# Hardware Implementation of an Efficient Guided Image Filter for Underwater Image Restoration

I-Chang Huang, Cheng-Hao Chang, Mu-Fan Lin, Hung-Hsiang Chen, and Shiann-Rong Kuang
Department of Computer Science and Engineering, National Sun Yat-sen University, Kaohsiung, Taiwan
Email: d122543064@gmail.com, srkuang@cse.nsysu.edu.tw

***Abstract*—Transmission estimation is an important step in underwater image restoration based on dark channel prior. To obtain good visual quality, guided image filter is usually adopted to refine the crude transmission map of underwater images. In this paper, we first develop a simplified guided image filter for underwater image restoration to reduce the computational complexity of transmission estimation while maintaining good visual quality. Subsequently, a low-cost architecture for the simplified guided image filter is proposed to achieve real-time Full-HD (1920 × 1080) underwater image restoration. The proposed architecture is synthesized with TSMC 90nm CMOS technology and the result shows that it can operate at 100MHz and support Full-HD image restoration at a throughput of 30frame/s with 10.5K gate counts and 1.35KB on-chip memory. Compared to previous hardware design of guided image filter, the proposed design uses 11.3% of gate counts and 42.2% of on-chip memory to achieve the same frame rate with almost no visual quality degradation.**

***Index Terms*—underwater image, dark channel prior, transmission estimation, guided image filter**

## I. Introduction

Clear and brilliant underwater image is crucial to many applications in ocean engineering, ocean science, and ocean biology [1]. A lot of approaches have been proposed to restore the underwater image, but most of them require high computing resource and long execution time. Recently, the work in [2] introduced a dark channel prior method to remove haze from a single input image efficiently. Several researches [3]-[8] have extended this method to restore the visual quality of underwater image. For this kind of image restoration methods, the concept of dark channel is applied to estimate the transmission map $\tilde{t}(x)$ which is reasonably good but contains some block effects. In [2], a soft matting method was used to refine $\tilde{t}(x)$ to remove the block effect, but this optimization has very high computational complexity and requires heavy computing resource. Therefore, many underwater image restoration systems (e.g., [5]-[8]) adopted the guided image filter (GIF) [9] to efficiently optimize $\tilde{t}(x)$ into $t(x)$.

The GIF can perform edge-preserving smoothing operation and has a fast and non-approximate linear-time algorithm. Due to its simplicity, efficiency and high quality, various applications have adopted the GIF as the filtering method. Moreover, some dedicated hardware implementations of GIF have been proposed for meeting the requirements of real-time applications. For example, an efficient VLSI architecture of GIF based on the double integral image technique was proposed in [10] to achieve real-time HD applications. In [10], integral image engines were utilized to compute the sum of the pixel values in the windows at all stages, leading to fewer computations. Moreover, the design in [10] employed the stripe-based method, which decomposes a frame into several vertical stripes and performs the computation stripe by stripe instead of on the whole frame, to significantly reduce the on-chip memory cost. As a result, the GIF in [10] implemented with TSMC 90-nm CMOS cell library can operate at 100MHz and support for Full-HD (1920×1080) 30 frame/s with 92.9K gate counts and 3.2KB on-chip memory. However, the double integral image architecture requires high off-chip memory bandwidth since it has to load and store the image data and the intermediate coefficients frequently.

The work in [11] proposed the VLSI design of moving sum based GIF, whose FPGA prototype can process 720P (1280 × 720) HD video at 60 frame/s. In the moving sum based GIF, the integral image engine in [10] is replaced by the mean filter with windows of radius $r$, where the column sum memory (its size depends on the image width) and a shift register (queue) with size $2r+1$ are employed to store the required intermediate column sums. Compared to the GIF architecture in [10], the moving sum based GIF requires fewer arithmetic units and smaller bit number since the column sum is smaller than the integral image when the window size is the same. However, the on-chip memory overhead of the moving sum based GIF may become unacceptable when the image width and the window size of GIF becomes larger.

