SELECTIVELY DETAIL-ENHANCED EXPOSURE FUSION VIA A GRADIENT DOMAIN CONTENT ADAPTIVE BILATERAL FILTER

Z. G. Li and J. H. Zheng

Signal Processing Department, Institute for Infocomm Research, 1 Fusionopolis Way, Singapore

ABSTRACT

Bilateral filters suffer from halo artifacts when they are applied for image enhancement. In this paper, a new bilateral filter is proposed in gradient domain to address this problem. Both spatial similarity parameter and intensity similarity parameter of the proposed filter are spatially varying instead of being fixed as in the existing bilateral filters. As a result, it can preserve edges and smooth flat areas better than the existing bilateral filters. The proposed filter is then adopted to design a selectively detail-enhanced exposure fusion algorithm. Fine details of multiple differently exposed images are extracted simultaneously using the proposed filter. Instead of amplifying and adding all extracted fine details to an intermediate image which is fused by an existing exposure fusion algorithm, the fine details in all areas except flat ones are amplified and added to the intermediate image. The resultant algorithm can reduce halo artifacts and prevent noise in flat areas from being amplified in the final image. Therefore, the proposed algorithm fuses images with much better visual quality.

1. INTRODUCTION

The bilateral filter is perhaps one of the most popular and simplest edge-preserving smoothing local filters [1]. The output of the bilateral filter at a pixel is a weighted average of its nearby pixels. There are two parameters with fixed values in the bilateral filter and their function is to adjust the sensitivity of the bilateral filter to the spatial similarity and the range sensitivity respectively. This filter is widely used in many applications including tone mapping of high dynamic range (HDR) images [2], multi-scale detail decomposition [3], image abstraction [4], etc. Its extension is also a popular research topic, such as the bilateral filter in gradient domain [5], trilateral filter [6], and their accelerated versions [3, 7].

The bilateral filters suffer from halo artifacts when they are applied for image enhancement [8, 9]. To overcome this problem, two content adaptive bilateral filters were introduced in [10, 11]. Range similarity parameter is adaptive to the content of filtered image in [10] while both spatial similarity parameter and range similarity parameter are adaptive to the content of filtered image in [11]. Both content adaptive bilateral filters preserve sharp edges and smoothes flat areas better than the existing bilateral filters in [1, 2, 3, 5]. However, as pointed out in [10], adaptation of the parameters will destroy the 3-D convolution form, and the content adaptive bilateral filters cannot be accelerated via the approach in [7]. It is time consuming to extract fine details from a set of differently exposed images by the content adaptive bilateral filters in [10, 11] because each input image needs to be decomposed individually.

Inspired by the gradient domain bilateral filter in [5], the content adaptive bilateral filter in [11] and the joint bilateral filter in [13, 14], a content adaptive bilateral filter in gradient domain is first proposed in this paper. Both spatial range parameter and range similarity parameter are spatially varying instead of being fixed in [5]. Their values are larger at a pixel position if the pixel is at an edge and smaller if it is in a flat area. As a result, the proposed filter can preserve edges and smooth flat areas better than the filter in [5]. Since the new filter can exact fine details from a set of images simultaneously, it is applied for detail enhancement of multiple differently exposed images. It is worth noting that all extracted fine details are enhanced by the detail-enhanced exposure fusion algorithm in [12]. Unfortunately, noise is also amplified when fine details are enhanced. The human visual system (HVS) can tolerate amplified noise in complex regions but is particularly sensitive to amplified noise in flat areas. In addition, it is an open problem to separate noise from fine details. To reduce amplification of noise which is inherent in the fully detail-enhanced exposure fusion algorithm, a selectively detail-enhanced exposure fusion algorithm is introduced to enhance fine details in all regions except those in flat ones. Experimental results show that the proposed selective detail enhancement algorithm can significantly reduce halo artifacts from appearing in the final images and improves visual quality of enhanced images produced by the full detail enhancement algorithm.

The rest of this paper is organized as follows. Existing bilateral filters are summarized in Section 2. A content adaptive bilateral filter in gradient domain is proposed in Sections 3. It is applied to design a selectively detail-enhanced exposure fusion algorithm in Section 4. Experimental results are given in Section 5 to illustrate the efficiency of the proposed exposure fusion algorithm. Concluding remarks are provided in Section 6.

