A Fusion Algorithm for Flash and No-Flash Images under a Low-Light Environment

Yan-li ZHANG
Computer Science Department
Guangdong AIB Polytechnic
Guangzhou, 510507, China
(86)13610048912
zhangyanli@gdaib.edu.cn

Jia-min WANG
School of Telecommunication and Information Engineering
XI'AN University of Posts & Telecommunications
Xian, 710121, China
(86)18700592397
18700592397@163.com

Guo-liang JIANG
The State Key Laboratory of Integrated Services Networks
Xidian University
Xian, 710000, China
(86)029-88204409
gljiang@stu.xidian.edu.cn

Bin SONG*
The State Key Laboratory of Integrated Services Networks
Xidian University
Xian, 710000, China
(86)13891880055
bsong@mail.xidian.edu.cn

ABSTRACT

Pictures which are taken under a dim light condition might be noisy or relatively dark. Generally, we open the flash to increase the exposure to obtain a flash image. It has high definition but often changes tone of the image. On the contrary, a no-flash image remains the tone but has low definition. Therefore, it is natural to explore a fusion algorithm to combine the advantages of the two kinds of images in order to improve the final image quality. A common fusion algorithm is composed of matching and filtering. However, the conventional matching algorithms are only suitable for the same illumination situations. There may be few feature points while matching based on flash and no-flash images of the same scene. After matching, we employ the filter based on multi-scale detail extraction to obtain the details and it may lead to color distortion. Despite such problems, the conventional matching and filtering algorithms have their own advantages. In order to inherit their advantages on matching and filtering and address the potential problems, we propose an advanced matching algorithm ASIFT and an advanced filtering algorithm AWLS in this paper. Experimental results demonstrate that ASIFT can increase the number of matching pairs and ALWS can eliminate color distortion and improve the details.

Keywords

ASIFT, AWLS, Flash, Image fusion, Low-Light, No-Flash

1. INTRODUCTION

Digital photography has made it possible to take a pair of images in low-light environments rapidly and efficiently: one with flash to capture details and the other without flash to capture color. Images captured in dim light are hardly satisfied. They could be very noisy when increasing ISO in short exposure duration. Moreover, it creates unwanted shadow and highlight, or changes the overall scene appearance to cold and gray (Adams et al., 1998) if we use flash to improve illuminance. Therefore, the method of restoring a color image based on flash and no-flash inputs of the same scene (Kang et al., 2003; Liu Gang, 2006; Wang Zhong-hua, 2007), which consists of two parts, namely the matching algorithm called Scale Invariant Feature Transform (SIFT) (Lowe, 2004) and the filtering algorithm called Weighted Least Squares (WLS) (Farbman et al., 2008; Rukhlenko, 2007), is proposed in this paper.

SIFT (Scale Invariant Feature Transform) is an algorithm in computer vision to detect and describe local features in images. SIFT features are invariant to image scale and rotation. For this reason, we use an affine transformation matrix to record the information of corresponding feature points of the image pair. However, the matrix is hard to compute without enough matching pairs because of the uneven illumination and the over-concentration of feature points. To address this problem, an Advanced SIFT (ASIFT) is proposed to increase the number of matching pairs. We firstly equalize the two images to enhance the image contrast. Then we use Random Sample Consensus (RANSAC) (Chum and Matas, 2008) to eliminate the mismatching pairs after calculating and matching features points. If the matching pairs are less than a given threshold, we add a certain amount of matching pairs uniformly in images for calculating the affine transformation matrix. Experimental results show that ASIFT works much better than SIFT subjectively.

WLS (Weighted Least Squares) is suitable for multi-scale detail extraction, which is quite different from the conventional detail decomposition techniques based on the bilateral filter (Tomasi and Manduchi, 1998; Durand and Dorsey, 2002). However, WLS takes advantages of RGB channels to extract details which may result in color distortion. For the purpose of eliminating color distortion, we propose an Advanced WLS (AWLS). It converts RGB channels to YCbCr channels. We only deal with the Y channel, which is also called brightness Y but not process color components Cb and Cr. For the images, there are two kinds of details and brightness in flash and no-flash. This means we can select the better details and brightness in blocks. The final images in different scenes show that ALWS enhances the brightness and has high definition without color distortion compared with WLS and cross-field filter.