In fact, none of previous VLSI architectures of GIF is customized and optimized according to the characteristic of underwater image restoration. To overcome the drawbacks of previous designs, we first simplify the GIF algorithm for underwater image restoration to significantly reduce the computational complexity and memory requirement with almost no quality degradation. Afterwards, a low-cost VLSI architecture for the simplified guided image filter is proposed by combining the advantages of GIF architectures in [10] and [11] to further reduce the cost of on-chip memory and computational resource. The synthesis result of the

Manuscript received June 10, 2017; revised November 23, 2017.


doi: 10.18178/ijsps.5.3.94-99

proposed GIF shows that our design uses much fewer gate counts and on-chip memory to achieve the same frame rate with almost no visual quality degradation than the previous design in [10].

*Algorithm*: Guided Image Filter [11]

*Input*: guidance image $I$, guided image $p$
*Output*: output $q$ (filtered image)
*Parameters*: $r$, $\varepsilon$
1: $mean_I = f_{mean}(I)$; $mean_p = f_{mean}(p)$;
$corr_I = f_{mean}(I.*I)$; $corr_{Ip} = f_{mean}(I.*p)$;
2: $var_I = corr_I - mean_I.* mean_I$;
$cov_{Ip} = corr_{Ip} - mean_I.* mean_p$;
3: $a = cov_{Ip}/(var_I + \varepsilon)$; $b = mean_p - a.* mean_I$;
4: $mean_a = f_{mean}(a)$; $mean_b = f_{mean}(b)$;
5: $q = mean_a.*I + mean_b$;

Figure 1. Pseudocode of GIF algorithm.

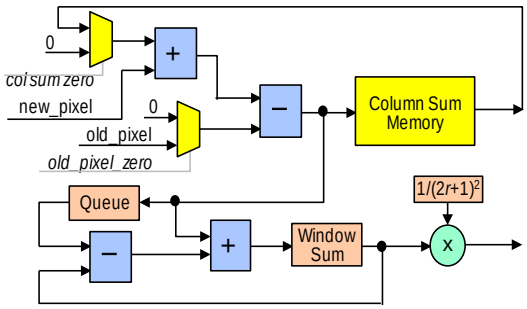



Figure 2. Hardware architecture of mean filter proposed in [11].

## II. Background

### A. Haze Removal

The image formation model [12] widely used to describe hazy images can be expressed as

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

where $I$ is the observed image (i.e., hazy image), $J$ is the scene radiance, $t(x)$ is the transmission along the ray, and $A$ is the global background light. The first term $J(x)t(x)$ is treated as direct attenuation and the second term $A(1-t(x))$ denotes airlight. The transmission map $t$ is defined as

$$t(x) = e^{-\beta d(x)} \quad (2)$$

where $\beta$ is the scattering coefficient of the atmosphere and $d$ is the scene depth at $x$. However, it is difficult to calculate the scattering coefficient and the scene depth exactly. Recently, He et al. [2] proposed a dark channel prior method to roughly estimate the transmission map $t$. They observed the haze-free scenes and found that at least the intensity of one color channel is far lower than others. The following equation describes this phenomenon.

$$J^{dark}(x) = \min_{y\in\Omega(x)}\left(\min_{c\in\{r,g,b\}} J^c(y)\right) \approx 0 \quad (3)$$

where $J^c$ is a color channel of $J$ and $\Omega(x)$ is a local patch centered at $x$. When the scene $J$ is a haze-free image, the intensity value of $J^{dark}$ is low and tends to be zero. Moreover, [2] assumed that the transmission in a local patch $\Omega(x)$ is constant. After rewriting (1) and applying (3) to (1), the transmission can be estimated as

$$\tilde{t}(x) = 1 - \min_{y\in\Omega(x)}\left(\min_{c}\frac{I^c(y)}{A^c}\right) \quad (4)$$

