Real-Time Visible-Infrared Image Fusion using Multi-Guided Filter

Woojin Jeong, Bok Gyu Han, Hyeon Seok Yang, and Young Shik Moon*
Department of Computer Science and Engineering, Hanyang University,
Ansan, Gyeonggi-do 426-791, Republic of Korea
[e-mail: wjjeong@visionlab.or.kr, bghan@visionlab.or.kr, hsyang@visionlab.or.kr, ysmoon@hanyang.ac.kr]
*Corresponding author: Young Shik Moon

Received August 21, 2018; revised December 1, 2018; accepted December 8, 2018; published June 30, 2019

Abstract

Visible-infrared image fusion is a process of synthesizing an infrared image and a visible image into a fused image. This process synthesizes the complementary advantages of both images. The infrared image is able to capture a target object in dark or foggy environments. However, the utility of the infrared image is hindered by the blurry appearance of objects. On the other hand, the visible image clearly shows an object under normal lighting conditions, but it is not ideal in dark or foggy environments. In this paper, we propose a multi-guided filter and a real-time image fusion method. The proposed multi-guided filter is a modification of the guided filter for multiple guidance images. Using this filter, we propose a real-time image fusion method. The speed of the proposed fusion method is much faster than that of conventional image fusion methods. In an experiment, we compare the proposed method and the conventional methods in terms of quantity, quality, fusing speed, and flickering artifacts. The proposed method synthesizes 57.93 frames per second for an image size of 320 × 270. Based on our experiments, we confirmed that the proposed method is able to perform real-time processing. In addition, the proposed method synthesizes flicker-free video.

Keywords: Visible-infrared image fusion, guided image filtering, real-time processing

http://doi.org/10.3837/tiis.2019.06.018 **ISSN : 1976-7277**

1. Introduction

Image fusion is the process of synthesizing two or more images into one image. It assists human perceptibility because it incorporates complementary information of images. Visible-infrared image fusion is a typical task of image fusion. The infrared image is able to capture the thermal characteristics of an object. This type of image is useful at night or in foggy environments but typically results in blurry images. By contrast, the visible image shows sharp details under normal lighting conditions. Thus, it can assist in the identification of the details of an object and its background. However, it is often impossible to identify a target at night or in foggy environments.

Combining these two complementary images is advantageous for identifying objects under different lighting conditions. This not only assists human perceptibility but is also useful for application to higher order algorithms involving object tracking and object recognition. In particular, in surveillance systems, the infrared image is advantageous for identifying the position of the target, and the visible image is useful for exploring the details of the surrounding environment. Therefore, visible-infrared image fusion has important applications that include surveillance systems.

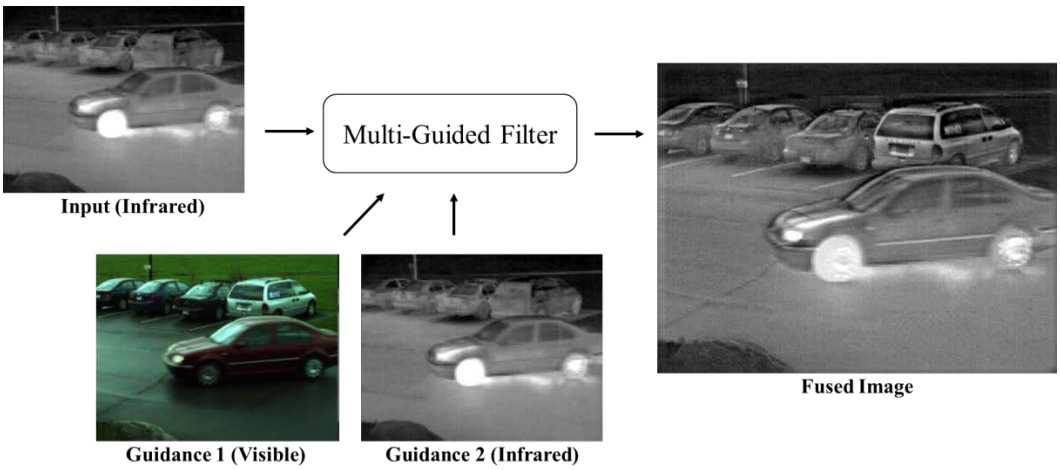


Fig. 1. Proposed image fusing process

There have been numerous investigations on visible-infrared image fusion. Conventional methods have shown superior results in terms of the sharpness of the target and the details of the surrounding environment. However, since a complicated algorithm is used, these approaches are time intensive. This problem is a disadvantage especially in real-time monitoring, which is an important function of modern surveillance systems.

In this paper, we propose a real-time visible-infrared image fusion method. For this purpose, we propose a multi-guided filter. The multi-guided filter is a modification of the well-known guided filter. First, we analyzed the problems associated with the use of a guided filter for image fusion. We discovered significant drawbacks of this filter and subsequently modified it to produce multi-guided filters. The guided filter is characterized by the gradient of the guidance image. This characterization is not suitable to visible-infrared image fusion in which

the information in both images is important. Therefore, we proposed the multi-guided filter based on two or more images as the guidance images. The multi-guided filter effectively synthesizes two images. At the same time, the proposed method exhibited a fast fusing speed.

Through experiments, we verified that the proposed method quickly synthesizes images. At the same time, we showed quantitative and qualitative performances comparable to the conventional methods. In addition, we confirmed that there are no flickering artifacts, which often occur in consecutive image processing. The proposed method synthesizes images with a size of 320×270 at a speed of 58 fps. This is fast enough for real-time processing. The contributions of this paper are as follows:

- We analyze the reasons why the guided filter is not suitable for image fusion.
- We propose a multi-guided filter. The multi-guided filter is a modified algorithm of the guided filter for multiple guidance images.
- We propose a real-time image fusion method using the multi-guided filter. The proposed image fusing method is faster than conventional methods.
- We compare the conventional methods and the proposed image fusion method in terms of quantitative, qualitative, fusing speed, and flickering artifacts.

