Applications of Flash and No-Flash Image Pairs in Mobile Phone Photography

Xi Luo

Stanford University

450 Serra Mall, Stanford, CA 94305

xluo2@stanford.edu

Abstract

The project explores various application of the Flash and No-flash image pairs that could be implemented in modern day mobile phones. By applying post-processing techniques using the image pair taken with an consumer grade cell phone, a better quality image can be constructed that can reduce noisy compared to the original as well as less color distortion compared to the image with flash.

I. Introduction

Rapid advances in modern day semiconductor processes and imaging sensors resulted in astonishing improvement of photo capturing capabilities of cell phones. May consumers question the necessity of a much heavier and expensive digital camera given the image quality offered by consumer-grade phones and their shallow learning curve.

Take the Apple iPhone 5S for example; it contains an 8MP camera with a pixel size of 1.5um and aperture of f/2.2. [1] It performs well under a broad range of lighting conditions. The shortcoming occurs under low light environment where the image can be noisy. (Fig. 1) Comparing the same scene captured with Canon 550D digital camera with 1 second exposure time, (Fig. 1) which shows much better image quality due to the low ISO setting of 200.

Figure 1: Same scene captured using iPhone 5S (left) and Canon 550D DSLR

II. Related Work

The Flash and No-Flash Image Pairs processing have been studied to show the combined image retains more detail than the No-Flash image and better color representation compared to the Flash image in study by Petschnigg et al. [2]. One caveat is that this study used a digital camera that is able to manually adjust ISO, f-stop, and exposure time to capture the image source. This is rarely possible while using the cell phone camera because the device automates the settings. The project aims to study the effectiveness of the same techniques but with a cell phone because that's where the application would provide huge potential in delivering better user experience.

Another post-processing applications, White Balancing, has been extensively investigated by Petschnigg *et al.* [2] and DiCarlo *et al.* [3]. It is an important step as it can be used as a way to remove color distortions that are caused by a light source with a specific spectrum density, such as the one from the integrated flash from the phone.

Special effects will also be investigated that build upon the post-processing chain where users of cell phones want to add special effects to the original image. This project attempts to implement a method that converts the image into a painting-like style while still retaining the original object boundaries. The blurring involved can be easily achieved by low pass filtering techniques used in noise reduction, such as bilateral filtering in Durand and Dorsey [4], but builds upon it with dynamic weight adjustments to the filter kernel parameters for better performance.

III. Problem Statement

Smaller sensor pixel size directly contributes to additional noise under low lighting conditions when high ISO is required to amplify the intensity due to fewer photons being captured per pixel. One possible solution to compensate the poor low lighting performance is to use the integrated flash on the phone, although this significantly improves detail preserved in the image, but doing so creates another problem, which is color distortion due to the flash.

IV. Approach

The project is divided into three distinctive parts. The first part deals with the denoising of the image pair and combining them into a single image that preserve the details from the Flash image while retaining the original object color from the No-Flash image. The second portion implement the White Balancing to simulate how the flash could impact the scene and illustrate how it can be used to restore the color to the combined image from the first part. The last section discusses possible techniques that can be implemented to add intentional effects to the image depending on user needs.

The image pair used in this project is acquired by using a consumer-grade iPhone, placed on the table at a fixed position, then use the flash on/off option to take the same scene by clicking on the screen to focus at the same object. The automated photo capturing of the iPhone does not allow manual setting of IOS or exposure time. The images are all captured with the iPhone, which automatically converted the file to .JPG image.

A. Denoising and detail transfer

Reducing noises in photographic images have been long studied. Many type of solutions exist that attempt to preserve edges while apply low filtering on the image, such as the bilateral filter in Durand and Dorsey [4]. This is a fast and non-iterative technique meaning that its implementation would not be resources intensive. This is desirable due to space and power requirements for cell phone electronics.