However, the red color component usually undergoes maximum attenuation in an underwater environment. Therefore, several works [5]-[8] for underwater image restoration modified the dark channel prior according to the absorption effect in the red channel to enhance the restoration performance. Nevertheless, this estimated transmission $\tilde{t}(x)$ is coarser and requires a smoothing operator which can preserve its edges. In [2], a soft matting procedure was employed to refine the coarse map $\tilde{t}(x)$ into $t(x)$, but very high computational complexity and heavy computing resource are required. Instead of soft matting, many underwater image restoration systems [5]-[8] applied GIF [9] to refine $\tilde{t}(x)$. After $A$ and $t$ have been obtained, the desired scene radiance $J$ can be recovered as

$$J = \frac{I(x)-A}{t(x)} + A \quad (5)$$

### B. Guided Image Filtering

Given a guidance image $I$ and an input image $p$, a filtered output image $q$ can be produced by GIF as follows:

$$q_i = a_k I_i + b_k, \quad \forall i \in w_k \quad (6)$$

where $a_k$ and $b_k$ are linear constant coefficients in a window $w_k$ with radius $r$, centered at pixel $k$. Note that a radius-$r$ window is a window with size $(2r+1) \times (2r+1)$. For the underwater image restoration based on dark channel prior, the guidance image is the intensity of observed image while the input image is the estimated transmission map $\tilde{t}(x)$. To find the appropriate coefficients, [9] defined a cost function in window $w_k$ as follows

$$E(a_k, b_k) = \sum_{i\in w_k}\left((a_k I_i + b_k - p_i)^2 + \varepsilon a_k^2\right) \quad (7)$$

where the parameter $\varepsilon$ prevents $a_k$ from being too large. The solutions of (7) for $a_k$ and $b_k$ are listed in (8) and (9), respectively.

$$a_k = \frac{\frac{1}{|w|}\sum_{i\in w_k} I_i p_i - \mu_k \overline{p_k}}{\sigma_k^2 + \varepsilon} \quad (8)$$

$$b_k = \overline{p_k} - a_k \mu_k \quad (9)$$

where $|w|$ is the number of pixels in window $w_k$, $\mu_k$ and $\sigma_k^2$ are the mean and the variance of $I$ in window $w_k$, and $\overline{p_k}$ is the mean of input image $p$ in window $w_k$. For a

window $w_k$ with radius $r$, $|w| = (2r+1) \times (2r+1)$. After the linear coefficients $a_k$ and $b_k$ are computed, the filtering output $q_i$ can be obtained by replacing them into (9). However, any pixel $i$ is covered by several overlapping windows, different output pixel value $q_i$ may be produced by different local windows. In general, this problem is solved using the averaging strategy of overlapping windows and the filtering output is redefined as

$$q_i = \frac{1}{|w|} \sum_{k, i \in w_k} (a_k I_i + b_k) = \overline{a_k} I_i + \overline{b_k} \quad (10)$$

where $\overline{a_k}$ and $\overline{b_k}$ are the mean values of $a_k$ and $b_k$ in the window $w_k$, respectively.

Fig. 1 shows the pseudocode of GIF algorithm adopted in [11]. For a full-HD frame, the memory demand for the GIF algorithm is (4 + 2) × 1920 × 1080 × 4 bytes ≈ 50 MB if $I$, $I^2$, $p$, $Ip$, $a$, and $b$ are represented with the IEEE 754 single-precision format. Fig. 2 and Fig. 3 show the hardware architectures of mean filter and GIF proposed in [11], respectively, where fixed-point calculation is used. The mean filter with windows of radius $r$ receives the new pixel and the old pixel, and outputs the mean corresponding to the window being filtered. The mean filter maintains a sum for each column in the image to be filtered, where each column sum accumulates $2r+1$ pixels and the window sum is obtained by adding $2r+1$ adjacent column sums.

