Research Article

Adaptive Total Variation Minimization-Based Image Enhancement from Flash and No-Flash Pairs

Sang Min Yoon,[1] Yeon Ju Lee,[2] Gang-Joon Yoon,[2] and Jungho Yoon[3]

[1] *School of Computer Science, Kookmin University, 77 Jeongneung-ro, Seogbuk-gu, Seoul 136-702, Republic of Korea*
[2] *Institute of Mathematical Sciences, Ewha Womans University, 52 Ewhayeodae-gil, Seodaemun-gu, Seoul 120-750, Republic of Korea*
[3] *Department of Mathematics, Ewha Womans University, 52 Ewhayeodae-gil, Seodaemun-gu, Seoul 120-750, Republic of Korea*

Correspondence should be addressed to Sang Min Yoon; smyoon@kookmin.ac.kr

We present a novel approach for enhancing the quality of an image captured from a pair of flash and no-flash images. The main idea for image enhancement is to generate a new image by combining the ambient light of the no-flash image and the details of the flash image. In this approach, we propose a method based on Adaptive Total Variation Minimization (ATVM) so that it has an efficient image denoising effect by preserving strong gradients of the flash image. Some numerical results are presented to demonstrate the effectiveness of the proposed scheme.

1. Introduction

A camera flash provides a practical and tractable lighting environment creating vivid and bright images. Yet, the camera flash produces some undesirable effects like uneven lighting and distracting sharp shadows [1]. To solve the dilemma, a flash photography method has been attracting great interest [1–5] to transfer the details from the flash image, while removing the ambient shadows captured from the no-flash image.

We briefly address the method of generating a high-quality image from two captured images: a flash image (Z) and a no-flash image (Y). The task in hand is to generate a new image (X) that contains the ambient light of the no-flash image and preserves the details of the flash image. To be more precise, since the given no-flash image Y might be noisy or blurry, we generate an estimated version $\widehat{Y}$ of Y, enhanced with the help of the flash image. Also, we produce a filtered version $\widehat{Z}$ from the flash image Z so that $Z - \widehat{Z}$ can provide details. Then, a new image is generated by the two images called a basis layer and a detail layer, respectively, as follows [1]:

$$X = \widehat{Y} + \tau \left(Z - \widehat{Z} \right). \tag{1}$$

Here τ is a parameter controlling the tradeoff.

The main issue in this method is how to generate the two images $\widehat{Y}$ and $\widehat{Z}$. Recently, numerous approaches for image denoising and enhancement from flash effects have been proposed by modeling the behavior of flash light or understanding the geometry of the scene. In order to extract $\widehat{Y}$ and $\widehat{Z}$ for image denosing, some techniques based on filtering method have been proposed. The bilateral filtering technique by Tomasi and Manduchi [4] and the total variation (TV) minimization technique by Wang et al. [6] have used edge-preserving and noise-reducing smoothing filters. Eisemann and Durand [5] proposed a cross-bilateral filter where they modified the bilateral filter [4] and computed the edge-preserving term as a function of the flash-image values. Their method preserves edges that do not really appear in the no-flash image. Eisemann and Durand [5] replace the intensity value at each pixel in an image by a weighted average of intensity values from nearby pixels.

PDE-based denoising methods have been used for image smoothing with edge preservation, which are based on variational approach of energy functional minimization. A popular variational denoising method is the TV minimizing process of Rudin-Osher-Fatemi [7]. In the literature of image restoration, the TV regularization term has an effect of preserving salient edges and removing noises. When the variation minimizing influence is too strong, however, it

causes smooth regions to be flat or constant so that the restored image looks unnatural. This is known as the staircasing effect [8], which is mainly due to the fact that the TV minimization method approximates the image to piecewise constant approximation. To avoid the staircasing effect and obtain a higher order approximation of the reconstructed image, several variants of the TV functional have been proposed [8–10].