2. BACKGROUND AND MOTIVATION

The task of edge-preserving smoothing is to decompose an image to be filtered X into two parts as follows:

$$X(p) = \hat{Z}(p) + e(p), \tag{1}$$

where $p(= (x, y))$ is a pixel position, $\hat{Z}$ is a reconstructed image formed by homogeneous regions with sharp edges and e is fine detail or noise.

Edge-preserving smoothing can be achieved by using the bilateral filter as [1]

$$\hat{Z}(p) = \sum_{p' \in \Omega_{\zeta_1}(p)} W_{p,p'}(X, \sigma_1, \sigma_2) X(p'), \tag{2}$$

where $\Omega_{\zeta_1}(p)$ is a square window centered at the pixel p of a radius

ζ_1. The weighting function $W_{p,p'}(X,\sigma_1,\sigma_2)$ is defined as

$$W_{p,p'}(X,\sigma_1,\sigma_2) = \frac{\exp^{\frac{-(p'-p)^2}{\sigma_1^2}} \exp^{\frac{-(X(p')-X(p))^2}{\sigma_2^2}}}{c_p(X,\sigma_1,\sigma_2)}, \quad (3)$$

$$c_p(X,\sigma_1,\sigma_2) = \sum_{p'\in\Omega_{\zeta_1}(p)} \exp^{\frac{-(p'-p)^2}{\sigma_1^2}} \exp^{\frac{-(X(p')-X(p))^2}{\sigma_2^2}}, (4)$$

and σ_1 and σ_2 are two constants which adjust the sensitivity of the spatial similarity and the range similarity respectively.

By enlarging the values of σ_1 and σ_2, the whole image is smoothed well while edges might not be preserved well. By reducing their values, edges are preserved well while the flat regions are not smoothed well. It is thus very challenging to select the values of σ_1 and σ_2 to simultaneously preserve edges and smooth flat regions. A sophisticated method was introduced in [13, 14] in presence of a pair of flash and non-flash images. The flash image I captures details of a scene while the non-flash one X captures ambient illumination. The flash image contains less noise than the non-flash, and it serves as the guidance image. A joint bilateral filter can be designed to reduce the noise from the non-flash image as follows:

$$\hat{Z}(p) = \sum_{p'\in\Omega_{\zeta_1}(p)} W_{p,p'}(I,\sigma_1,\sigma_2)X(p'). \quad (5)$$

It is worth noting that the weight of $X(p')$ is $W_{p,p'}(I,\sigma_1,\sigma_2)$ rather than $W_{p,p'}(X,\sigma_1,\sigma_2)$. It is thus called a joint bilateral filter.

As pointed out in [8, 9], the bilateral filter produces halo artifacts and/or gradient reversal artifacts around some edges due to unwanted smoothing of these edges. To reduce halo artifacts, a content adaptive bilateral filter was proposed in [11] as follows:

$$\hat{Z}(p) = \sum_{p'\in\Omega_{\zeta_1}(p)} W_{p,p'}(X,\tilde{\sigma}_{1,X}(p),\tilde{\sigma}_{2,X}(p))X(p'), \quad (6)$$

where the values of $\tilde{\sigma}_{i,X}(p)(i=1,2)$ are computed as

$$\tilde{\sigma}_{i,X}(p) = \frac{\sigma_i}{\sqrt{w_X(p)}},$$

and the value of $w_X(p)$ is computed by using local variance of input image X as follows:

$$w_X(p) = \frac{1}{N}\sum_{p'=1}^{N} \frac{\sigma^2_{X,\zeta_2}(p)+1}{\sigma^2_{X,\zeta_2}(p')+1}, \quad (7)$$

$\sigma^2_{X,\zeta_2}(p)$ is the variance of a block centered at pixel p with the radius as ζ_2, i.e., $\Omega_{\zeta_2}(p)$. $\zeta_2(\geq \zeta_1)$ is a constant, and N is the total number of pixels in an image.

Larger $\tilde{\sigma}_{1,X}(p)$ and $\tilde{\sigma}_{2,X}(p)$ are adopted for the pixel p if it is in a smooth region and smaller $\tilde{\sigma}_{1,X}(p)$ and $\tilde{\sigma}_{2,X}(p)$ are used for the pixel p in a complex region. As such, the filter (6) preserves edges and smoothes flat areas better than the filter (2). However, it is time consuming to extract fine details from multiple input images using the filter (6) because each input image needs to be decomposed individually. In the following section, a content adaptive bilateral filter will be proposed in gradient domain and the filter can be applied to extract fine details from multiple input images simultaneously.