The structure of the paper is as follows. In Chapter 2, we review studies related to visible-infrared image fusion. Chapter 3 briefly reviews the guided filter. Chapter 4 analyzes the problem of image fusion using the guided filter. In Section 5, we derive the multi-guided filter from the guided filter. In Chapter 6, we compare the proposed method with other methods through experimentation. Chapter 7 is the conclusion.

2. Related Work

In this chapter, we briefly review previous visible-infrared image fusion methods.

2.1 Image Decomposing and Recomposing based Image Fusion

Image decomposing and recomposing based image fusion methods are mainstream approaches for visible-infrared image fusion. These methods decompose visible-infrared images and reconstruct a new image according to the rules for fusion. Pyramid-based methods [1][2][3], wavelet-based methods [4][5][6], curvelet transform-based methods [7], multi-resolution singular value decomposition-based method [8], bilateral filter-based method [9], guided filtering-based methods [10][11][12], sparse representation-based methods [13], fourth order partial differential based method [14], and visual saliency based method [15][16] have been studied. Image decomposing and recomposing based fusion methods have shown superior results in terms of the quality of the fused image. However, these methods can be quite complex when high-quality results are needed. Therefore, they are usually not suitable for real-time processing.

2.2 Iterative Optimization based Image Fusion

Ma et al. [17] proposed an image fusion method using iterative optimization. The objective of their method is to transfer the details of the visible image into the infrared image. The method used by Ma et al. successfully synthesizes the visible-infrared image. However, it requires a long computational period because it is an iterative optimization-based method,

which is an inherently time-intensive algorithm. This method is suitable for military purposes, where speed is an issue.

2.3 Deep Neural Network based Image Fusion

Recently, deep neural network techniques have demonstrated superior results in various fields. Thus, deep neural network based image fusion methods have been proposed. Liu et al. [18] use a convolutional neural network to construct a focus map and synthesize images of well-focused regions. The main objective in this case was to synthesize differently focused images. But, their method was also useful for visible-infrared image fusion, medical image fusion, and multiple exposure image fusion. In spite of these results, this method is unsuitable for real-time image fusion. When a GPU was used, their method processed a 520×520 image in 0.33 s.

Xu et al. [19] also used deep neural networks for multi focused image fusion. The method of Xu et al. demonstrated better results than other conventional methods. However, this method is also a time-consuming algorithm. Ma et al. [20] proposed a generative adversarial network based image fusion method named FusionGAN. Their method finds the optimal fused image that keeps thermal radiation and textural details in the source without a manually tuned fusion rule. The method of Ma et al. demonstrated superior fusion performance but is not suitable for real-time processing owing to its speed.

2.4 Guided Filter based Image Fusion

The guided filter is an edge preserving-smoothing filter [21]. This filter is used in a wide variety of applications such as soft matting, image enhancement, and multi-exposure image combination. The guided filter inherits the conventional bilateral filter [22][23] and the joint bilateral filter [24] but is faster than conventional filters. The guided filter is fast because it is based on a simple one-pass algorithm. Gan et al. [11] used the guided filter to fuse images in a stage of weight refinement. The guided filter was also used for visible-infrared image fusion in a previous investigation by Li et al. [10]. Li et al. used the guided filter as a sub process of the image decomposing and recomposing processes.

A gradient domain guided filter was proposed for improving the edge-preserving ability via exploiting constraints [25]. Zhu et al. [12] proposed an image fusion method using the gradient domain guided filter and the image decomposing and recomposing. Their method demonstrated superior performance in both visual perceptibility and quantity. However, the method of Zhe et al. also used the guided filter as a sub process of the image decomposing and recomposing.

In the conventional image fusion methods, the guided filter was a sub stage and a sub process. Thus the conventional guided filter based methods shared common characteristics. They are complicated algorithms and have tunable parameters by hands.

2.5 Fused Image Evaluation Metric

In order to measure the fused image, various metrics have been introduced. Mutual information (MI) [26] and normalized mutual information (NMI) [27] represent the average of the mutual information between images. Mutual information is the entropy of information that is mutually shared between two data. We evaluate how much the fused image expresses the mutual information of the source image by using the average mutual information. The gradient-based metric Q_g (also known as Q_{abf}) [28] is also used. Q_g measures the similarity of the magnitude and orientation of two image gradients.

3. Guided Filter Review

In this chapter, we briefly review the guided filter. The guided filter is an edge preserving-smoothing algorithm proposed by He et al. [21]. The guided filter replaces conventional filters such as the bilateral filter [22][23] and the joint-bilateral filter [24], and it has a faster computation speed. In addition, the guided filter can be utilized for soft matting, image enhancement, and so on. The guided filter fuses the intensity of the input image with the gradient of the guidance image. Equations (1), (2), and (3) are the key operations of the guided filter.

$$a_n = \frac{cov(I_k, p_k)}{var(I_k) + \varepsilon} \tag{1}$$

$$b_n = \bar{p}_n - a_n \bar{I}_n \tag{2}$$

$$q_n = \bar{a}_n I_n + \bar{b}_n \tag{3}$$

where q denotes a fused image, p denotes an input image, I denotes a guidance image, n denotes a pixel, and k denotes a corresponding window. $\overline{(\cdot)}$ is the local average of a window, $cov(\cdot)$ represents the covariance, and $var(\cdot)$ represents the variance. ε is a user parameter. In general, a very small value is used for ε. Equations (1) and (2) calculate a and b, and (3) defines how to obtain fused pixels. In other words, the guided filter sequentially calculates a_n, b_n, and q_n for each sliding window in the input image p and the guidance image I.