In this work, we especially employ an adapted TV minimization (ATVM) technique [11] with high accuracy so that it can preserve the details of flash image. To preserve strong edges while smoothing noises, we add the TV regularization term to the moving least squares method with the weight functions that consider the similarity of the local areas between the evaluation and the reference positions. In our proposed method, we construct a local polynomial approximation p by solving the TV based minimization at each pixel position $\mathbf{x}$ as follows:

$$
f(\mathbf{x}) := \operatorname*{argmin}_{p \in \Pi_L} \left\{ \sum_{n=1}^{N} |\nabla p(\mathbf{x}_n)| + \frac{\mu}{2} |p(\mathbf{x}_n) - I(x_n)|^2 \right.
$$
$$
\left. \times \theta(\|\mathbf{x} - \mathbf{x}_n\|) \right\}, \tag{2}
$$

where $\nabla g = (\partial g/\partial x, \partial g/\partial y)$ indicates the gradient vector of g, Π_L is the space of bivariate polynomials of degree $\leq L$, and θ is a weighting function. The adaptive TV minimization scheme retains the merit of the original method which preserves the geometrical information and performs the image denoising, while obtaining a solution image of high accuracy.

2. Our Proposed Approach

Suppose that our observed data is a discrete sampling of a function at a point set in a domain $\mathscr{D} \subset \mathbb{R}^2$. That is, an image I is regarded as a set of function values $I := \{I(i,j) : i = 1,\ldots,M_1, j = 1,\ldots,M_2\}$. In this paper, we use the notation $[1,\ldots,M_1] \times [1,\ldots,M_2] = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)$, where $N = M_1 \times M_2$ is the size of the image. In most cases, images are contaminated by the noise introduced by the image acquisition process. So, an image can be written as $I(\mathbf{x}_n) = f(\mathbf{x}_n) + \epsilon_n, n = 1,\ldots,N$, where $f(\mathbf{x}_n)$ is the value of an underlying function f and ϵ_n indicates the additive noise at the location $\mathbf{x}_n$. The denoising method to construct a new image from no-flash and flash images is discussed below.

(i) Denoising Method. In this section, we propose a new energy functional to be minimized for image denoising which is useful for flash photography. Given a non-flash image Y and a flash image Z, our image denoising method to extract $\widehat{Y}$ and $\widehat{Z}$ is designed to generate a new high-quality image defined on the domain Ω. Since flash and non-flash images have different features, it is desirable to construct a denoising scheme considering the characteristics of both images.

For a given noisy image I, the TV minimization technique generates a denoised image $\widehat{I}$ by solving the following minimization problem:

$$
\widehat{I} = \operatorname*{argmin}_{u} \int_{\mathscr{D}} |u - I|^2 d\mathbf{x} + \lambda \int_{\mathscr{D}} |\nabla u| \, d\mathbf{x}, \tag{3}
$$

where $\mathbf{x}$ denotes the position vector of the pixel and ∇u is the gradient operator. This model is particularly useful for flash photography since the minimization of the TV norm eliminates small scaled structures in images very fast from the image. The incorporation of the TV minimization term reduces the influence of outliers which is mainly due to the uneven lighting. The second term in (3) is called the total variation norm, and its minimizing has the property of preserving sharp discontinuities (edges) in images while removing noise. This is a desirable property for images, since the visual quality of an image greatly depends on the preservation of edges. However, this TV scheme processes the observed image toward a piecewise constant image, which exhibits many false jump discontinuities and is visually unpleasant. This is mainly due to the fact that the TV minimization variation method approximates to an image with a first-order accuracy. In this case, it may lose the specific features of flash images. In this point of view, we employ an adapted TV minimization technique which can preserve the details of flash image [11]. Let I be a given reference image defined on a domain Ω, and let $\mathbf{x}$ be a fixed point in Ω. We seek a solution by using local polynomial of degree L in $\mathbb{R}^2 : p(\mathbf{r}) := p_{\mathbf{x}}(\mathbf{r}) := \sum_{|\alpha|_1 \leq L} c_\alpha \mathbf{r}^\alpha$. The coefficients c_α are obtained by minimizing the following energy functional:

$$
\operatorname*{argmin}_{p \in \Pi_L} \left\{ \sum_{n=1}^{N} |\nabla p(\mathbf{x}_n)| + \frac{\mu}{2} |p(\mathbf{x}_n) - I(\mathbf{x}_n)|^2 \theta(\mathbf{x}, \mathbf{x}_n) \right\}, \tag{4}
$$

where Π_L is the space of bivariate polynomials of degree $\leq L$ and θ is a weighting function. In this construction, we use a weight function θ defined as