The project re-implemented the Joint Bilateral Filter discussed in Petschnigg *et al.*[2] (Fig. 2). This is a modification of the bilateral filter that acts as a low-pass filter with an edge stopping function when the adjacent pixel intensity is large from the image with flash. The following equation shows the computation of each pixel p in image A^{Simple} using the notation of Durand and Dorsey [4]:

$$A_p^{Base} = \frac{1}{k(p)} \sum_{p' \in \Omega} g_d(p' - p) \, g_r\left(A_p - A_{p'}\right) A_{p'} \quad (1)$$

where $k(p)$ is a normalization term:

$$k(p) = \sum_{p' \in \Omega} g_d(p' - p) \, g_r\left(A_p - A_{p'}\right) \quad (2)$$

The spatial weight is set by g_d based on the distances between the pixels where the edge-stopping function g_r compute the weight of the intensity differences. Both functions are Gaussian with standard deviation parameters σ_d and σ_r respectively.
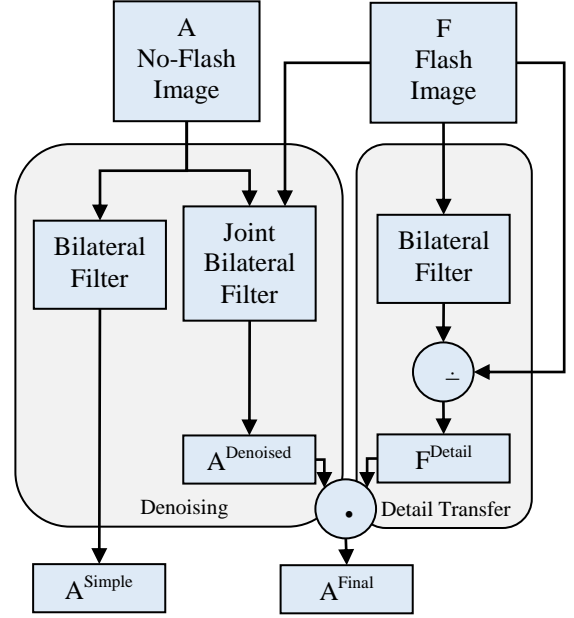

Figure 2: Denoising and Detail Transfer with joint Bilateral Filter

The bilateral filter is modified to use the edge-stopping function gr from image F. This creates the joint bilateral filter in [2]:

$$A_p^{Base} = \frac{1}{k(p)} \sum_{p' \in \Omega} g_d(p' - p) \, g_r\left(A_p - A_{p'}\right) A_{p'} \quad (3)$$

where $k(p)$ is also modified accordingly to use g_r from image F.

The σ for the bilateral filter is different from the joint bilateral filter. This is because since the edge-stopping function is from the F image, this means that a smaller σ is required for noise filtering while the bilateral which operates on the A image need to have a higher value in order to reduce noise.

The joint bilateral filter only act as an improvement over the noise reduction step, in order to take advantage of the flash image, the detail layer can also be extracted from image F due to more details retained than the no-flash image. The following equation computes the detail layer as shown in Shashua and Riklin-Raviv [5]:

$$F^{Detail} = \frac{F + \varepsilon}{F^{Base} + \varepsilon} \quad (4)$$

where F^{base} is the result of applying the bilateral filter on image F.

At low image intensities, F may contain noise that can generate spurious detail, the error term is used to reject such artifacts and also avoid division by zero case shown in Petschnigg *et al.* [2]. For this project the value is determined to be a small value of 0.1 experimentally.

B. White balancing

Both the flash and no-flash image can portrait the same image to a different color tone due to the low light level and introduction of the flash. For this project, the ground truth, or the desired image is the one that would form under moderate light level is present.

The study by DiCarlo *et al.* [3] used flash/no-flash pairs to estimate the scene illumination by using discrete point searches that closely match between the two images. Based on the comparison results from Petschnigg *et al.* [2], this project will re-implement the approach that calculates the albedo estimate at each pixel because it produced excellent result in the study.

The following calculation steps by Petschnigg *et al.* [2] compute the white balanced image for a given scale K:

Step	Calculation
1	$\Delta = F - A$
2	$Cp = \dfrac{Ap}{\Delta p}$ where $|A_p| < \zeta_1$ or $|\Delta_p| < \zeta_2$
3	c = mean value of C_p
4	$A^{wb} = \dfrac{1}{c} * Ap * K$

The above calculation assumes that the difference image in step 1 is a direct measure of the flash from the camera. But this may not be true as the surfaces of objects can be transparent or that the surface specular color may not match its diffuse color.[2] This means that the estimation carries inherent error in the albedo estimation.