Mathematically, it is not easy to explain the guided filter. Therefore, we will attempt to do so intuitively. If the window of the guided filter is in a flat region, the window has a very small variance, and result q_n is the local average of the input image. Equations (4), (5), and (6) represent the case where the window of the guidance image is flat. If I is flat, i.e., $var(I_k) \ll \varepsilon$ and $cov(I_k, p_k) \ll \varepsilon$, a is equal to or similar to 0, b is equal to or similar to the local average of p_n, and the fused pixel q_n is equal to or similar to the local average of p_n. As a result, the guided filter facilitates smoothing if the window is in a flat region.

$$a_n \cong \frac{0}{\varepsilon} = 0 \tag{4}$$

$$b_n \cong \bar{p}_n \tag{5}$$

$$q_n = \bar{a}_n I_n + \bar{b}_n \approx \bar{\bar{p}}_n \tag{6}$$

Conversely, when the window is in an edge region, it has a very large variance value, and the resulting pixel represents the gradient of the guidance image. Therefore, the guided filter maintains the edge of the guidance image. Equations (7) and (8) explain the case where the window is on the edge. If $var(I_k) \gg \varepsilon$ and $cov(I_k, p_k) \gg \varepsilon$, then a_n is almost 1. As a result, the gradient of the resulting pixel becomes similar to the gradient of the guide image.

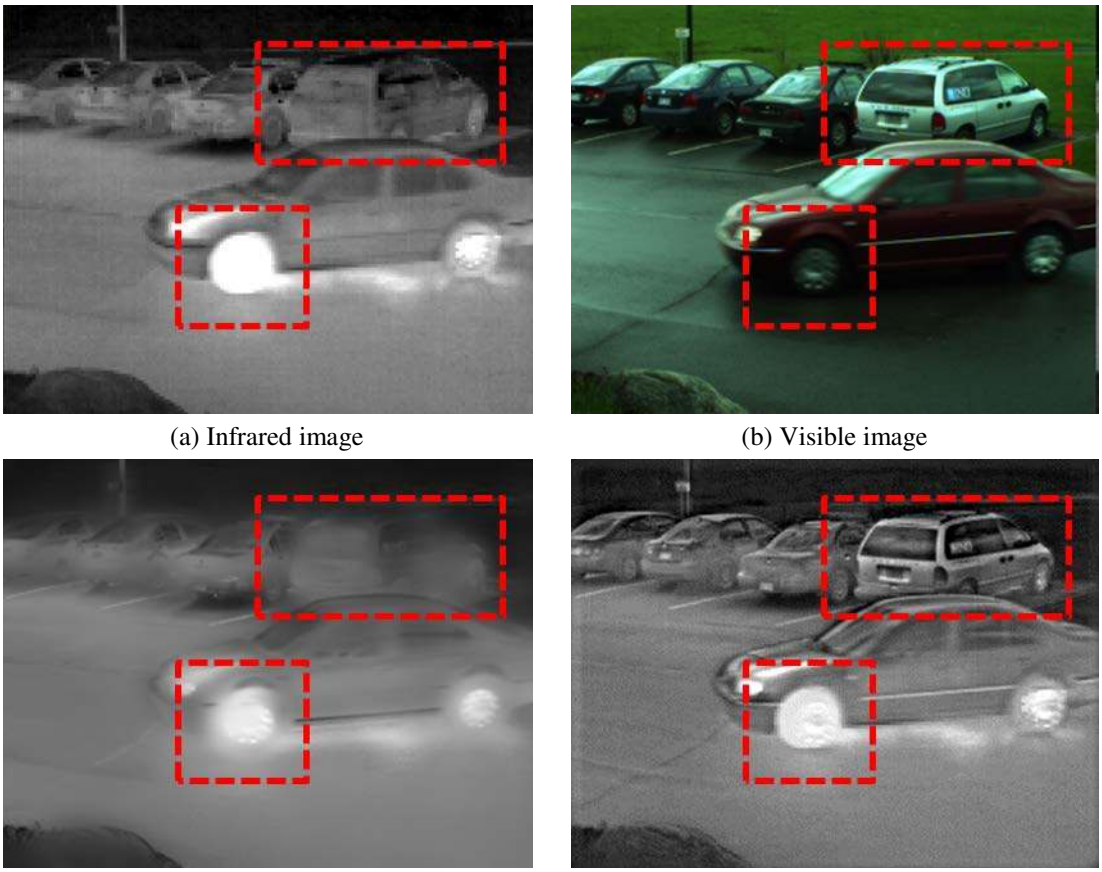
$$a_n \approx 1 \tag{7}$$

$$\nabla q_n = \bar{a} \nabla I_n \approx \nabla I_n \tag{8}$$

4. Use and Analysis of Guided Filter in Image Fusion

The guided filter transfers the gradient of the guidance image to the input image. Therefore, we expected that it would be applicable to image fusion. However, we observed critical defects when the guided filter was used for image fusion. **Fig. 2** shows the result of fusing images using the guided filter. We expected that the result of the guided filter would be similar to (d), but the guided filter did not work as expected. (c) is the result of the guided filter. It appears blurry, and the detail of the front wheel is not properly represented. This is because the wheel of the guidance image is dark and there are few gradients.

In addition, parked cars are blurred in (c). The edge of the parked cars is almost black in the infrared image and is not represented in result (c). Conversely, the moving car at the center of (b) is dark, and there is less edge. As a result, the complementary information of the two images is not represented in (c). On the other hand, the result of the proposed method (d) represents both the details of the infrared image and those of the visible image. The proposed method is explained in Chapter 5.



(a) Infrared image

(b) Visible image

(c) Fused result of guided filtering, input: (a), guidance: (b)

(d) Fused result of multi-guided filtering, input: (a), guidance: (a) and (b)

Fig. 2. Fused results of guided filtering and multi-guided filtering 1

We observed another problem in the guided filter related to texture transferrence. **Fig. 3** shows this problem. (c) has a detailed outline of a person. However, the texture of the bushes is not represented. This is because the guided filter determined that the texture is a noise to

remove it. This smoothing characteristic is a disadvantage for image fusion. On the other hand, in (d), the proposed method transfers the texture. The texture transfer of the proposed method is explained in Chapter 5.