$$
\theta(\mathbf{x}, \mathbf{r}) = \exp \left\{ -\frac{G_a * (\|I(\mathbf{x} + \cdot) - I(\mathbf{r} + \cdot)\|^2)(0)}{h_0^2} \right\}, \tag{5}
$$

where h_0 is a small positive value, G_a is the Gaussian function with standard deviation a, and $+\cdot$ denotes the shifting operator. The weight function is data adaptive and considers the similarity of the local areas between two positions $\mathbf{x}$ and $\mathbf{r}$. In our proposed method, we construct p at each pixel in the image, and so we are solving the minimization problem in (4) locally at each pixel in Ω. Then, the overall approximation function $\widehat{I}$ becomes $\widehat{I}(\mathbf{x}) := p(\mathbf{x}) := p_{\mathbf{x}}(\mathbf{x})$ for all $\mathbf{x} \in \Omega$.

For solving our proposed minimization model (4), we employ the split Bregman iteration algorithm [12]. By placing

(a) Noisy input image

(b) Image denoising using BF [4]

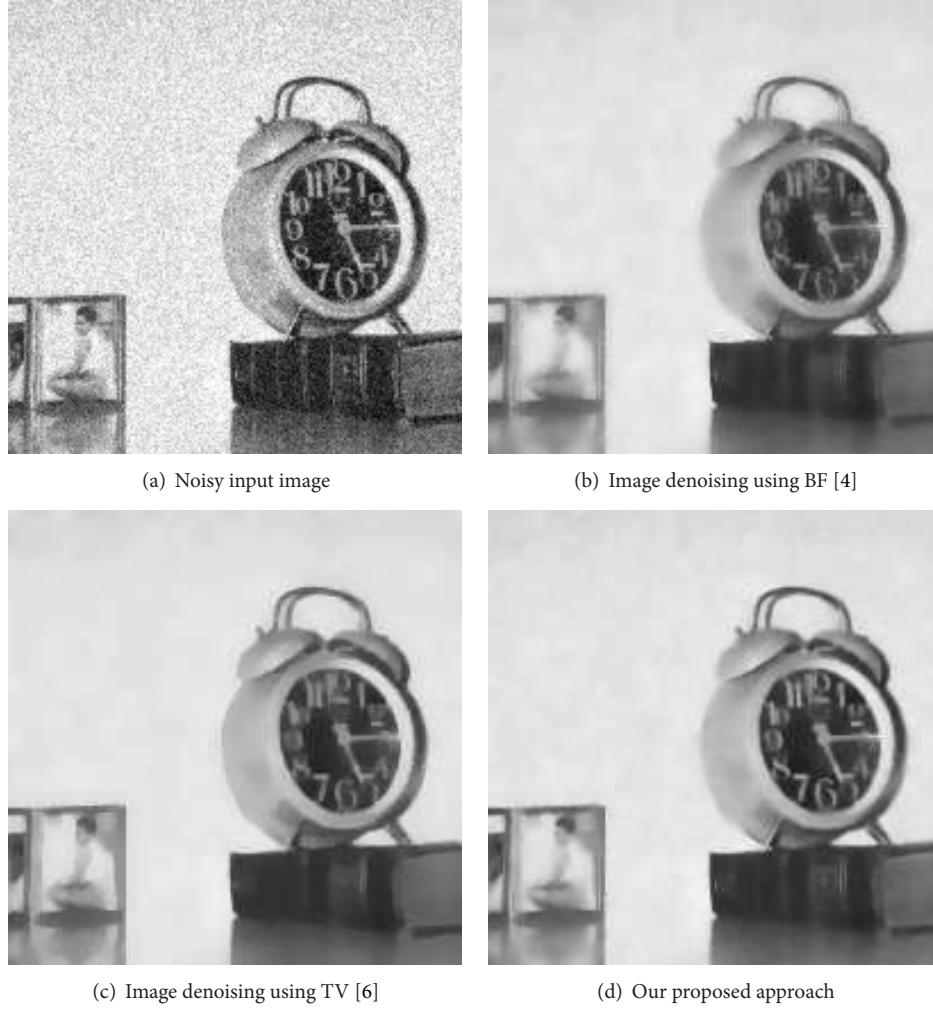(c) Image denoising using TV [6]

(d) Our proposed approach

FIGURE 1: Comparison of image denoising between previous approaches and our proposed approach.