2. ADVANCED SIFT

Generally, we use RANSAC to eliminate the mismatching pairs after calculating SIFT features. In order to increase the number of matching pairs (PF_i ,PN_i) uniformly, first of all, we divide the flash image into $M\times M$ (M=2) sub-blocks and each sub-block

MOBIMEDIA 2016, June 18-19, Xi'an, People's Republic of China

DOI 10.4108/eai.18-6-2016.2264317

should be added with N×N (N=5) feature points. Hence, we have $D_x = \lfloor\frac{W}{M(N+1)}\rfloor$, $D_y = \lfloor\frac{H}{M(N+1)}\rfloor$, $W_{sub} = \lfloor\frac{W}{M}\rfloor$, $H_{sub} = \lfloor\frac{H}{M}\rfloor$, where W, H are the width and height of flash image, D_x, D_y represent the horizontal and vertical distance between any two adjacent feature points, W_{sub} ,H_{sub} stand for the width and height of each sub-block. Thus, the new image added with feature point of the flash image PF_i (i=1, 2…n，n is the sum of the matching pairs) can be derived as:

$$PF^x_{n+(2i+j)\times N\times N+i'\times N+j'} = j\times W_{sub} + j'\times D_x,$$
$$PF^y_{n+(2i+j)\times N\times N+i'\times N+j'} = i\times H_{sub} + i'\times D_y \qquad (1)$$

where i, j(i, j=0,1,2…M) are the row number and column number of the sub-block and i', j'(i', j'=0,1,2,… ,N) represent the row number and column number of the feature points in each sub-block. Since we have increased PF^x_i and PF^y_i, new image with feature point of the no-flash image PN_i should be computed with offsets, which are of the corresponding feature points in flash and no-flash images. The offsets horizontal Δ_x and $\Delta_y$vertical can be written as:

$$\Delta_x = \frac{1}{n}\sum_{i=1}^{n}(PF^x_i - PN^x_i),$$
$$\Delta_y = \frac{1}{n}\sum_{i=1}^{n}(PF^y_i - PN^y_i) \qquad (2)$$

Then PN_i can be acquired as:

$$PN^x_i = PF^x_i - \Delta_x,$$
$$PN^y_i = PF^y_i - \Delta_y \qquad (3)$$

As shown in Eq.(1), the number of matching pairs increases from n to n+(2i+j)×N×N+i'×N+j'. With the increased matching pairs, we can easily match the images through the affine transformation matrix for the latter filtering process. The red lines in Figure 1(c) and Figure1(d) show the number of matching pairs using SIFT and ASIFT. Figure 1(a) and Figure 1(b) are the mismatching and matching results caused by Figure 1(c) and Figure 1(d).

3. ADVANCED WLS

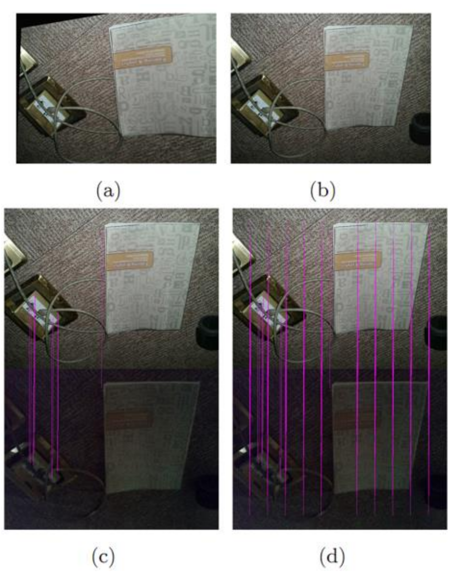

Figure 1: A comparison of the number of matching pairs between SIFT and ASIFT: (a) Unmatched (b) matched (c) SIFT (d) ASIFT