3. CONTENT ADAPTIVE BILATERAL FILTERING IN GRADIENT DOMAIN

In this section, we propose a content adaptive bilateral filter in gradient domain. Inspired by the content adaptive bilateral filter (6), a content adaptive joint bilateral filter is proposed as

$$\hat{Z}(p) = \sum_{p'\in\Omega_{\zeta_1}(p)} W_{p,p'}(I,\tilde{\sigma}_{1,I}(p),\tilde{\sigma}_{2,I}(p))X(p'). \quad (8)$$

It can be derived from Equations (1) and (8) that

$$e(p) = \sum_{p'\in\Omega_{\zeta_1}(p)} W_{p,p'}(I,\tilde{\sigma}_{1,I}(p),\tilde{\sigma}_{2,I}(p))(X(p)-X(p')).$$

A vector field $\vec{G}(\nabla X,p,p')$ is defined for two pixels $p(=(x,y))$ and $p'(=(x',y'))$ in image X as [5]

$$G_x(\nabla X,p,p') = \begin{cases} \sum_{r=x'}^{x-1} \frac{\partial X(r,y)}{\partial x}; & \text{if } x \geq x' \\ -\sum_{r=x}^{x'-1} \frac{\partial X(r,y)}{\partial x}; & \text{otherwise} \end{cases},$$

$$G_y(\nabla X,p,p') = \begin{cases} \sum_{r=y'}^{y-1} \frac{\partial X(x',r)}{\partial y}; & \text{if } y \geq y' \\ -\sum_{r=y}^{y'-1} \frac{\partial X(x',r)}{\partial y}; & \text{otherwise} \end{cases}.$$

Similarly, a vector field $\vec{G}(\nabla I,p,p')$ can be defined for two pixels $p(=(x,y))$ and $p'(=(x',y'))$ in the image I. Both definitions are based on the four possible relationships between two pixels $p(=(x,y))$ and $p'(=(x',y'))$ as shown in Fig. 1.

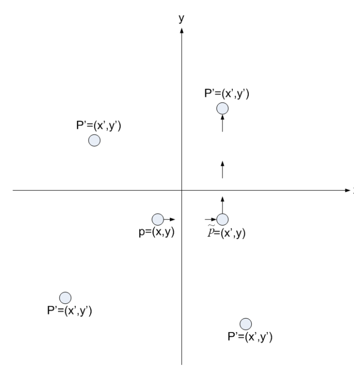


Fig. 1. Four possible relationships between two pixels $p(=(x,y))$ and $p'(=(x',y'))$.

Let $\tilde{p}=(x',y)$ as illustrated in Fig. 1 and define $\frac{\partial X(x,y)}{\partial x}$ and $\frac{\partial X(x,y)}{\partial y}$ as $(X(x+1,y)-X(x,y))$ and $(X(x,y+1)-X(x,y))$, respectively. It can be derived that

$$\begin{aligned} X(p)-X(p') &= X(p)-X(\tilde{p})+X(\tilde{p})-X(p') \\ &= G_x(\nabla X,p,p')+G_y(\nabla X,p,p'). \end{aligned}$$

Similarly, it can be derived that

$$I(p)-I(p')=G_x(\nabla I,p,p')+G_y(\nabla I,p,p').$$

Therefore, $e(p)$ can also be computed as

$$e(p)=\frac{\sum\limits_{p'\in\Omega_{\zeta_1}(p)}[\chi_w(\nabla I,p,p')\sum\limits_{q\in\{x,y\}}G_q(\nabla X,p,p')]}{\sum\limits_{p'\in\Omega_{\zeta_1}(p)}\chi_w(\nabla I,p,p')}, \quad (9)$$

where $\chi_w(\nabla I,p,p')$ is defined as

$$\chi_w=\exp(-\frac{(p'-p)^2}{\sigma_1^2/w_g(p)})\exp(-\frac{(\sum\limits_{q\in\{x,y\}}G_q(\nabla I,p,p'))^2}{\sigma_2^2/w_g(p)}). \quad (10)$$

Replacing the gradient fields ∇I and ∇X by two general vector fields $\vec{U}=(U_x,U_y)$ and $\vec{V}=(V_x,V_y)$ respectively, a content adaptive bilateral filter in gradient domain is then obtained as

$$e(p)=\frac{\sum\limits_{p'\in\Omega_{\zeta_1}(p)}[\chi_w(\vec{U},p,p')\sum\limits_{q\in\{x,y\}}G_q(\vec{V},p,p')]}{\sum\limits_{p'\in\Omega_{\zeta_1}(p)}\chi_w(\vec{U},p,p')}. \quad (11)$$