(a) Infrared image (b) Visible image

(c) Fused result of guided filtering, input: (a), guidance: (b)

(d) Fused result of multi-guided filtering, input: (a), guidance: (a) and (b)

Fig. 3. Fused results of guided filtering and multi-guided filtering 2

5. Multi-Guided Filter

The proposed multi-guided filter is a modification of the guided filter. In this chapter, we analyze the guided filter and describe a process of converting it to the multi-guided filter.

Equation (3) is the key formula that describes the guided filter. In order to extend the multi-guided filter, we separate (3) into two components: a term related to the guidance image and the other terms. Equation (9) outlines the derivation process.

$$\begin{aligned} q_n &= \bar{a}_n I_n + \overline{(\bar{p}_n - a_n \bar{I}_n)} \\ &\approx \bar{a}_n I_n + \bar{\bar{p}}_n - \overline{(a_n \bar{I}_n)} \\ &= \bar{\bar{p}}_n + \{\bar{a}_n I_n - \overline{(a_n \bar{I}_n)}\} \end{aligned} \tag{9}$$

Strictly, $\overline{(\bar{p}_n - a_n\bar{I}_n)}$ in (9) is not $\bar{\bar{p}}_n - \overline{(a_n\bar{I}_n)}$. However, this does not present any practical issues during experiments. In (9), the first term $\bar{\bar{p}}_n$ is not related to the guidance image, whereas the second term $\{\bar{a}_n I_n - \overline{(a_n\bar{I}_n)}\}$ relates to the guidance image. Next, we convert the guidance image term to the multi-guidance images term.

$$q_n = \bar{\bar{p}}_n + \sum_m w^{(m)} \left\{ \overline{a^{(m)}}_n I^{(m)}{}_n - \overline{\left(a^{(m)}{}_n \overline{I^{(m)}}_n\right)} \right\} \tag{10}$$

Here, m represents the m-th guidance image. $w^{(m)}$ is the weight of the guidance image, which is a user parameter, and we set it to 1.0 during our experiments.

In order to fuse the images, we modify the guided filter into the multi-guided filter based on (10). However, we identified a problem where texture is not represented in the fused image. **Fig. 3** highlights this problem. The guidance image **Fig. 3 (b)** has the texture of bushes, but this does not appear in **Fig. 3 (c)**. It is noted that this problem is caused by the characteristics of the guided filter. If the input image p is flat, $cov(I_k, p_k)$ in (1) is close to 0, and then the texture of I is not properly represented. To solve this problem, we changed $cov(I_k, p_k)$ of (1) to $var(I_k)$. Equation (11) is a modification of (1). As a result, our method can represent textures properly.

$$a^{(m)}{}_n = \frac{var\left(I^{(m)}{}_k\right)}{var(I^{(m)}{}_k) + \varepsilon} \tag{11}$$

Finally, the proposed method synthesizes two or more images using the multi-guided filtering of (10) and (11). This filter is a variation of the guided filter, which is a fast algorithm. Therefore, it is suitable for real-time processing. **Fig. 1** shows the process of image fusion using the multi-guided filter. Since the conventional guided filter uses only one guidance image, the information of the visible or infrared image can be lost. However, the proposed multi-guided filter can use two guidance images. Therefore, it is possible to utilize all information of the visible-infrared images. In addition, it has a fast processing speed similar to the guided filter. In Chapter 6, we verify that the proposed method is suitable for real-time image fusion by comparing computation times with the other methods.

6. Experimental Results

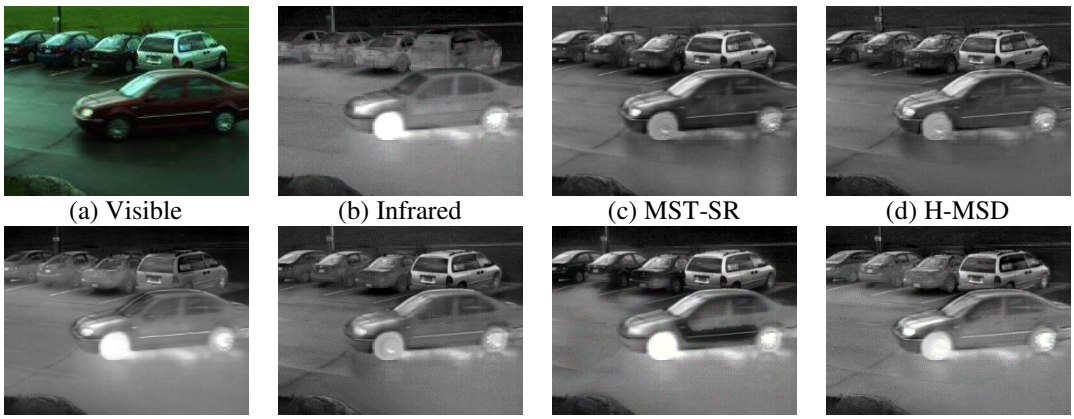
6.1 Experimental Setting

In this chapter, we compare the proposed method with conventional methods in terms of quality, quantity, fusing speed, and flickering artifacts. The conventional methods include MST-SR [13], H-MSD [9], GTF [17], GFF [11], and CNNF [18]. We used dataset from [29], [30], and [31]. The quantitative matrices are the normalized mutual information (NMI) [27] and the gradient-based metric Q_g [28]. We used MATLAB code provided by the authors. The fusing speed was measured using an Intel i5-6600 3.3 GHz CPU. In this experiment, we set the parameters of the proposed method as follows: weight of visible = 1.0, weight of infrared = 1.0, window size = 5, and $\varepsilon = 0.001^2$.

6.2 Qualitative Evaluation

We compare the proposed method with the conventional methods. **Fig. 4, 5, and 6** show the fused results of the visible and infrared images.

