A Novel Imaging Method for Cell Phone Camera in Low Ambient Light Conditions Using Flash and No-Flash Image Pairs

Lin-bo LUO†,††, Jun CHEN†††, Sang-woo AN††, Chang-shuai WANG††, Jong-joo PARK††, Ying-chun LI††, Nonmembers, and Jong-wha CHONG††b, Member

SUMMARY In lowlight conditions, images taken by phone cameras usually have too much noise, while images taken using a flash have a high signal-noise ratio (SNR) and look unnatural. This paper proposes a novel imaging method using flash/no-flash image pairs. Through transferring the natural tone of the former to the latter, the resulting image has a high SNR and maintains a natural appearance. For realtime implementation, we use two preview images, which are taken with and without flash, to estimate the transformation function in advance. Then we use this function to adjust the tone of the image captured with flash in real time. Thus, the method does not require a frame memory and it is suitable for cell phone cameras.

key words: imaging method, cell phone camera, low ambient light, flash and no-flash image pairs

1. Introduction

Due to their low-cost and easy integration, CMOS image sensors (CIS) are popular nowadays. Especially with the coming of the smart phone era, every cell phone has one or more CISs. Due to their portability, people often use phone cameras to take photos. With these phone cameras, high-quality images can be captured in normal light environments, but usually only low-exposed and noisy images can be obtained in low-light environments. Using a fixed aperture, phone cameras have to collect more light in low ambient light by extending exposure time. However, extended exposure will cause motion blur due to camera movement or object motion [1].

Therefore, some algorithms have tried to deblur the motion-blurred images [2], [3] and some algorithms have denoised the noisy images [4], [5]. Though these algorithms achieved good performance, they need huge computational power and they cannot address the case of serious noise and blur.

Recently, many phone cameras are equipped with a flash for taking photos in lowlight conditions. These flash images have a high signal-to-noise ratio (SNR), but they appear unnatural and do not create a real atmosphere. For example, if you want to capture a photo with an intimate mood in a candlelit restaurant, there is no doubt that the flash will destroy the mood you want. Petschigg et al. proposed digital cameras using flash and no-flash image pairs [6]. This study achieved a good performance; unfortunately, it was designed for general digital cameras and implemented with software. If this method were conducted through hardware, real-time performance cannot be achieved and one or more frame memories would be necessary. As Table 1 shows, the area of a frame memory is 5~20 times that of a sensor chip. Therefore a frame memory must be eliminated to reduce the size and cost of the CIS chip.

In this study, through modifying the method of [6] and using the imaging framework without a frame memory which was proposed in our previous work [7], we propose a novel imaging method for cell phone cameras in low ambient light conditions using flash and no-flash image pairs. The method can preserve the real atmosphere of ambient light and keep a low noise level for the final image. The method can be implemented by hardware in real time and does not require a frame memory.

2. Proposed Method

As shown in Fig. 1, we use three consecutive images in our method: two preview images and one snapshot. The first image (P1) is previewed without flash the second preview image (P2) and the snapshot (S) are consecutively acquired using flash and the same camera setting. P1 has a low exposure and is noisy. However it looks like natural and is consistent with the real scene color. P2 and S have clear edges and textures but it looks unnatural. Our idea is that transferring the natural tone of P1 to S let S not only have clear details but also is comfortable to the human visual system and mental feeling.
To save on hardware cost, we first calculated the color mapping function of the no-flash and flash images through the two preview images. Then we used the function to modify the tone of the snapshot. Because both the calculation of the tone mapping function and the modification of snapshot can be implemented in real time, no image data need to be stored. Thus, the proposed method does not require a frame memory.

The innovation of this study is that we transfer the natural tone of P1 to the low-noise but unnatural image S, while the diﬀerence of the reference are almost smaller than 1 except for one outlier. Actually, it is easy to satisfy the experiment conditions of [7] and [8]. Many high-end image sensors even outperform the experimental devices and can achieve smaller displacements between two consecutive frames along X and Y axes are only 0.40 and 0.28 pixel. In our previous work [8], we performed another experiment that took 11 consecutive frames using a hand-held camera and calculated the difference of the mean and standard deviation between every two consecutive frames. The results showed the diﬀerence are almost smaller than 1 except for one outlier. Actually, it is easy to satisfy the experiment conditions of [7] and [8]. Many high-end image sensors even outperform the experimental devices and can achieve smaller diﬀerence. Since the displacements and the diﬀerence of the parameters of two consecutive frames are almost very small, it is reasonable to using the parameters of P2 to replace the coordinates of S.