Since WLS deals with RGB directly, which causes color distortion (Eisemann and Durand, 2004), AWLS converts the RGB channels to YCbCr channels to remove the color distortion and only handles the brightness Y while the color components remain unchanged. We employ division to extract the details with the Y channels by computing a detail layer from the flash image as the following ratio

$$F_{detail} = \frac{F_{origin} + \varepsilon}{F_{base} + \varepsilon}, \qquad NF_{detail} = \frac{NF_{origin} + \varepsilon}{NF_{base} + \varepsilon} \qquad (4)$$

where F_{origin} is the brightness of the flash image, F_{base} represents the brightness of the denoised flash image and F_{detail} stands for the details of the flash image. The ratio captures the local detail variation in F and is commonly called a quotient image (Shashua and Riklin-Raviv, 2001) or ratio image (Liuet al., 2001) in computer vision.

At low signal values, the flash image contains noise that can generate spurious details. We add ε to both the numerator and denominator of the ratio to reject these low signal values and also avoid division by zero. In practice we use ε=0.02 (Petschnigg et al., 2004) across all our results.

NF_{origin}, NF_{base} and NF_{detail} are defined similarly but in terms of the no-flash image.

Although F_{detail} captures more details in quantity, NF_{detail} may have higher quality details in the same place. Hence, details selecting must be taken into consideration. Each image can be split into blocks in size of N×N (N=3). By computing the sums of the details of the corresponding blocks in the image pair, we can obtain a dominating outcome, which has a larger sum and superior details.

Similarly, NF_{base} which has a more significant brightness can be computed by comparing each pixel value of F_{base} and NF_{base}.

Since we have selected F_{detail} and NF_{base} the final image is given as:

$$Final = NF_{base} \times F_{detail}. \qquad (5)$$

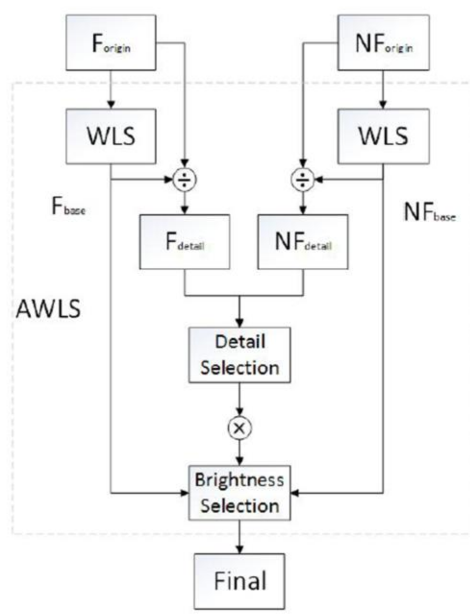

Figure 2: Flow chart of AWLS

Figure 2 is the flow chart of ALWS. As shown in Figure 3 (e), AWLS removes color distortion while Figure 3(c) (WLS) is greenish in the right side and red dish in the left side. We employ the cross-field filter for comparison. Cross-field filter designs functions to form an optimal scale map considering adaptive smoothing, edge preservation and guidance strength manipulation to solve the cross-field issues: Gradient Magnitude Variation, Gradient Direction Divergence, Gradient Loss, Shadow and Highlight by Flash. Figure 3(d) is the result of cross-field filter.