Fig. 4 shows a moving car and parked cars. In (b), the front wheel of the car is bright at the center. This indicates that the wheel is hot. This is evidence that the car is moving or recently moved. Thus, the fused image must represent the thermal properties of the wheel. (c) MST-SR, (d) H-MSD, and (f) GFF fail to represent the thermal information of the wheel. (g) CNNF renders an incorrect backline on the center of the car. This is a critical defect. (e) GTF correctly represents the thermal properties but fails to represent the detail of the front wheel. However, the proposed method represents this information correctly and has superior detail.

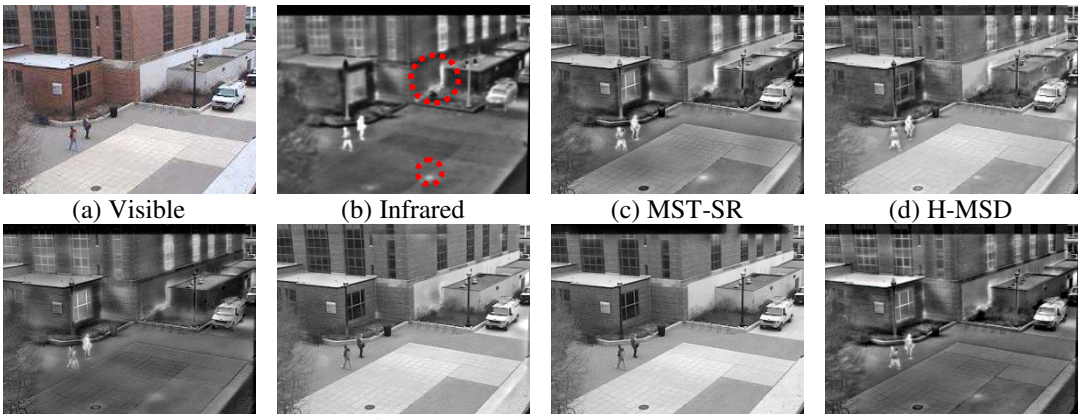


(a) Visible (b) Infrared (c) MST-SR (d) H-MSD

(e) GTF (f) GFF (g) CNNF (h) Proposed

Fig. 4. Fused result 1

In **Fig. 5**, we highlight important points with red circles. With the visible image only, we cannot determine that are thermal sources. Thus, it is important to highlight these regions in **Fig. 5**. (d) H-MSD, (f) GFF, and (g) CNNF appear similar to the visible image, and these methods fail to identify the thermal sources. The proposed method identifies the marked points and the detail of the image correctly.



(a) Visible (b) Infrared (c) MST-SR (d) H-MSD

(e) GTF (f) GFF (g) CNNF (h) Proposed

Fig. 5. Fused result 2

Fig. 6 shows a man and bushes. In (a), a man is not shown, but he does appear in (b). In this case, the objective of the image fusion is to show this man and the detail of the surrounding bush simultaneously. (g) CNNF makes a critical error, and the man is not visible. The other methods represent this man correctly, but the detail of the bush is degraded. The proposed method correctly represents the man and the background details.

In this experiment, conventional methods often failed to represent important information. However, the fused image of the proposed method includes important information and details without creating critical artifacts.

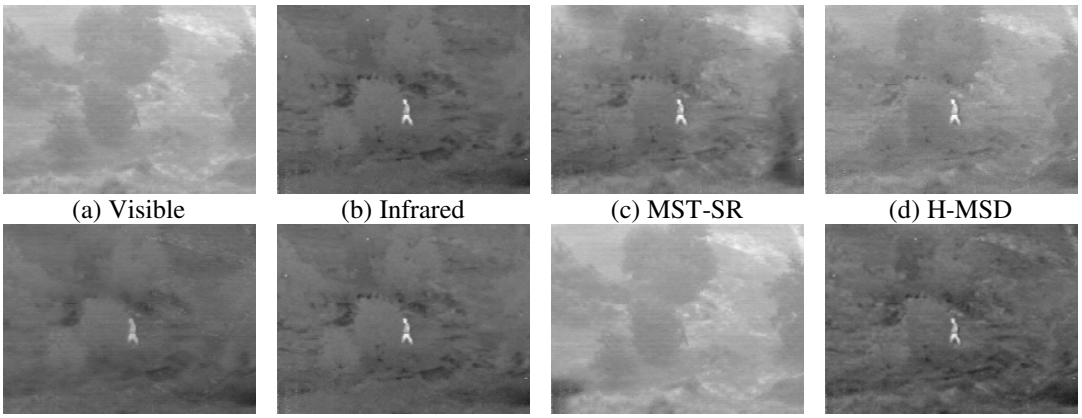


Fig. 6. Fused result 3

6.3 Running Time Evaluation

The biggest advantage of the proposed method is its running speed. **Table 1** lists the running times and speeds of the other methods. The input image sizes were 320×270, 640×480, 1280×720, and 520 × 520.

In this experiment, we added another neural network based method named FusionGAN (F-GAN) [20]. F-GAN is a state-the-of-art method that uses a generative adversarial neural network for visible and infrared image fusion. We implemented F-GAN according to [20].

We measured the speed using MATLAB code provided by the authors of MST-SR, H-MSD, GTF, GFF, and CNNF (CPU). CNNF (GPU) results were referred from [18]. F-GAN was measured using our implementation with TensorFlow and GTX 1080ti. The proposed method was written in MATLAB.

Table 1 shows the averages of the image fusing speeds. We measured only the computation time excluding the I/O time. The fusing speed of the proposed method for the 320 × 270 size image is 57.9 fps. This is faster than 30 fps, and thus the proposed image fusion method can be utilized for real-time processing. The second fastest algorithm was F-GAN, which has a rate of 50.07 fps. However, this method has a much slower speed without GPU at a rate of 0.2 fps. The third fastest algorithm was GFF, which is 3.4 times slower than the proposed method and is not suitable for real-time processing.