the iterative minimization scheme into our problem (4), we get the following iteration steps for each $\mathbf{x}$:

(a) $p^{k+1}(\mathbf{x}) = \min_{p} \left\{ \sum_{n=1}^{N} \frac{\mu}{2} |p(\mathbf{x}_n) - I(\mathbf{x}_n)|^2 \theta(\mathbf{x}, \mathbf{x}_n) + \frac{\lambda}{2} \right.$

$\left. \times \left| d^k(\mathbf{x}_n) - \nabla p(\mathbf{x}_n) - b^k(\mathbf{x}_n) \right|^2 \right.$

$\left. \times \theta(\mathbf{x}, \mathbf{x}_n) \right\}$

(b) $d^{k+1}(\mathbf{x}) = \min_{d} |d| + \frac{\lambda}{2} \left| d - \nabla p^{k+1}(\mathbf{x}) - b^k(\mathbf{x}) \right|^2$

$= \text{shrink}\left(\nabla p^{k+1}(\mathbf{x}) + b^k(\mathbf{x}), \frac{1}{\lambda} \right)$

(c) $b^{k+1}(\mathbf{x}) = b^k(\mathbf{x}) + \nabla p^{k+1}(\mathbf{x}) - d^{k+1}(\mathbf{x}).$

$$(6)$$

Without the first term in (4), the energy functional in (4) becomes just as that used in the conventional least square

approximation which fits data by local polynomial approximation [13–15]. In our scheme (4), the TV regularization term endows the better denoising property. Normally, a least square approximation is weak against noise, since, in general, least squares methods are weak against outliers. On the other hand, TV based methods are strong against outliers, since outliers have large variation values. The TV regularization term eliminates noise very fast and helps the regularized least squares framework to produce a better approximation to the original noise-free image. Therefore, the proposed scheme can also provide a better solution than conventional TV schemes.

(ii) General Algorithm. Let Y and Z be the non-flash and the flash images, respectively. The algorithm to construct a new image X is given as follows.

Algorithm 1.

Step 1. Apply the ATVM algorithm to generate denoised images $\widehat{Y}$ and $\widehat{Z}$ from the given images non-flash image Y and the flash image Z.

(a) No-flash image (Y)   (b) Flash image (Z)   (c) Enhanced Image (X)

FIGURE 2: Image enhancement from real flash/no-flash image pairs.

Step 2. Choose a suitable tension parameter τ (We use τ = 0.02 for the all experiments).

Step 3. Construct a new image by using the following equation

$$X = \widehat{Y} + \tau \left(Z - \widehat{Z} \right). \tag{7}$$

3. Experiments

In this section, we conduct experimental results on two flash and no-flash image pairs to show the effectiveness of our proposed flash photography with a two-step approach.

First, we need to efficiently denoise the no-flash image while transferring the fine detail of the flash image and maintaining the ambient lighting of the no-flash image. ATVM based denoising method yields much better results than remarkable existing approaches like bilateral filter and traditional TV. Figure 1 provides the comparison that our proposed ATVM method is very powerful to remove the noise while keeping the details of the target object. For detailed comparison of image denoising, additive white Gaussian noise (AWGN) with σ = 0.6 is added as shown in Figure 1(a). We can remove the noise by using bilateral filter in Figure 1(b), traditional TV in Figure 1(c), and our proposed approach in Figure 1(d). Our proposed approach was better than the previous approaches in emphasizing the details and edges of the target object. This characteristic is very effective

TABLE 1: Quantitative comparison of image denoising between our approach and previous approaches using PSNR.

	Noisy image	BF [4] approach	TV [6] approach	Our approach
PSNR	24.6150	27.5591	27.4700	29.8629

to transfer the details of the target object from that captured from the flash image. The comparison of image denoising and recovering as shown in Figure 1 shows that our proposed ATVM (in Figure 1(d)) also removes the staircasing effect within the object. Table 1 shows a quantitative comparison of image denoising and deblurring using Peak Signal-to-Noise Ratio (PSNR) values between our result and the previous approaches. Table 1 proves that our proposed ATVM method out performs the previous approaches directly by transferring the details while removing the ambient shadows.