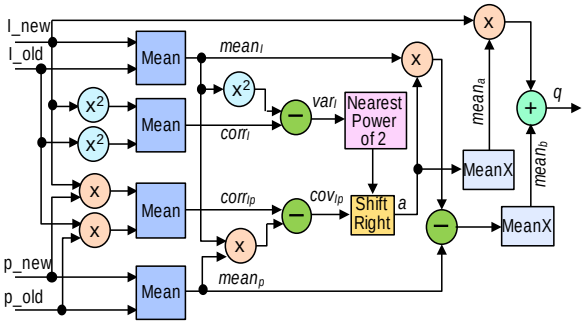



Figure 3. Hardware architecture of GIF proposed in [11].

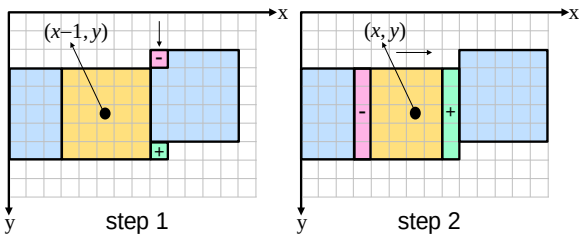



Figure 4. The mean filtering process with $r$ = 2.

The mean filtering process with $r$ = 2 is illustrated in Fig. 4. When the filter window moves from pixel ($x$–1, $y$) to ($x$, $y$) in the current image line, the first step consists of updating the current column sum of the filter window by subtracting the topmost old pixel and adding the bottommost new pixel. In the second step, the window sum of current pixel ($x$, $y$) can be calculated by subtracting its leftmost column sum (old column sum), and adding the updated column sum computed in step 1 (new column sum).

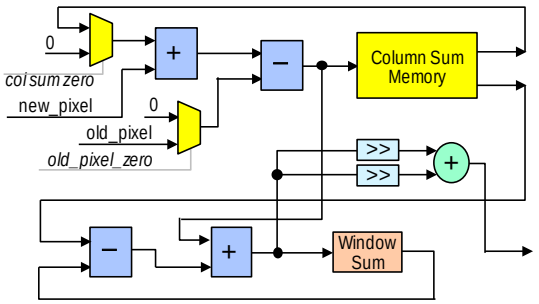



Figure 5. Hardware architecture of proposed mean filter.

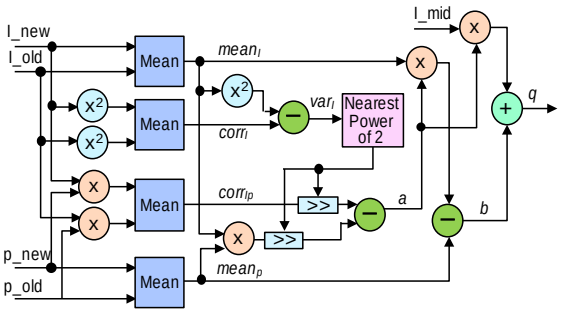



Figure 6. Hardware architecture of proposed GIF.