Table 1. Running speed evaluation

	Size	MST-SR	H-MSD	GTF	GFF	CNNF (CPU)	CNNF (GPU)	F-GAN (CPU)	F-GAN (GPU)	Proposed
Average running speed (FPS)	**320 × 270**	5.27	2.05	0.66	17.10	0.03	-	0.20	50.07	**57.93**
	640 × 480	1.27	0.70	0.08	3.48	0.01	-	0.06	10.04	**17.51**
	1280 × 720	0.40	0.24	0.02	1.16	< 0.01	-	0.02	4.03	**7.07**
	520 × 520	1.48	0.81	0.12	4.21	0.01	3.03	0.06	13.97	**21.38**

The reason for the notable difference in speed between the conventional methods and the proposed method is that the proposed method uses a simple algorithm. MST-SR and H-MSD are methods to decompose images and recompose them based on specific rules. Therefore, it takes a significant amount of time to decompose into multiple images and to reconstruct them. GTF uses iterative optimization techniques. GFF also uses the guided filter, but it involves decomposing and recomposing images. CNNF is a deep artificial neural network based method, which requires a significant amount of time.

For the 640×480 and 1280×720 images, the proposed method is the fastest fusing method. In order to compare this with the GPU performance of CNNF, a 520×520 size image was also tested. The GPU performance of CNNF was obtained from [18]. The proposed method had a rate of 21.38 fps, which is faster than that of the CNNF using GPU. F-GAN demonstrated 50.06 fps at 320×270 using GPU. Thus, F-GAN is suitable for real-time processing with GPU. However, it can process at 0.2 fps without GPU, which is not sufficient for real-time processing.

The proposed method is much faster at a frame rate over 30 fps. This is because the multi-guided filter is a modification of the guided filter, which is a fast algorithm. F-GAN is suitable for real-time processing, but without GPU, F-GAN is slower than conventional methods. By experimentation, we have shown that our method is able to perform real-time image fusion.

6.4 Quantitative Evaluation

Table 2 shows the average quantitative results for three sequences: duine, nato, and tree.

Table 2. Quantitative evaluation of test sequences

	MST-SR	H-MSD	GTF	GFF	CNNF	Proposed
NMI	1.01	1.03	1.04	**1.22**	1.21	1.06
Q_g	0.51	0.45	0.38	**0.56**	0.55	0.43

In the case of quantitative evaluation, GFF and CNNF show superior results. MST-SR, H-MSD, GTF, and the proposed method show similar performances. However, the quantitative evaluation is not important for surveillance systems. CNNF's quantitative performance is good, but CNNF introduces a critical artifact in the previous qualitative evaluation. The proposed method is not superior with regard to quantitative evaluation, but considering the qualitative evaluation, it yields competitive results for surveillance systems. In particular, the proposed method is sufficiently useful in terms of its high speed.

6.5 Flickering Artifact

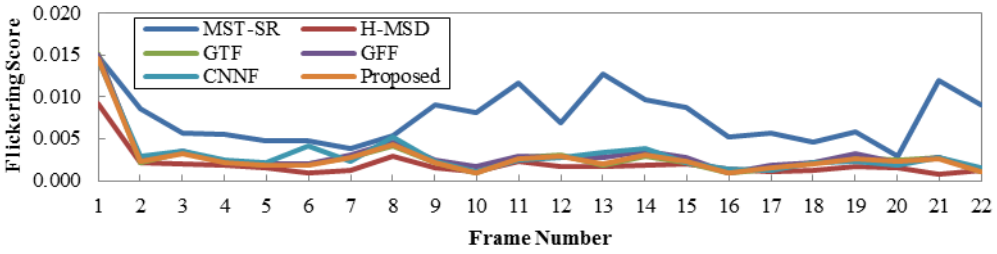
Typically, surveillance systems require consecutive image processing for real-time monitoring. Flickering artifacts often appear during consecutive image processing. **Fig. 7** shows the consecutive image fusion result of CNNF and the proposed method. **Fig. 7 (a)** shows a flickering artifact. In the region marked with a red rectangle, it is bright at frame # 21, but is suddenly darker at frame # 22. The rightmost figure is an overlapped image of frames # 22 and # 23. At the marked region, CNNF has a noticeable brightness change. This artifact causes flickering in the real-time monitoring system. By contrast, the proposed method did not generate flickering artifacts.

Fig. 7. Example of flickering artifact

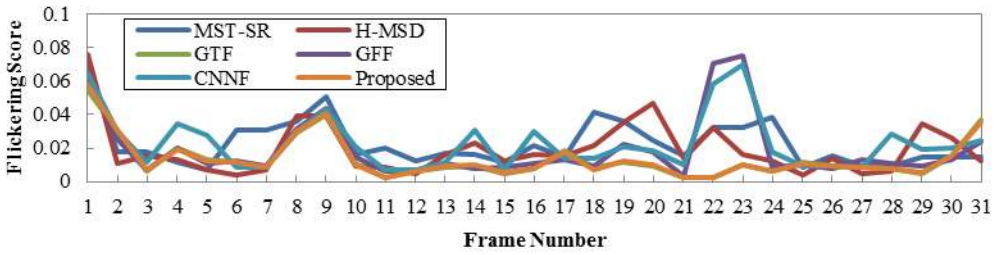
In order to measure the flickering artifacts, we measure the luminance change using (12):

$$\text{Flickering Score} = \frac{1}{MN}\sum_{m=0}^{M}\sum_{n=0}^{N}\frac{\left|l_{m,n}^{(t)} - l_{m,n}^{(t+1)}\right|}{min(l_{m,n}^{(t)}, l_{m,n}^{(t+1)})} \tag{12}$$