The filter (8) and the proposed filter (11) are equivalent if $w_g(p)$, $\vec{V}$ and $\vec{U}$ are selected as $w_I(p)$, ∇X and ∇I, respectively. The filter (6) and the proposed filter (11) are equivalent if $w_g(p)$, $\vec{V}$ and $\vec{U}$ are selected as $w_X(p)$, ∇X and ∇X, respectively. Since the proposed filter (11) is applicable for other choices of $w_g(p)$, $\vec{V}$ and $\vec{U}$, both the filter (6) and the filter (8) are special cases of the proposed filter (11).

It is worth noting that three key components in the proposed bilateral filter (11) are the vector field $\vec{V}=(V_x,V_y)$, the vector field $\vec{U}=(U_x,U_y)$ and the weighting function $w_g(p)$. In the next section, fusion of differently exposed images [15] is taken as an example to illustrate how to determine them.

4. SELECTIVELY DETAIL-ENHANCED FUSION OF DIFFERENTLY EXPOSED IMAGES

Let input images be denoted as $X_k(1\leq k\leq L)$ and their luminance components denoted as $Y_k(1\leq k\leq L)$. The gradient field of $Y_k(p)(1\leq k\leq L)$ is denoted as $(\nabla Y_{k,x}(p),\nabla Y_{k,y}(p))$ [12]. Normally, the gradient of a pixel with the largest absolute value among different exposures corresponds to the most desirable detail at a position. However, there is a high likelihood that the maximum gradient is noisy, especially in dark regions of an HDR scene. A vector field is thus constructed by using a weighted average of gradients over all exposures.

As indicated in [12, 15], a well exposed pixel includes more reliable information than an under/over-exposed pixel. Therefore, weighting factor of a well exposed pixel is larger than that of an under/over-exposed pixel. Such a weighting function $\gamma(z)$ is defined as [12]

$$\gamma(z)=\begin{cases} z+1; & \text{if } z\leq 127\\ 256-z; & \text{otherwise}\end{cases}.$$

Two weighting factors of a gradient vector $(\nabla Y_{k,x}(p),\nabla Y_{k,y}(p))$ are computed as

$$\Gamma_{k,x}(p)=\gamma(Y_k(p))\gamma(Y_k(p_r)),$$
$$\Gamma_{k,y}(p)=\gamma(Y_k(p))\gamma(Y_k(p_b)),$$

where $p_r=(x,y+1)$ and $p_b=(x+1,y)$ are the right pixel and the bottom pixel of pixel p, respectively.

The logarithmic conversion allows us to measure local contrast by using spatial difference [3]. The desired vector fields $\vec{V}$ and $\vec{U}$ are thus computed in *log* domain. Let $\mathbf{V}_q$, $\mathbf{U}_q$, $\mathbf{\Gamma}_{k,q}$, and $\mathbf{Y}_{k,q}$ be vectors containing all $V_q(p)$'s, $U_q(p)$'s, $\Gamma_{k,q}(p)$'s, and $Y_{k,q}(p)$'s respectively. The desired vector fields are computed as

$$\mathbf{V}_q=\mathbf{U}_q=\frac{\sum_{k=1}^{L}\mathbf{\Gamma}_{k,q}\nabla\log(\mathbf{Y}_{k,q})}{\sum_{k=1}^{L}\mathbf{\Gamma}_{k,q}}, q\in\{x,y\}. \quad (12)$$

The value of $w_g(p)$ is computed by using all luminance components $Y_k(1\leq k\leq L)$ in log domain. Let $\sigma^2_{\log(Y_k),1}(p)$ be the local variance of $\log(Y_k)$ in a 3×3 square window centered at the pixel p. Due to different exposures, a well exposed pixel in one input image could be under/over-exposed in another image. This implies that the value of $\sigma^2_{\log(Y_k),1}(p)$ is different for different k. On the other hand, gradient magnitude becomes larger when a pixel get better exposed, and it decreases as the pixel becomes under/over-exposed. Therefore, the largest value of $\sigma^2_{\log(Y_k),1}(p)$ along all k's is chosen to represent the overall local variance of pixel p. The value of $w_g(p)$ is then given as

$$w_g(p)=\frac{1}{N}\sum_{p'=1}^{N}\frac{\max\limits_{1\leq k\leq L}\{\sigma^2_{\log(Y_k),1}(p)\}+0.001}{\max\limits_{1\leq k\leq L}\{\sigma^2_{\log(Y_k),1}(p')\}+0.001}. \quad (13)$$