The ζ terms are used to eliminate error that could come from very weak intensity pixels and both are set to 0.1 experimentally. A scaling factor K which can to applied to view the different amount of white balancing in the final image so the effect could be compared.

C. Dynamic edge preservation filtering

The goal here is to come up with a method that could introduce intentional effects into the original image, which a user may want to use for aesthetic needs.

For the scope of the project, the bilateral filter is modified which dynamically adjust the weight of the filter for each pixel depending on its intensity gradient, thus setting a high bilateral filter σ value is able to produce an image that smoothen the regional area while strongly enforce the preservation of object edges the image. The following steps summarizes the calculation steps:

1. For each pixel p, determine the mean intensity gradient based on the 8 surrounding pixels.

2. Create an intensity mask, with each pixel has the value of:
$$Mask(x,y) = \begin{cases} 1.5, & gradient < T \\ 0.01, & gradient \geq T \end{cases}$$

3. Apply the bilateral filter, but for each pixel the kernel σ is multiplied by the weighted mask value respectively.

From this calculation, the edge mask will ensure that the actual σ for an edge pixel the smoothing effect is kept to a minimum while push up the effect in local regions. The T (threshold) is used to polarize the mask values thus basket each pixel to either be an edge, or non-edge.

V. Evaluation

The following image pair is used to test the effectiveness of the implemented methods. (Fig. 3)
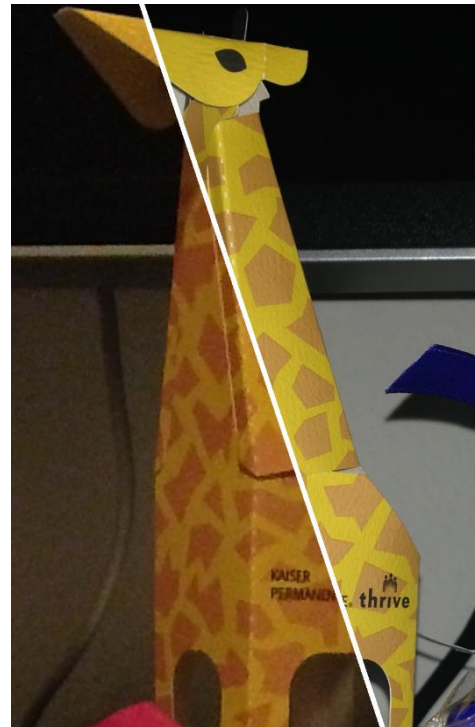

Figure 3: No-Flash (left) and Flash (right) image pair.

A. Denoising and detail transfer

First, the denoised image from the bilateral filter and the joint bilateral filter is compared to see if better noise reduction has been achieved for the second case. (Fig. 4)

Figure 4: Denoise with bilateral filter (left) vs with joint bilateral filter (Right).

The zoomed in shot provide better insight into the denoise quality comparison. (Fig. 5)

Figure 5: Denoised image from the bilater filter (top), the joint bilateral filter (middle), and the original no-flash image (bottom).

The bilateral filter at the top clearly shows the smoothing effect to reduce noise, however some details were lost. The joint bilateral filter reduced the noise compared to the original while preserved more edge information as intended.

Next step is to look at the calculated detail layer and the combined image with the detail transfer. (Fig. 6)

Figure 6: Detail layer (left) and combined image with detail transfer (right).

The detail layer shows that surfaces which reflected greater amount of light, for example the silver edge of the computer screen running horizontally across the image has a very heavy weight in the detail layer due to bigger intensity difference which was used in the calculation. Also the combined image contains a non-existent ghosting at the center right side of the result image. This is due to the shadow caused by the flash which was transferred to final image.