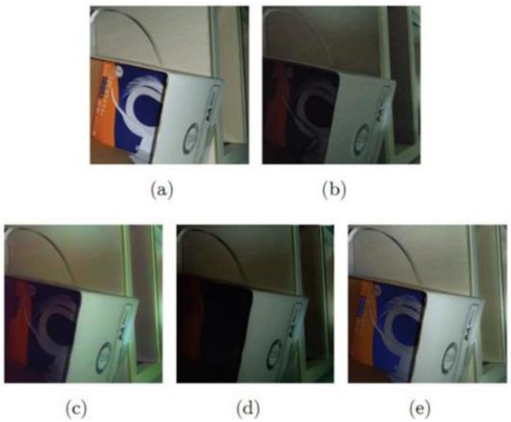

Figure 3: A comparison of color distortion between WLS, Cross and AWLS: (a) Flash (b) No-Flash (c) WLS (d) Cross (e) AWLS

4. SIMULATION RESULTS

A comparison between a fusion image and each of its originals can be measured subjectively. The fusion image combines the advantages of the two kinds of images through matching and filtering as mentioned earlier. Another way to measure the image similarity is the Structural Similarity Index Measurement system (SSIM), which comprehensively outperforms the Signal to Noise Ratio (SNR) and Peak Signal to Noise Ratio (PSNR). The SSIM values are close to 1, indicating the high similarity. Table 1 shows the values of structural similarity index measurement in different scenes. As shown in Figure 4 and Figure 6, AWLS eliminates the shadow and highlight by flash compared with cross-field filter. In addition, AWLS enhances the brightness compared with no-flash.

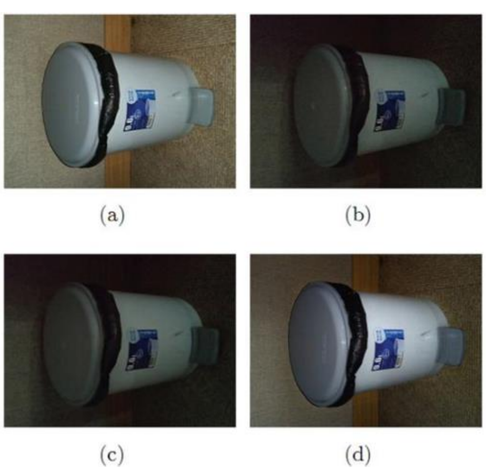

Figure 4: A comparison of results in scene 1: (a) Flash (b) No-Flash (c) ASIFT and Cross (d) ASIFT and AWLS

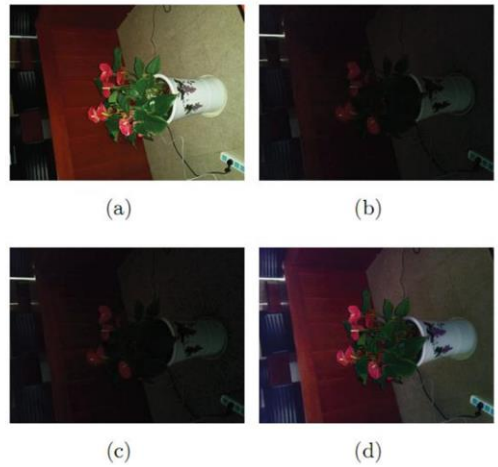

Figure 5: A comparison of results in scene 2: (a) Flash (b) No-Flash (c) ASIFT and Cross(d) ASIFT and AWLS

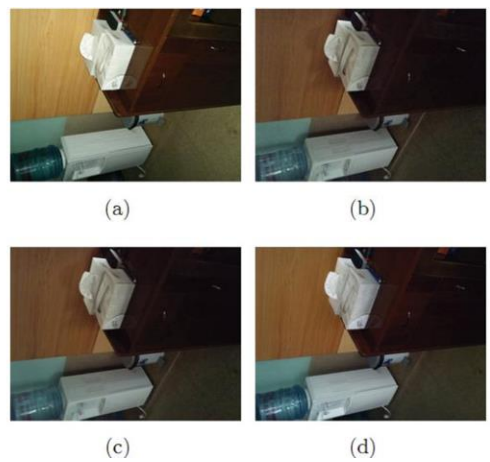

Figure 6: A comparison of results in scene 3: (a) Flash (b) No-Flash (c) ASIFT and Cross (d) ASIFT and AWLS

Moreover, it not only brings no color distortion and noise but also remains high definition.