The second step in our experiments is to create a new image from denoised flash and no-flash image pair using (1). Based on the experimental results using ATVM, we can dramatically improve the quality of the image as shown in Figure 2. It shows the enhanced image using real flash and no-flash image pairs as represented in Figures 2(a) and 2(b). The noise from no-flash image as shown in Figure 2(b) is significantly reduced, but the ambient lighting in no-flash image in Figure 2(a) is still kept in our proposed approach. Figure 2(c) also shows that our proposed approach preserves

both details of the object and smooth continuity between foreground and background even though the borders of the foreground and background are not clearly separated.

4. Conclusion

In this work, we propose a novel image enhancement approach from flash and no-flash image pairs. To combine the strengths of the flash and no-flash images, we adopt the MLS method and the TV minimization together. The MLS method performs fast denoising and enables us to find a sufficiently smooth approximation. And the TV based image denoising and debluring approach shows robustness in reducing the noise while keeping the edges from flash/no-flash pairs. The experiments using real images show that ATVM based image enhancement answers our intention considerably. Enhancing the resolution from the flash and no-flash pairs, the proposed scheme retains the merit of the original TV minimization method which preserves the geometrical information while obtaining an image of higher order approximation. In particular, our proposed ATVM works well on maintaining the detail of the target object while reducing the noise without prior information.
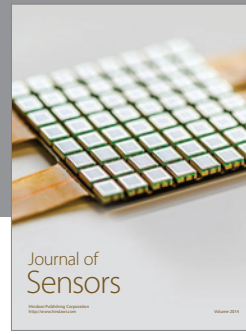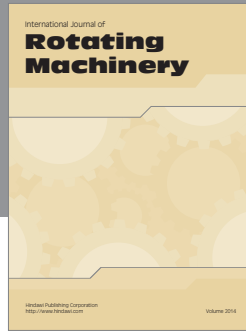
Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper

Acknowledgments

References

[1] 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*, vol. 23, no. 3, pp. 664–672, 2004.

[2] F. Durand and J. Dorsey, "Fast bilateral filtering for the display of high-dynamic-range images," in *Proceedings of the ACM Transactions on Graphics (SIGGRAPH '02)*, pp. 257–266, July 2002.

[3] B. M. Oh, M. Chen, J. Dorsey, and F. Durand, "Image-based modelling and photo editing," in *Proceedings of the Computer Graphics Annual Conference (SIGGRAPH '01)*, pp. 433–442, August 2001.

[4] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proceedings of the IEEE 6th International Conference on Computer Vision*, pp. 839–846, January 1998.

[5] E. Eisemann and F. Durand, "Flash photography enhancement via intrinsic relighting," in *Proceedings of ACM Transactions on Graphics (SIGGRAPH '04)*, pp. 673–678, August 2004.

[6] Y. Wang, W. Yin, and Y. Zhang, "A fast algorithm for image deblurring with total variation regularization," Tech. Rep., CAM, 2007.

[7] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, no. 1–4, pp. 259–268, 1992.

[8] M. Nikolova, "Weakly constrained minimization: application to the estimation of images and signals involving constant regions," *Journal of Mathematical Imaging and Vision*, vol. 21, no. 2, pp. 155–175, 2004.

[9] V. Caselles, A. Chambolle, and M. Novaga, "The discontinuity set of solutions of the TV denoising problem and some extensions," *Multiscale Modeling and Simulation*, vol. 6, no. 3, pp. 879–894, 2007.

[10] M. Nikolova, "Local strong homogeneity of a regularized estimator," *SIAM Journal on Applied Mathematics*, vol. 61, no. 2, pp. 633–658, 2000.

[11] Y. Lee, S. Lee, and J. Yoon, "A framework for moving least squares method with total variation minimizing regularization," *Journal of Mathematical Imaging and Vision*.

[12] T. Goldstein and S. Osher, "The split Bregman method for l1 regularized problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 2, pp. 323–343, 2009.

[13] P. Lancaster and K. Salkauskas, "Surface generated by moving least squares methods," *Mathematics of Computation*, vol. 37, no. 155, pp. 141–158, 1981.

[14] D. Levin, "The approximation power of moving least-squares," *Mathematics of Computation*, vol. 67, no. 224, pp. 1517–1531, 1998.

[15] D. Levin, "Mesh-independent surface interpolation," in *Geometric Modeling For Scientific Visualization*, G. Brunnett, B. Hamann, H. Mueller, and L. Linsen, Eds., Springer, 2003.