSAC [15]. The result is shown in Figure 7. From the Figure, we can found that the proposed method has better performance in underwater dehazing. As can be observed compared with the other methods our proposed method is able to amplify better than the others while the loss of details is reduced. At the same time, it cost less computation time.

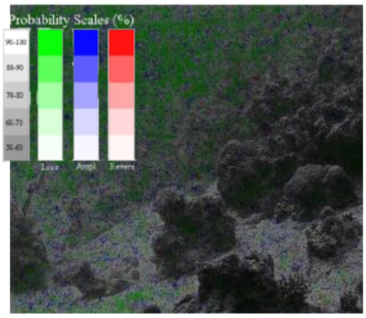


(a) (b) (c) (d)

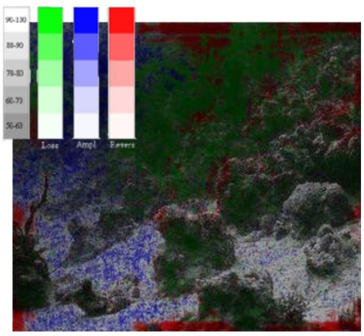
(e)

(f)

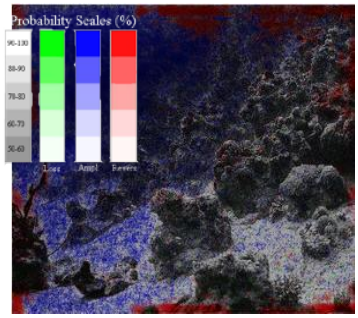
Figure 5. Comparison results with different methods in coral reefs image, (a) Input coral reefs image (image size 345×292), (b) Y.Y. Schechner's method, (c) Histogram equalization, (d) Fattal's method, (e) He's method, (f) Our proposed method.



(a)



(b)



(c)

Figure 6. Amplification and loss graphs of contrast induced by several operators by applying HDR-VDP2-IQA measure, (a) Fattal's, (b) He's, (c) Our proposed.

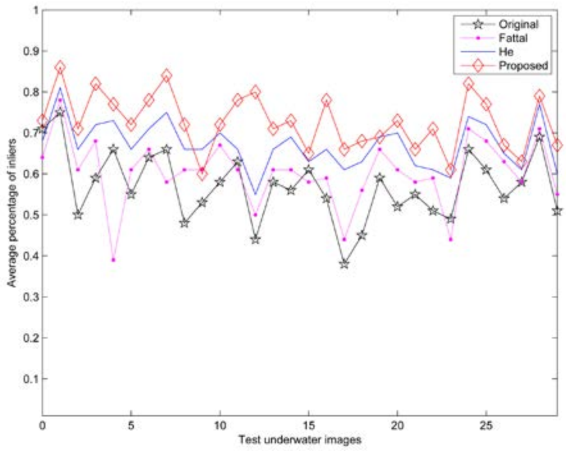


Figure 7. Average percentage of inlier correspondences for 30 underwater images.

5. Conclusions

In this paper we explored and successfully implemented a novel image dehazing techniques for underwater images. We proposed a simple prior based on the difference in attenuation among the different color channels, which inspire us to estimate the transmission depth map. Another contribution is to compensate the attenuation discrepancy along the propagation path, and to take the influence of the possible presence of an artificial lighting source into consideration. After these, our algorithm is faster than the state of the art algorithms. That is suitable for real-time computing in practice. In future, we consider utilizing the fusion [12] based method for dehazing (e.g. C.Ancuti et al. [13] by using two input images, and T. Treibitz et al. [4] to fuse different illumination images). Source code and test datasets of the proposed method is available at http:/www.boss.ecs.kyutech.ac.jp/~luhuimin/index.html.

Acknowledgments

This work was partially supported by Grants-in-Aid for Scientific Research of Japan (No. 19500478) and Research Found of Japan Society for the Promotion of Science (No.25・10713). The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers.

References

[1] Y.Y. Schechner, N. Karpel, "Recovery of Underwater Visibility and Structure by Polarization Analysis", IEEE Journal of Oceanic Engineering, vol.30, no.3, pp.570-587, 2005.

[2] C. Ancuti, C.O. Ancuti, T. Haber, P. Bekaert, "Enhancing Underwater Images and Videos by Fusion", In Proceeding of IEEE Conference on Computer Vision and Pattern Recognition, pp.81-88, 2012.

[3] Y.Y. Schechner, Y. Averbuch, "Regularized Image Recovery in Scattering Media", IEEE Transactions on Pattern Analysis & Machine Intelligence, vol.29, no.9, pp.1655-1660, 2007.

[4] T. Treibitz, Y.Y. Schechner, "Turbid Scene Enhancement Using Multi-Directional Illumination Fusion", IEEE Transactions on Image Processing, vol.21, no.11, pp.4662-4667, 2012.

[5] W. Hou, D.J. Gray, A.D. Weidemann, G.R. Fournier, J.L. Forand, "Automated Underwater Image Restoration and Retrieval of Related Optical Properties", In Proceeding of IEEE Internation Symposium of Geoscience and Remote Sensing, pp.1889-1892, 2007.

[6] A. Levin, D. Lishinski, Y. Weiss, "A Closed-form Solution to Natural Image Matting", IEEE Transactions on Pattern Analysis & Machine Intelligence, vol.30, no.2, pp.228-242, 2008.