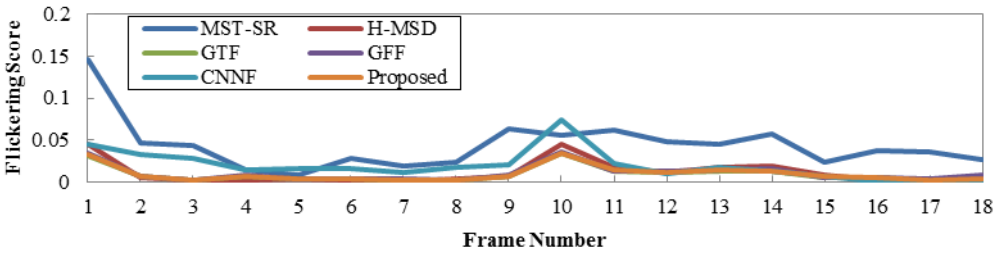
where m,n is a block and a pixel, t is the frame number, and l represents the luminance. At each block, we measure the ratio of the luminance changes and compute the average. We use the non-overlapped 64 × 64 size block. The flickering score is higher if the consecutive images exhibit flickering. **Fig. 8** shows the flickering scores of three sequences. In **Fig. 8**, the proposed method achieves low flickering scores for all three sequences. In (a), the other methods except MST-SR show low flickering scores. In frame # 9 of (b), every method shows high scores because the luminance of the source frame was changed, but frame # 23 does not. In (b), it is noted that the proposed method has the lowest flickering score at frame # 23. The other methods exhibited flickering. In (c), the proposed method shows a low flickering score.



(a) Duine



(b) Nato



(c) Tree

Fig. 8. Graph of flickering score

7. Conclusion

Image fusion is a process of synthesizing two or more images into one image. The image fusion of visible-infrared images is especially useful because the visible and infrared images contain complementary information. Many methods have been proposed for visible-infrared image fusion with good success. However, conventional methods are not suitable for real-time processing because of the complexity of the implemented algorithms. To address the speed problem, we proposed a multi-guided filter and a real-time image fusion method based on this filter.

The proposed method combines visible-infrared images without critical artifacts in terms of image quality and flickering. In particular, the proposed method demonstrated a fast fusing speed of 57.93 fps, which is sufficient for real-time image processing. The proposed method has room for further improvement such as automated parameter tuning. Further research will include an adaptive algorithm and deep neural networked based algorithm.

References

[1] P. Burt and E. Adelson, "The Laplacian Pyramid as a Compact Image Code," *IEEE Transactions on Communications*, vol. 31, no. 4, pp. 532-540, 1983. Article (CrossRef Link).

[2] A. Toet, L. J. van Ruyven, and J. M. Valeton, "Merging Thermal And Visual Images By A Contrast Pyramid," *Optical Engineering*, vol. 28, no. 7, 1989. Article (CrossRef Link).

[3] A. Toet, "Image Fusion by a Ratio of Low-Pass Pyramid," *Pattern Recognition Letters*, vol. 9, no. 4, pp. 245-253, 1989. Article (CrossRef Link).

[4] H. Li, B. S. Manjunath, and S. K. Mitra, "Multisensor Image Fusion Using the Wavelet Transform," *Graphical Models and Image Processing*, vol. 57, no. 3, pp. 235-245, 1995. Article (CrossRef Link).

[5] J. J. Lewis, R. J. O'Callaghan, S. G. Nikolov, D. R. Bull, and N. Canagarajah, "Pixel-and region-based image fusion with complex wavelets," *Information Fusion*, vol. 8, no. 2, pp. 119-130, 2007. Article (CrossRef Link).

[6] B. Jin, Z. Jing, and H. Pan, "Multi-modality image fusion via generalized resize-wavelet transformation," *KSII Transactions on Internet and Information Systems*, vol. 8, no. 11, pp. 4118-4136, 2014. Article (CrossRef Link).

[7] F. Nencini, A. Garzelli, S. Baronti, and L. Alparone, "Remote Sensing Image Fusion using the Curvelet Transform," *Information Fusion*, vol. 8, no. 2, pp. 143-156, 2007. Article (CrossRef Link).

[8] V.P.S. Naidu, "Image fusion technique using multi-resolution singular value decomposition," *Defence Science Journal*, vol. 61, no. 5, pp. 479-484, 2011. Article (CrossRef Link).

[9] Z. Zhou, B. Wang, S. Li, and M. Dong, "Perceptual Fusion of Infrared and Visible Images through a Hybrid Multi-Scale Decomposition with Gaussian and Bilateral Filters," *Information Fusion*, vol. 30, pp. 15-26, 2016. Article (CrossRef Link).

[10] S. Li, X. Kang, and J. Hu, "Image Fusion With Guided Filtering," *IEEE Transactions on Image Processing*, vol. 22, no. 7, pp. 2864-2875, 2013. Article (CrossRef Link).

[11] W. Gan, X. Wu, W. Wu, X. Yang, C. Ren, X. He, and K. Liu, "Infrared and Visible Image Fusion with the Use of Multi-Scale Edge-Preserving Decomposition and Guided Image Filter," *Infrared Physics and Technology*, vol. 72, pp. 37-51, 2015. Article (CrossRef Link).

[12] J. Zhu, W. Jin, L. Li, Z. Han, and X. Wang, "Multiscale infrared and visible image fusion using gradient domain guided image filtering," *Infrared Physics & Technology*, vol. 89, pp. 8-19, 2017. Article (CrossRef Link).

[13] Y. Liu, S. Liu, and Z. Wang, "A General Framework for Image Fusion based on Multi-Scale Transform and Sparse Representation," *Information Fusion*, vol. 24, pp. 147-164, 2015. Article (CrossRef Link).

[14] D. P. Bavirisetti, G. Xiao, and G. Liu, "Multi-sensor image fusion based on fourth order partial differential equations," in *Proc. of International Conference on Information Fusion*, 2017. Article (CrossRef Link).

[15] J. Ma, Z. Zhou, B. Wang, and H. Zong, "Infrared and visible image fusion based on visual saliency map and weighted least square optimization," *Infrared Physics & Technology*, vol. 82, pp. 8-17, 2017. Article (CrossRef Link).