5. CONCLUSIONS AND FUTURE WORK

In this paper, we present the ASIFT to increase the number of matching pairs for calculating the affine transformation matrix and the AWLS to reinforce the details and eliminate color distortion. Different from the conventional fusion algorithm, this framework is suitable for restore a color image based on flash and no-flash images taken a low-light environment, especially indoors. Experimental results demonstrate that such an image pair can be synthesized into a new image with higher quality than either of the originals.

An interesting problem requires to be investigated in our future work. Actually, the framework cannot achieve universality for all the light conditions. It works better indoors than outdoors due to the reason that illumination like reflected light, diffused light and even transmission light are more complex, non-uniform and unexpected in the night. For further study, we will concentrate on the outdoors research to eliminate the influence of the light.

Table 1: The values of structural similarity index measurement in different scenes

Scene	Cross	AWLS	Subjective quality
1	0.71	0.79	Better
2	0.89	0.98	Good
3	0.73	0.88	Good

Cross: cross-field filter; AWLS: Advanced WLS filter

6. ACKNOWLEDGMENTS

This work has been supported by the National Natural Science Foundation of China (Nos.61271173 and 61372068)

7. REFERENCES

[1] J.Adams, K. Parulski, K. Spaulding. Color processing in digital cameras. *IEEE Micro*, 18(6):20-30, 1998.

[2] O.Chum, J. Matas. Optimal randomized ransac. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(8):1472-1482, 2008.

[3] F.Durand, J. Dorsey. Fast bilateral filtering for the display of high-dynamic-range images. *ACM Transactions on Graphics*, 21(3):257-266, 2002.

[4] E. Eisemann, F. Durand. Flash photography enhancement via intrinsic relighting.*ACM Transactions on Graphics*, 23(3):673-678. [doi:10.1145/1015706.1015778], 2004.

[5] Z. Farbman, R. Fattal, D.Lischinski, R. Szeliski.Edge-preserving decompositions for multi-scale tone and detail manipulation. *ACM Transactions on Graphics* (TOG), 27(3):67, 2008.

[6] S. B.Kang, M. Uyttendaele, S. Winder, R. Szeliski. High dynamic range video. *ACM Transactions on Graphics*, 22(3):319-325, 2003.

[7] Liu Z, Shan Y, Zhang Z. Expressive expression mapping with ratio images [J]. *International Conference on Computer Graphics and Interactive Techniques*, p.271-276. [doi:10.1145/383259.383289], 2001.

[8] Liu G, Jing Z, Sun S. Multiresolution image fusion scheme based on fuzzy region feature [J]. *Journal of Zhejiang University Scienc*e, 7(2):117-122. [doi:10.1631/jzus.2006.A0117], 2006.

[9] D.G.Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91-110. [doi:10.1023/B:VISI.0000029664.99615.94], 2004.

[10] G.Petschnigg, R.Szeliski, M.Agrawala, M.Cohen, H.Hoppe, K.Toyama. Digital photography with flash and no-flash image pairs. *ACM transactions on graphics* (TOG), 23(3):664-672. [doi:10.1145/1015706.1015777], 2004.

[11] A. S.Rukhlenko. Iterative wls design of saw bandpass filters. IEEE Transactions on Ultrasonics. *Ferroelectrics and Frequency Control*, 54(10):1930-1935, 2007.

[12] A.Shashua, T.Riklin-Raviv. The quotient image: class based re-rendering and recognition with varying illuminations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(2):129-139, 2001.

[13] C. Tomasi, R.Manduchi. Bilateral filtering for gray and color images. *IEEE International Conference on Computer Vision*, p.839-846.[doi:10.1109/ICCV.1998.710815], 1998.

[14]Wang, Zhonghua, Qin. A framework of region-based dynamic image fusion [J]. *Journal of Zhejiang University Science*, 8(1): 56-62, 2007.