It should be pointed out that our method is not suitable again if there is a drastic camera movement or object motion during imaging. However, as we know, it is not possible to obtain satisfactory images in this case even we use any other methods.

2.1 Tone Transformation

The essence of tone transformation (TT) is mapping the color palette of a source (reference) image to a target image. Among many TT algorithms [9], we chose a global color correction algorithm [10] that is conducted in Lαβ color space to perform the tone mapping from image P1 to image S. This global color correction algorithm can achieve good performance and more important is that this algorithm can be easily implemented by hardware. In addition, why we chose the orthogonal Lαβ color space [10] is because R, G, and B have a strong correlation which forces all color channels to be modiﬁed in tandem.

The procedure of the color space conversion is as follows. First, RGB data are converted to device dependent XYZ space:

\[
\begin{bmatrix}
X \\
Y \\
Z
\end{bmatrix} =
\begin{bmatrix}
0.4124 & 0.3576 & 0.1805 \\
0.2126 & 0.7152 & 0.0722 \\
0.0193 & 0.1192 & 0.9505
\end{bmatrix}
\begin{bmatrix}
R \\
G \\
B
\end{bmatrix},
\]

Then the XYZ space is converted to LMS color space:

\[
\begin{bmatrix}
L \\
M \\
S
\end{bmatrix} =
\begin{bmatrix}
0.3897 & 0.6890 & -0.0787 \\
-0.2298 & 1.1834 & 0.0464 \\
0.0000 & 0.0000 & 1.0000
\end{bmatrix}
\begin{bmatrix}
X \\
Y \\
Z
\end{bmatrix}.
\]

Finally, the LMS data are logarithmically compressed to reduce skew:

\[
\begin{bmatrix}
\alpha \\
\beta
\end{bmatrix} =
\begin{bmatrix}
\frac{1}{\sqrt{3}} & 0 & 0 \\
0 & \frac{1}{\sqrt{6}} & 0 \\
0 & 0 & \frac{1}{\sqrt{2}}
\end{bmatrix}
\begin{bmatrix}
\log L \\
\log M \\
\log S
\end{bmatrix},
\]

After the conversion of color space, take the L component as an example the transformation process is described as follows:

\[
L_{out} = \frac{\sigma_s}{\sigma_s}(L_s - \mu_s) + \mu_1,
\]

where \(L_{out}\) and \(L_s\) are the L components of the output images and the image S, \((\mu_s, \sigma_s)\) and \((\mu_1, \sigma_1)\) are the mean and standard deviation of S and P1. The transformation of \(\alpha\) and \(\beta\) components can be performed with same manner as L.

To eliminate a frame memory, we use the parameters of
P₂ instead of S to perform (4). Because the calculation of the parameters of P₂ can be completed before processing image S, the TT can be performed in real time. Otherwise, if the parameters of S itself are used, the mean and standard deviation need to be calculated first, and then use them to perform the transformation (4). However, when calculating the parameters, the image has to be buffered for later processing. Then, using at least one frame memory is unavoidable.

Once the TT function is used to correct one succeeding captured images, the function is not valid again. Because different illumination condition, objects and environmental conditions can result in different transformation function, the TT function should be calculated again when a new captured image needs to be corrected.

2.2 Pipelined Hardware Design

To implement the proposed algorithm with hardware, we propose some optimization schemes and to design a three-stage pipelined framework. As shown in Fig. 2, the framework consists of three stages: the estimation of μ₁ and σ₁ of P₁, the estimation of μ₂ and σ₂ of P₂, and tone transformation of S using the parameters calculated in stages I and II.

First, we use the method of [11] to accelerate the RGB2Labβ module. By omitting the logarithm, Labβ color space is converted to an approximation of the CIE Lab color space, and the complex color space conversion of (1)–(3) are simplified to a combined 3×3 matrix calculation whose forward and inverse transformation are described as (5) and (6), respectively. Then we use a fixed-point operation and right-shifting operation to further speed up the matrix calculations.

\[
\begin{bmatrix}
L \\
α \\
β
\end{bmatrix} = \begin{bmatrix}
0.3475 & 0.8231 & 0.5559 \\
0.2162 & 0.4316 & -0.6411 \\
0.1304 & -0.1033 & -0.0269
\end{bmatrix} \begin{bmatrix}
R \\
G \\
B
\end{bmatrix}
\] (5)

\[
\begin{bmatrix}
R \\
G \\
B
\end{bmatrix} = \begin{bmatrix}
0.5773 & 0.2621 & 5.6947 \\
0.5774 & 0.6072 & -2.5444 \\
0.5832 & -1.0627 & -0.2073
\end{bmatrix} \begin{bmatrix}
L \\
α \\
β
\end{bmatrix}
\] (6)

Through optimization stage I, II, and III can be completed in real-time with an acceptable latency L₁, L₂, and L₃, respectively. All of these latencies are much smaller than the interval between the any two images.