[16] X. Zhang, Y. Ma, F. Fan, Y. Zhang, and J. Huang "Infrared and visible image fusion via saliency analysis and local edge-preserving multi-scale decomposition," *Journal of the Optical Society of America A*, vol. 34, no. 8, pp. 1400-1410, 2017. Article (CrossRef Link).

[17] J. Ma, C. Chen, C. Li, and J. Huang, "Infrared and visible image fusion via gradient transfer and total variation minimization," *Information Fusion*, vol. 31, pp. 100-109, 2016. Article (CrossRef Link).

[18] Y. Liu, X. Chen, H. Peng, and Z. Wang, "Multi-focus image fusion with a deep convolutional neural network," *Information Fusion*, vol. 36, pp. 191-207, 2017. Article (CrossRef Link).

[19] K. Xu, Z. Qin, G. Wang, H. Zhang, K. Huang, and S. Ye, "Multi-focus Image Fusion using Fully Convolutional Two-stream Network for Visual Sensors," *KSII Transactions on Internet and Information Systems*, vol. 12, no. 5, pp. 2253-2272, 2018. Article (CrossRef Link).

[20] J. Ma, W. Yu, P. Liang, C. Li, and J. Jiang, "FusionGAN: A generative adversarial network for infrared and visible image fusion," *Information Fusion*, vol. 48, pp. 11-26, 2019. Article (CrossRef Link).
[21] K. He, J. Sun, and X. Tang, "Guided Image Filtering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 6, pp. 1397-1409, 2013. Article (CrossRef Link).
[22] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proc. of Sixth International Conference on Computer Vision*, pp. 839-846, 1998. Article (CrossRef Link).
[23] F. Banterle, M. Corsini, P. Cignoni, and R. Scopigno, "A Low-Memory, Straightforward and Fast Bilateral Filter Through Subsampling in Spatial Domain," in *Proc. of Computer Graphics Forum*, vol. 31, no. 1, pp. 19-32, 2012. Article (CrossRef Link).
[24] G. Petschnigg, R. Szeliski, M. Agrawala, M. Cohen, H. Hoppe, and K. Toyama, "Digital photography with flash and no-flash image pairs," *ACM Transactions on Graphics (TOG)*, vol. 23, no. 3, pp. 664-672, 2004. Article (CrossRef Link).
[25] F. Kou, W. Chen, C. Wen, and Z. Li "Gradient Domain Guided Image Filtering," *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 4528-4539, 2015.
[26] G. Qu, D. Zhang, and P. Yan, "Information measure for performance of image fusion," *Electronics Letters*, vol. 38, no. 7, pp. 313-315, 2002. Article (CrossRef Link).
[27] N. Cvejic, C. N. Canagarajah, and D. R. Bull, "Image fusion metric based on mutual information and Tsallis entropy," *Electronics Letters*, vol. 42, no. 11, pp. 626-627, 2006. Article (CrossRef Link).
[28] C. S. Xydeas and V. Petrovic, "Objective image fusion performance measure," *Electronics Letters*, vol. 36, no. 4, pp. 308-309, 2000. Article (CrossRef Link).
[29] A. Toet, J.K. Ijspeert, A.M. Waxman, and M. Aguilar, "Fusion of visible and thermal imagery improves situational awareness," *Displays*, vol. 18, no. 2, pp. 85-95, 1997. Article (CrossRef Link).
[30] J. W. Davis and V. Sharma, "Background-subtraction using contour-based fusion of thermal and visible imagery," *Computer Vision and Image Understanding*, vol. 106, no. 2-3, pp. 162-182, 2007. Article (CrossRef Link).
[31] J. J. Lewis, S. G. Nikolov, A. Loza, E. F. Canga, N. Cvejic, J. Li, A. Cardinali, C. N. Canagarajah, D. R. Bull, T. Riley, D. Hickman, and M. I. Smith. "The Eden Project multi-sensor data set," *Technical report TR-UoB-WS-Eden-Project-Data-Set, University of Bristol and Waterfall Solutions Ltd*, 2006.

Woojin Jeong received the B.S degree in the Department of Computer Science and Engineering from Hanyang University, Korea, in 2012. He is currently working towards PhD. Degree at the Department of Computer Science and Engineering from Hanyang University, Korea, From 2012. His research interests include computer vision and machine learning.
Email : wjjeong@visionlab.or.kr

Bok Gyu Han received his B.S. degree in the Department of Computer Science and Engineering from Hallym University, Korea, in 2016. He is currently working towards PhD. Degree at the Department of Computer Science and Engineering from Hanyang University, Korea, From 2016. His research interests include computer vision and machine learning.
Email : bghan@visionlab.or.kr

Hyeon Seok Yang received his B.S. degree in the Department of Electronics and Information Engineering from Yeungnam University, Korea, in 2010. He received the M.S. degrees in the Department of Computer Science & Engineering from Hanyang University, Korea, in 2012. He is studying for his PhD. degree in the Department of Computer Science & Engineering from Hanyang University, Korea. His research interests include computer vision, pattern recognition, and deep learning.
Email : hsyang@visionlab.or.kr

Young Shik Moon received the B.S. and M.S. degrees in Electronics Engineering from Seoul National University and Korea Advanced Institute of Science and Technology, Korea, in 1980 and 1982, respectively, and PhD. degree in Electrical and Computer Engineering from the University of California at Irvine, CA, in 1990. From 1982 to 1985, he had been a researcher at the Electronics and Telecommunication Research Institute, Daejon, Korea. In 1992, he joined the Department of Computer Science and Engineering at Hanyang University, Korea, as an Assistant Professor, and is currently a Professor. Dr. Moon served as General Chair of 2014 IEEE International Symposium on Consumer Electronics, and worked as the President of the Institute of Electronics and Information Engineer, Korea.
Email : ysmoon@hanyang.ac.kr