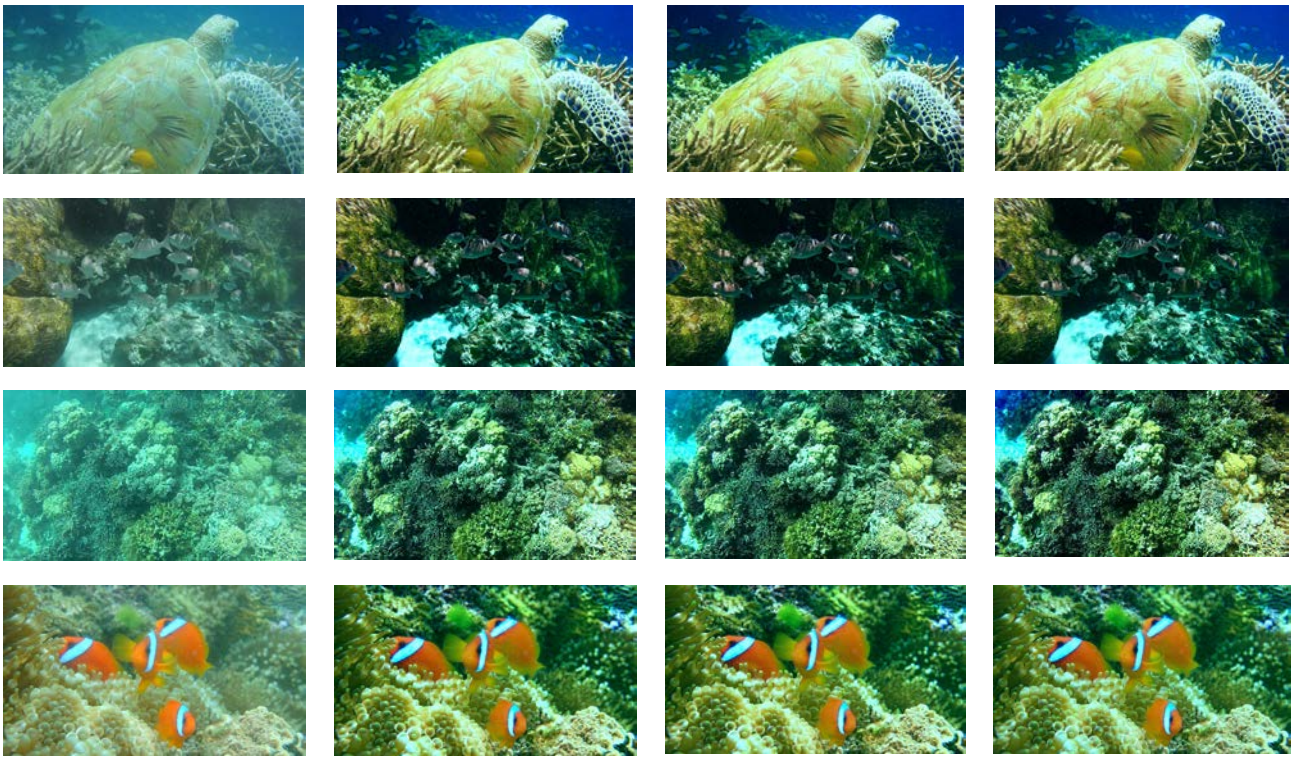
## III. Roposed Guided Image Filter

As can be seen in Fig. 1, GIF algorithm performs six mean filtering processes to obtain the filtered output image $q$, where the last two mean filters are used to calculate $\overline{a_k}$ and $\overline{b_k}$ in (10). Fortunately, we find that the transmission map optimization is insensitive to the mean filtering processes of step 4 in Fig. 1 according to a large number of experiments. That is, the step 4 in Fig. 1 can be removed during the underwater image restoration with almost no quality degradation. Note that step 5 must be revised into $q = a * I + b$ if step 4 is removed. Obviously, the simplified algorithm which only includes four mean filtering processes results in lower computational complexity and less memory and bandwidth demand.

Nevertheless, the on-chip memory demand of each mean filter shown in Fig. 2 is still very high. The size of column sum memory in Fig. 2 depends on the image width and hence dual-port SRAM is used to store the column sums. Moreover, a shift register (queue) with size $2r+1$ is used to store the old column sums to avoid the access conflicts of column sum memory. To lower the on-chip memory demand of mean filter in Fig. 2, we employ the stripe-based method used in [10] to reduce the size of column sum memory. The stripe-based method decomposes a frame into several vertical stripes, thus the size of column sum memory is reduced from the width of frame to the width of stripe plus the extended region. Given a frame with width $W_f$ and height $H_f$, a window with radius $r$, and the stripe width $W_s$, the size of each input stripe will be $(W_s + 2r) \times H_f$ and the size of column sum memory will be reduced from $W_f$ to $W_s$+2$r$.