The value of $w_g(p)$ is larger than 1 if p is at an edge and smaller than 1 if p in a smooth area. Clearly, larger weights are assigned to pixels at edges than those pixels in flat areas by using the proposed weight $w_g(p)$ in Equation (13). Applying this content-aware weighting, it is observed that the halo artifacts are greatly reduced. However, there might be blocking artifacts in final images. To prevent possible blocking artifacts from appearing in the final image, the value of $w_g(p)$ is smoothed by using a Gaussian filter.

With the vector fields $\mathbf{V}_q$ and $\mathbf{U}_q$ in Equation (12) and the weighting function $w_g(p)$ in Equation (13), fine details are extracted from all input images $X_k(1\leq k\leq L)$ simultaneously by using the proposed filter (11). Instead of adding all fine details to an intermediate image as in the detail-enhanced exposure fusion algorithm in [12], the extracted fine details are selectively added to the intermediate image. The proposed algorithm is based on the following two observations:

Observation 1: There is a fundamental limitation for fully detail-enhanced exposure fusion algorithms, i.e., noise is also amplified when fine details are enhanced. The HVS can tolerate amplified noise in complex regions but is particularly sensitive to amplified noise in flat areas.

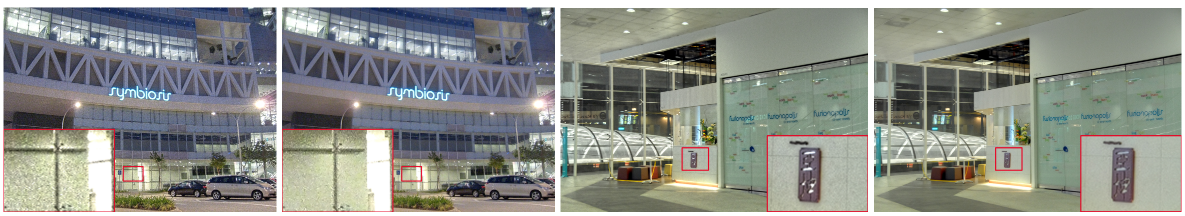
Observation 2: It is very challenging to separate noise from fine details.

With the proposed detail enhancement algorithm, fine details in all regions except flat ones are amplified and added to the intermediate image. Mathematically, the proposed selectively detail-

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

(e) (f) (g) (h)

Fig. 2. Comparison of the proposed algorithm with a fully detail-enhanced exposure fusion algorithm based on the filter in [5]. (a, b, c, d) by the filter in [5] and (e, f, g, h) by the proposed filter.



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

Fig. 3. Comparison of the proposed algorithm with the algorithm in [12]. (a, c) by the algorithm in [12] and (b,d) by the proposed algorithm.

enhanced exposure fusion algorithm is represented as

$$X_f(p) = X_{int}(p) \exp^{\theta e(p)\eta(p)}, \qquad (14)$$

where X_{int} is an intermediate image that is fused by using the exposure fusion scheme in [15]. $\theta(\geq 0)$ is a constant and its value is selected as 1.5 in this paper. The value of $\eta(p)$ is computed by using $w_g(p)$ in Equation (13). Its value is almost 0 if the pixel p is in a flat region and 1 otherwise.

5. EXPERIMENTAL RESULTS

The proposed exposure fusion algorithm is compared with two fully detail-enhanced exposure fusion algorithm. One is given in [12] and the other is obtained by using the gradient domain bilateral filter in [5] to replace the l2 norm based optimization method in [12]. Readers are invited to view to the electronic version of the full-size figures and zoom in these figures in order to better appreciate the differences among images.

The proposed exposure fusion algorithm is first compared with a fully detail-enhanced exposure fusion algorithm based on the bilateral filter in [5] by testing four sets of differently exposed images. Both the values of σ_1 and σ_2 in Equation (10) are selected as 2's. It is shown in Fig. 2 that the proposed filter can be applied to reduce halo artifacts significantly from the final images than the filter in [5].

The proposed exposure fusion algorithm is then compared with the exposure fusion algorithm in [12] by testing two sets of differently exposed images. The extracted fine details are also amplified by 1.5 times and added to the intermediate image by using the algorithm in [12]. It is illustrated in the zoom-in parts of Fig. 3 that the proposed exposure fusion algorithm does not amplify noise in flat areas as heavily as the fully detail-enhanced exposure fusion algorithm in [12].