3. Simulation Results

As shown in Fig. 3, we test the CISs of four smart phones: Apple iPhone 4S, Samsung Galaxy Note LTE, HTC Desire G7, and Motorola XT800 W.

We can see that the no-flash images have a lot of noise and are not bright enough so that the details in some black area are not clear, while the flash images have clear details but are unnatural. The resulting images of our method not only have clear details, but also keep the mood well. Because the corrected images have the similar histogram, mean and stand deviation with no-flash images, we can say that the tone of corrected images is similar with the no-flash images.

We also used the three test images (carpet, lamp, and puppet) of [6] to simulate our algorithm. As shown in Fig. 4, both the resulting images of the Ref. [6] and our proposed method have the similar histogram, mean and stand deviation. We can say that both the two algorithms correct the tone of flash images well.

Comparing to [6], our results have less noise, and background details can be seen clearer, though the mood is not as vivid as [6].

From Fig. 3, we can see that our algorithm can be conducted even if camera movement occurs or object motion exists between the flash and no-flash image pairs. However [6] requires the two images capture exactly the same points in the scene.

Comparing to [2]–[6], our method can reduce calculations and be implemented with hardware. In our simulation, the latency of L₁, L₂ and L₃ are only a couple of system clocks, while the interval between two images is longer than 1 ms. The proposed algorithm can be implemented with hardware in real time and does not need any frame memory.

In addition, to illustrate the efficiency of the TT, we compared this algorithm to conventional histogram equalization (HE). As shown in Fig. 5, the results of HE are not ideal as the TT. When we correct the low-exposed images (e.g. Fig. 3 (a) and (b)), the resulting images are noisy. When we correct the images captured with flash (e.g. Fig. 3 (i) and (j)), the resulting images still appear unnatural and do not create a real atmosphere.

4. Conclusions

Through improving a previous algorithm [6] and using a hardware-oriented imaging mode, this paper proposed a
low-cost and efficient imaging method for cell phone cameras using flash and no-flash pairs. This method can be easily used by cell phone cameras without any other equipment. When user half press the shutter, two preview images are obtained. Then we use these two images to correct the captured image when the shutter releases. This method can be implemented by hardware in real time and does not require a frame memory. By not using a frame memory, the size and cost of the CIS chip can be greatly reduced, thus the proposed method can be easily integrated as a module of CIS.

To further improve the performance, some post-processing such as red-eye correction, flash hot spot or specular reflections removal, can be considered. For example,
Fig. 4 Results of [6] and the proposed method. (a)–(c) are no-flash images, (d)–(f) are flash images, (g)–(i) are the results of [6], while (j)–(l) are the results of our proposed algorithm. (m)–(o) are the histograms of the no-flash image, flash image, the resulting image of [6] and the corrected image of our proposed algorithm.

some areas which are close to the flash of camera are over-exposed. This issue has not been addressed in this paper and needs further research.

Acknowledgments

This research was supported by the Ministry of Knowledge Economy, Korea, under the ITRC support program supervised by the National IT Industry Promotion Agency (NIPA-2012-C1090-1200-0010). The IDEC provide research facilities for this study. This work was also supported by the Brain Korea 21 Project in 2012 and the IT R&D program of MKE/KEIT (10035570). This work was also supported by Natural Science Foundation of Hubei Province (2012FFB06408), China.