Because the size of column sum memory is significantly reduced from $W_f$ to $W_s$+2$r$, register file rather than SRAM can be used to enhance the performance of memory accesses. For the mean filter shown in Fig. 2, two column sums have to be read from and one column sum has to be written to column sum memory in every operating cycle. Therefore, one register file with two read ports and one write port is adopted in the proposed mean filter to store ($W_s$+2$r$) column sums. As a result, shift register (queue) with size 2$r$+1 in Fig. 2 can be removed from the proposed mean filter and a memory controller is required to handle the access of column sum memory. Fig. 5 and Fig. 6 illustrate the hardware architectures of proposed mean filter and GIF, respectively. In Fig. 5, the shift register (queue) in Fig. 2 is removed, and the multiplier in Fig. 2 is replaced with shifts and additions/subtractions since $1/(2r+1)^2$ is a constant if $r$ is known. In Fig. 6, the word number and access sequence of column sum memory in each mean filter are the same. Therefore, the four column sum memories can be merged into one register file with bigger bit number.

For the example of a full-HD frame ($W_f$ and $H_f$ are 1920 and 1080, respectively) with window size 31× 31 (i.e., $r$ = 15) and stripe width $W_s$ = 120, the frame will be decomposed into 16 vertical stripes and the size of column sum memory is reduced from 1920 to 150. Moreover, if the input pixels of mean filter are 8 (16) bits, the column sum and the window sum will be 14 (22) and 18 (26) bits, respectively. That is, one 150-word (each word 72 bits) register file is required in the proposed GIF.



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

Figure 7. Qualitative Comparison of different GIF designs (a) Original underwater images with a size 1920×1080, from top to bottom: Image 1, Image 2, Image 3, and Image 4, (b) Enhanced results by Trad_GIF, (c) Enhanced results by Simp_GIF, and (d) Enhanced results by HW_GIF.

TABLE I. QUANTITATIVE COMPARISON OF DIFFERENT GIF DESIGNS IN TERMS OF ENTROPY, SIFT, AND COLORFULNESS

| Images | GIF design | Entropy | SIFT | Colorfulness |
|---|---|---|---|---|
| Image 1 | Trad_GIF | 5.6699 | 3208 | 90.0030 |
| | Simp_GIF | 5.6747 | 3197 | 89.9981 |
| | HW_GIF | 5.6567 | 3088 | 89.9528 |
| Image 2 | Trad_GIF | 6.5754 | 9149 | 58.9561 |
| | Simp_GIF | 6.5889 | 8938 | 58.9529 |
| | HW_GIF | 6.4969 | 9194 | 59.3980 |
| Image 3 | Trad_GIF | 5.0883 | 37351 | 68.7135 |
| | Simp_GIF | 5.0876 | 37042 | 68.7330 |
| | HW_GIF | 5.0428 | 33778 | 68.0130 |
| Image 4 | Trad_GIF | 6.1360 | 1869 | 88.4697 |
| | Simp_GIF | 6.1434 | 1794 | 88.1461 |
| | HW_GIF | 6.2000 | 2008 | 87.1271 |

## IV. Experimental Results

To verify the efficiency of the proposed design, an underwater image restoration system similar to the one in [5] was implemented in C++ language with single-precision floating-point (FP) format. The traditional GIF algorithm in Fig. 1 (denoted as Trad_GIF), the simplified GIF algorithm (denoted as Simp_GIF), and the hardware design of proposed GIF (denoted as HW_GIF) are embedded in this restoration system to evaluate the restoration performance. Note that Trad_GIF and Simp_GIF are software implementation with single-precision FP format and HW_GIF is the hardware design implemented in SystemC with the fixed-point format as described in Section III. Furthermore, several Full-HD underwater images (Image 1 to 4) in different water conditions and different scene configurations were used for the experiments. The output results shown in Fig. 7 demonstrate that the visual quality produced by Simp_GIF and HW_GIF is comparable to that of Trad_GIF.