[7] K. He, J. Sun, X. Tang, "Guided Image Filtering", In Proceeding of the 11th European Conference on Computer Vision, vol.1, pp.1-14, 2010.

[8] S. Paris, F. Durand, "A Fast Approximation of the Bilateral Filter using a Signal Processing Approach", International Journal of Computer Vision, vol.81, no.1, pp.24-52, 2009.

[9] Alpha Matting Website. http://www.alphamatting.com

[10] Z. Wang, H.R. Sheikh, A.C. Bovik, "No-reference Perceptual Quality Assessment of JPEG Compressed Images", In Proceeding of 2002 International Conference on Image Processing, vol.1, pp.477-480, 2002.

[11] R. Mantiuk, K.J. Kim, A.G. Rempel, W. Heidrich, "HDR-VDP2: A Calibrated Visual Metric for Visibility and Quality Predictions in All Luminance Conditions", ACM Transactions on Graphics, vol.30, no.4, pp.40-52, 2011.

[12] H. Lu, L. Zhang, S. Serikawa, "Maximum Local Energy: An Effective Approach for Image Fusion in Beyond Wavelet Transform Domain", Computers & Mathematics with Applications, vol.64, no.5, pp.996-1003, 2012.

[13] C. Ancuti, C.O. Ancuti, T. Haber, P. Bekaert, "Enhancing Underwater Images by Fusion", In Proceeding of ACM SIGGRAPH2011, Article No.32, 2011.

[14] M.E. Munichi, P. Pirjanian, E.D. Bernardo, L. Goncalves, et al., "SIFT-ing Through Features with ViPR: Application of Visual Pattern Recognition to Robotics and Automation", IEEE Robotics and Automation Magazine, vol.13, no.3, pp.72-77, 2006.

[15] M.A. Fischler, R.C. Bolles, "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography", Communications of the ACM, vol.24, no.6, pp.381-395, 1981.

[16] R. Fattal, "Signal Image Dehazing", In: SIGGRAPH, pp.1-9, 2008.

[17] K. He, J. Sun, X. Tang, "Single Image Haze Removal Using Dark Channel Prior", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.33, no.12, pp.2341-2353, 2011.

[18] K. Sun, B. Wang, Z. Zhou, Z. Zheng, "Real Time Image Haze Removal Using Bilateral Filter", Transactions of Beijing Institute of Technology, vol.31, no.7, pp.810-814, 2011.

[19] C. Xiao, J. Gan, "Fast Image Dehazing Using Guided Joint Bilateral Filter", The Visual Compute, vol.28, no.6/8, pp.713-721, 2012.

[20] Q. Yang, K. Tan, N. Ahuja, "Real Time O(1) Bilateral Filtering", In: IEEE Conference on

Computer Vision and Pattern Recognition, pp.1-8, 2009.

[21] J. Tarel, N. Hautiere, “Fast Visibility Restoration From A Single Color or Gray Level Image”, In: IEEE Conference of Computer Vision and Pattern Recognition, pp.2201-2208, 2008.

Huimin Lu

was born in Yangzhou, China, on November, 1985. He received the B.S. degree in Electronics Information Science and Technology from Yangzhou University in 2008. And he received M.S. degrees in Electrical Engineering from Kyushu Institute of Technology and Yangzhou University in 2011, respectively. Recently, he is a Ph.D. researcher in Kyushu Institute of Technology. His current research interests include computer vision, communication and deep-sea information processing. He is a member of IEEE, IEICE, ACM and JSPS.

Yujie Li

was born in Yangzhou, China, on March, 1988. She received the B.S. degree in Electronics Information Science and Technology from Yangzhou University in 2009. And she received M.S. degrees in Electrical Engineering from Kyushu Institute of Technology and Yangzhou University in 2012, respectively. Recently, he is a Ph.D. candidate in Kyushu Institute of Technology. Her current research interests include computer vision, social welfare system. She is a member of IEEE.

Shiyuan yang

was born in Tianjing, China, on August, 1963. He received the B.S. and M.S. degrees in Electrical Engineering from Kyushu Institute of Technology in 1990 and 1993, respectively. And he received Ph.D. degree in Electrical Engineering from Kyushu Institute of Technology, in 2000. Recently, he is an assistant Professor in Kyushu Institute of Technology. His current research interests include holography optics, laser design. He is a member of IIAE, SPIE, and OSA.

Seiichi Serikawa

was born in Kumamoto, Japan, on June, 1961. He received the B.S. and M.S. degrees in Electronic Engineering from Kumamoto University in 1984 and 1986. During 1986 -1990, he stayed in Tokyo Electron Company. From 1990 to 1994, he was an assistant in Kyushu Institute of Technology. He received the Ph.D. degree in Electronic Engineering from Kyushu Institute of Technology, in 1994. From 1994 to 2000, he was an Assistant Professor at the Kyushu Institute of Technology. From 2000 to 2004, he was an Associate Professor at the Kyushu Institute of Technology. Science 2004, he has been a Professor at the Kyushu Institute of Technology. He worked as Vice President of School of Engineering in Kyushu Institute of Technology from 2010 to 2012. Recently, he is the Dean of Department of Electrical Engineering and Electronics of School of Engineering in Kyushu Institute of Technology. His current research interests include computer vision, sensors, and robotics. He is a member of IEEJ, IEICE, and the president of IIAE.