References

[1] B.-D. Choi, S.-W. Jung, and S.-J. Ko, “Motion-blur-free camera system splitting exposure time,” IEEE Trans. Consum. Electron., vol.54, no.3, pp.981–986, Aug. 2008.
[2] H. Ji and K. Wang, “Robust image deblurring with inaccurate blur kernels,” IEEE Trans. Image Process., vol.21, no.4, pp.1624–1634, 2012.
[3] H. Hong and I.-K. Park, “Single-image motion deblurring using adaptive anisotropic regularization,” Optical Eng., vol.49, no.9, pp.097008-1–097008-13, 2010
[4] S. Lee, V. Maik, J. Jang, J. Shin, and J. Paik, “Noise-adaptive spatio-temporal filter for real-time noise removal in low light level images,” IEEE Trans. Consum. Electron., vol.51, no.2, pp.648–653, 2005.
[5] J.-H. Lee, I.-Y. Shin, H.-G. Lee, T.-Y. Kim, and Y.-S. Ho, “Anti-shaking algorithm for the mobile phone camera in dim light conditions,” Pacific Rim Conf. Multimedia: Adv. Multimedia Inf., pp.968–973, 2009.
[6] G. Petschnigg, M. Agrawala, H. Hoppe, R. Szeliski, M. Cohen, and K. Toyama, “Digital photography with flash and no-flash im-
age pairs” SIGGRAPH, pp.664–672, 2004.

[7] R. Sabaee-Rad and M. Aleksic, “Handshake characterization and image stabilization for cell-phone cameras,” Proc. SPIE-IS&T Electronic Imagery, vol.7241, no.1, pp.72410V-1–72410V-15, Jan. 2009.

[8] L.-B. Luo and J.-W. Chong, “Real-time digital image stabilization for cell phone cameras in low-light environments without using frame memory,” ERTI J., vol.34, no.1, pp.138–141, Feb. 2012.

[9] W. Xu and J. Mulligan, “Performance evaluation of color correction approaches for automatic multi-view image and video stitching,” Proc. CVPR 2010, pp.263–270, 2010.

[10] E. Reinhard, M. Ashikhmin, B. Gooch, and P. Shirley, “Color transfer between images,” IEEE Comput. Graph. Appl., vol.21, no.5, pp.34–41, 2001.

[11] E. Reinhard, A.O. Akuyz, M. Colbert, C.E. Hughes, and M. O’Connor, “Real-time color blending of rendered and captured video,” Interservice/Indus. Training, Simul. and Educ. Conf., pp.1–9, 2004.

[12] W.-C. Kao, L.-W. Cheng, C.-Y. Chien, and W.-K. Lin, “Robust brightness measurement and exposure control in real-time video recording,” IEEE Trans. Instrum. Meas., vol.60, no.4, pp.1206–1216, April 2011.

Lin-bo Luo received his Ph.D. degree in Electronics and Computer Engineering from Hanyang University, Seoul, Korea, in 2012. He is currently a researcher in the Department of Science and Technology, China University of Geosciences, Wuhan. His research interests include image warping, digital image stabilization for image sensor and their embedded implementation.

Jun Chen received the B.S. degree in Electronics and Information Engineering from China University of Geosciences (CUG) in 2002 and the ME degree in Electronics and Communication Engineering from CUG in 2004. And now she is a Ph.D. student at the Department of Electronics and Information Engineering in Huazhong University of Science and Technology. Her research interest is image processing.

Sang-woo An received a B.S. degree in Department of Information and Communication Engineering from Hanyang University, Seoul, Korea, in 2010 and He is currently working on the unified Master’s and Doctor’s course in Department of Nanoscale Semiconductor Engineering from Hanyang University. His research interests include image processing, image stabilization.

Chang-shuai Wang received his B.S. degree in Communication Engineering from Harbin Institute of Technology University, Weihai, China, in 2010. He is currently working toward the M.S. degree in Electronics and Computer Engineering at Hanyang University, Seoul, Korea. His research interests include computer vision, image processing.

Jong-joo Park received a B.S. degree in Department of Information and Communication Engineering from Hanyang University, Seoul, Korea, in 2011 and he is currently on the Master’s course in Department of Information Display Engineering from Hanyang University. His research interests include image processing, image stabilization.

Ying-chun Li received his B.S degree in Information Security from Harbin Institute of Technology University, Weihai, China, in 2010. He is currently working toward the M.S. degree in Electronics and Computer Engineering from Hanyang University, Seoul, Korea. His research interests include wireless location, image processing and multimedia communication.

Jong-wha Chong was born in Nonsan, Korea, on March 10, 1950. He received B.S. and M.S. degrees in Electronics Engineering from Hanyang University, Seoul, Korea, in 1975, and 1979 respectively. He received his Ph.D. degree in Electronics & Communication Engineering from Waseda University, Japan, in 1981. Since 1981, he has been a professor at the Department of Electronics Engineering, Hanyang University, Seoul, Korea.