In addition to the aforementioned qualitative comparison, the quantitative evaluation is carried out to further validate the results of proposed design. The image quality metrics in terms of entropy, SIFT (Scale-Invariant Feature Transform) local feature points [13] and colorfulness [14] are adopted, and Table I lists the comparative values for the underwater images shown in Fig. 7. The value of entropy represents the valuable information contained in the recovered images. The SIFT local feature points indicate the global contrast and local features while the colorfulness metric expresses the colorfulness of an image. The results in Table I show that the visual quality in terms of entropy, SIFT, and colorfulness of underwater images produced by Simp_GIF and HW_GIF is very close to that of Trad_GIF. That is, the simplifications in Simp_GIF and HW_GIF result in almost no quantitative performance degradation compared with Trad_GIF.

TABLE II. Comparison with Previous VLSI Design of GIF

| GIF Design | Kao [10] | Proposed |
|---|---|---|
| Technology | TSMC 90nm | TSMC 90nm |
| Frame Size | 1920 × 1080 | 1920 × 1080 |
| Frame Rate (fps) | 30 | 30 |
| Filter Window Size | 31 × 31 | 31 × 31 |
| Stripe Width | 120/180 | 120/150 |
| Operating Frequency | 100 MHz | 100 MHz |
| Gate Counts | 92.9 K | 10.5 K |
| On-chip Memory (Byte) | 3.2 K | 1.35 K |
| Bandwidth (bit) | 262.31M | 116.12M |

To evaluate the hardware efficiency, the hardware architecture of the proposed HW_GIF in Fig. 6 with $r$ = 15 and $W_s$ = 120 was implemented in Verilog HDL and synthesized by using the Synopsys Design Compiler with the TSMC 90nm CMOS standard cell technology library. Synthesis results show that our design spends 10.5K gate counts (excluded register file) and 1.35KB on-chip memory. It can operate at 100MHz clock frequency and achieve the throughput of 30 frame/s Full-HD 1080p. The synthesis results are summarized and compared with [10] in Table II. As can be seen in Table II, the proposed design uses 11.3% of gate counts, 42.2% of on-chip memory, and 44.3% of bandwidth to achieve the same frame rate.

## V. Conclusion

This paper has proposed a simplified GIF and its low-cost hardware architecture to achieve real-time Full-HD underwater image restoration while maintaining good visual quality. The synthesis result in TSMC 90nm CMOS technology has shown that the proposed VLSI design of GIF can operate at 100MHz and support Full-HD image restoration at a throughput of 30 frame/s with only 10.5K gate counts and 1.35KB on-chip memory. The visual quality of the proposed GIF was also evaluated through extensive experiments.

## Acknowledgment

This work was supported in part by the Ministry of Science and Technology, Taiwan, under Grant MOST 105-2221-E-110-088. The authors would like to thank the contributions of Taiwan Semiconductor Manufacturing Company Limited and National Chip Implementation Center, Taiwan, for their support in technology data.

## References

[1] J. K. Guo, C. C. Sung, and H. H. Chang, "Restoration of underwater vision using a two-phase regularization mechanism," in *Proc. 7th International Congress on Image and Signal Processing*, 2014, pp. 243-247.

[2] K. M. He, J. Sun, and X. O. Tang, "Single image haze removal using dark channel prior," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 33, no. 12, pp. 2341-2353, 2011.

[3] H. Y. Yang, P. Y. Chen, C. C. Huang, Y. Z. Zhuang, and Y. H. Shiau, "Low complexity underwater image enhancement based on dark channel prior," in *Proc. Second International Conference on Innovations in Bio-inspired Computing and Applications*, 2011, pp. 17-20.