6. CONCLUSION REMARKS AND DISCUSSIONS

A content adaptive bilateral filter is proposed in gradient domain by taking the characteristics of the human visual systems into consideration. The proposed bilateral filter can be applied to extract fine details from a set of images simultaneously. As such, it is adopted to design a selectively detail-enhanced exposure fusion algorithm. The proposed exposure fusion algorithm reduces halo artifacts significantly and prevents noise in flat areas from being amplified. Experimental results show that the resultant algorithm can produce images with better visual quality. It is worth noting that the similar idea can be used to improve the trilateral filter in [6] and the guided image filtering in [9].

Similar to the content adaptive bilateral filters in [10, 11], the acceleration of the proposed filter could be an issue. Fortunately, the idea in [16] might be borrowed to accelerate the proposed filter. This will be studied in our future research.

References

[1] C. Tomasi and R. Manduchi, “Bilateral filtering for gray and color images,” In *IEEE International Conference on Computer Vision*, pp. 836-846, 1998.

[2] F. Durand and J. Dorsey, “Fast bilateral filtering for the display of high-dynamic-range images,” ACM Trans. on Graphics, vol.21, no.3, pp.257-266, Aug. 2002.

[3] R. Fattal, M. Agrawala, and S. Rusinkiewicz, “Multiscale shape and details enhancement for multi-light image collections,” ACM Trans. on Graphics, vol. 26, no.3, pp.51:1-51:10, Aug. 2007.

[4] H. Winnemoller, S. C. Olsen, and B. Gooch, “Real-time video abstraction,” In *ACM SIGGRAPH*, 2007.

[5] Z. G. Li, J. H. Zheng, Z. J. Zhu, S. Q. Wu, and S. Rahardja, “A bilateral filter in gradient domain,” In *2012 International Conference on Acoustics, Speech, and Signal Processing*, pp. 1113-1116, 2012.

[6] P. Choudhury and J. Tumblin, “The trilateral filter for high contrast images and meshes,” In *Eurographics Symposium on Rendering*, 2003.

[7] J. Chen, S. Paris, and F. Durand, “Real-time edge aware image processing with the bilateral grid,” ACM Trans. on Graphics, vol. 26, no. 3, pp. 103-111, Aug. 2007.

[8] Z. Farbman, R. Fattal, D. Lischinshi, and R. Szeliski, “Edge-preserving decompositions for multi-scale tone and details manipulation,” ACM Trans. on Graphics, vol. 27, no. 3, pp. 249-256, Aug. 2008.

[9] K. He, J. Sun, and X. Tang, “ Guided image filtering,” IEEE Trans. On Pattern Analysis and Machine Learning, vol. 35, no. 6, pp. 1397-1409, Jun. 2013.

[10] B. Y. Zhang and J. P. Allebach, “Adaptive bilateral filter for sharpness enhancement and noise removal,” IEEE Trans. on Image Processing, vol. 17, no. 5, pp. 664-678, May 2008.

[11] Z. G. Li, J. H. Zheng, Z. J. Zhu, S. Q. Wu, W. Yao, and S. Rahardja, “Content adaptive bilateral filtering,” In *IEEE 2013 International Conference on Multimedia and Expo*, 2013.

[12] Z. G. Li, J. H. Zheng, and S. Rahardja, “Detail-enhanced exposure fusion,” IEEE Trans. on Image Processing, vol. 21, no. 11, pp. 4672-4676, Nov. 2012.

[13] G. Petschnigg, M. Agrawala, H. Hoppe, R. Szeliski, M. Cohen, and K. Toyama, “Digital photography with flash and non-flash image pairs”, ACM Trans. on Graphics (Proc. SIGGRAPH), vol. 22, no. 3, pp. 1-9, Aug. 2004.

[14] E. Elsemann and F. Durand, “Flash photography enhancement via intrinsic relighting”, ACM Trans. on Graphics (Proc. SIGGRAPH), vol. 22, no. 3, pp. 1-9, Aug. 2004.

[15] T. Mertens, J. Kautz, and F. V. Reeth, “Exposure fusion: a simple and practical alternative to high dynamic range photography,” Computer Graphics Forum, vol. 28, pp. 161-171, 2009.

[16] J. Baek and David E. Jacobs, “Accelerating spatially varying Gaussian filters,” In *SIGGRAPH Asia 2010*, 2010.