[4] S. J. Horng, P. J. Liu, and J. Sh. Lin, "Improving the contrast enhancement of oceanic images using modified dark channel prior," in *Proc. International Symposium on Computer, Consumer and Control*, 2016, pp. 801-804.

[5] A. Galdran, D. Pardo, A. Picon, and A. A. Gila, "Automatic red-channel underwater image restoration," *Journal of Visual Communication and Image Representation*, vol. 26, pp 132-145, Jan. 2015.

[6] S. Borkar and S. V. Bonde, "Underwater image restoration using single color channel prior," in *Proc. International Conference on Signal and Information Processing*, 2016, pp. 1-4.

[7] C. Li, J. Quo, Y. Pang, S. Chen, and J. Wang, "Single underwater image restoration by blue-green channels dehazing and red channel correction," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, 2016, pp. 1731-1735.

[8] C. Y. Cheng, C. C. Sung, and H. H. Chang, "Underwater image restoration by red-dark channel prior and point spread function deconvolution," in *Proc. IEEE International Conference on Signal and Image Processing Applications*, 2015, pp. 110-115.

[9] K. He, J. Sun, and X. Tang "Guided image filtering," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 35, no. 6, pp. 1397-1409, 2013.

[10] C. C. Kao, J. H. Lai, and S. Y. Chien, "VLSI architecture design of guided filter for 30 frames/s full-HD video," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 24, no. 3, pp. 513-524, 2014.

[11] C. Ttofis, C. Kyrkou, and T. Theocharides, "A low-cost real-time embedded stereo vision system for accurate disparity estimation based on guided image filtering," *IEEE Trans. on Computers*, vol. 65, no. 9, pp. 2678-2693, 2016.
[12] R. Fattal, "Single image dehazing," *ACM Transactions on Graphics*, vol. 27, no. 3, pp. 1-9, 2008.
[13] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91-110, 2004.
[14] D. Hasler and S. Suesstrunk, "Measuring colorfulness in natural images," in *Proc. IS&T/SPIE Electronic Imaging 2003: Human Vision and Electronic Imaging VIII*, 2003, pp. 87-95.

**I-Chang Huang** received the B.S. degree in computer science and engineering from National Sun Yat-sen University, Kaohsiung, Taiwan, in 2016. He is currently working toward the M.S. degree at the Department of Computer Science and Engineering, National Sun Yat-sen University, Kaohsiung, Taiwan. His research interests include VLSI architecture design of image processing and digital signal processing.

**Cheng-Hao Chang** received the B.S. and M.S. degrees in computer science and engineering from National Sun Yat-sen University, Kaohsiung, Taiwan, in 2015 and 2017, respectively. His research interests include VLSI design and implementation for digital signal processing systems.

**Mu-Fan Lin** received the B.S. degree in electrical engineering from Feng Chia University, Taichung, Taiwan, in 2016. He is currently working toward the M.S. degree at the Department of Computer Science and Engineering, National Sun Yat-sen University, Kaohsiung, Taiwan. His research interests include VLSI architectures for public-key cryptosystems.

**Hung-Hsiang Chen** received the B.S. degree in electrical engineering from National University of Kaohsiung, Kaohsiung, Taiwan, in 2016. He is currently working toward the M.S. degree at the Department of Computer Science and Engineering, National Sun Yat-sen University, Kaohsiung, Taiwan. His research interests include VLSI design and RSA cryptosystems.

**Shiann-Rong Kuang** received the B.S. degree from National Central University, Jhongli, Taiwan, in 1990, and the M.S. and Ph.D. degrees from National Cheng Kung University, Tainan, Taiwan, in 1992 and 1998, respectively, all in electrical engineering. He is currently a Professor with the Department of Computer Science and Engineering, National Sun Yat-sen University, Kaohsiung, Taiwan. His current research interests include low-power/high-performance VLSI architectures for public-key cryptosystems and digital signal processing systems.