Another closer look at the combined image with the original non-flash image. (Fig. 7)

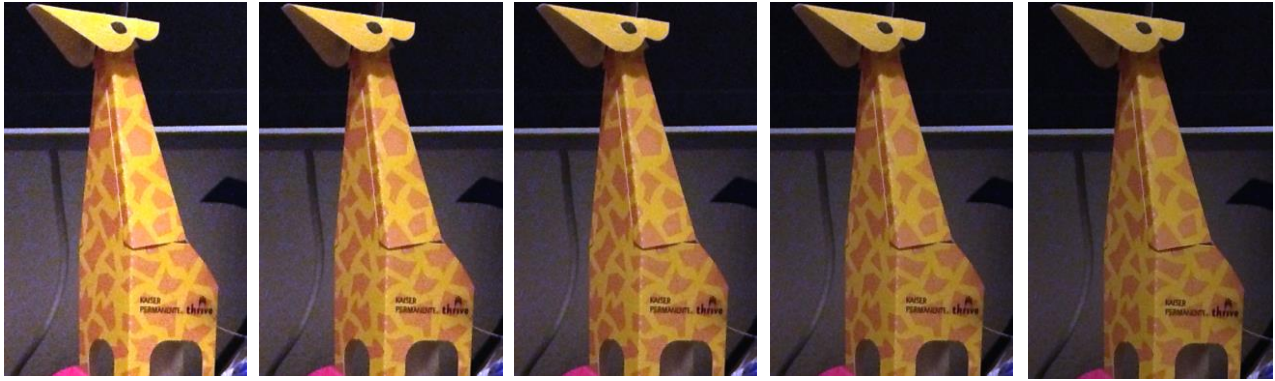Figure 7: Original no-flash image (top) and combined image with detail transfer (bottom).

Figure 8: Scaling factor k = 3to 5 from left to right in 0.5 steps.

This shows that although the detail is transferred to the final image, there is a misalignment between the detail layer vs the original image. This is a direct result of the source image pair captured by the iPhone which is not exactly aligned. It demonstrates that tiny amount of misalignment could manifest into huge error in the combined image, the effect may not be noticeable when the image size is small but once blown up, the quality degrades greatly.

B. White balancing

The images (Fig. 8) illustrate the result from the white balance calculation between scaling factor *K* of 3 to 5 where the 'white' illuminance trend is illustrated.

The middle represents the range where the white balance is not exaggerating which makes it unappealing to the eye. To the far right where the tone is dark is also not optimal. Since there is no one consensus as to the 'perfect' white balance image, it depends hugely on personal taste and the scene itself.

C. Dynamic edge preserving filtering

With an experimentally design filtering σ of 20 and kernel radius of 5, the mask layer shows the gradient edge detected and the absolute pixel mask value assigned. (Fig. 9) A conventional bilateral filtering is performed to serve as a base line and it is compared to the dynamic adjustment method. (Fig. 10)
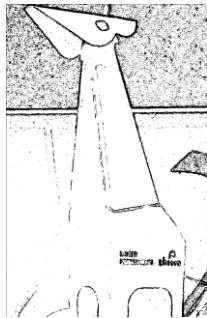

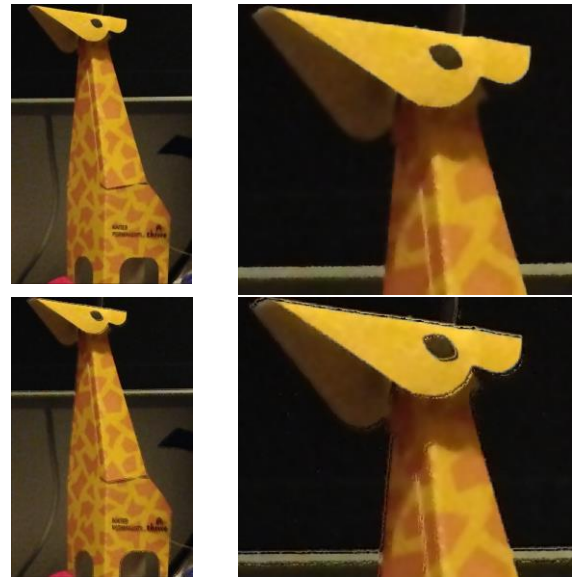Figure 9: Gradient edge mask used to weigh the kernel σ.

Figure 10: Bilateral filer result at the top row, and the dynamic adjustment to the bilateral filter result at the bottom row.

The enforcement edges showed that in the result image, the determined location produces the originally thought after special effect which is to add a painting style manipulation to the original but at the same time return the shape of the objects.

To determine the applicability of this modified filter, additional image were tested that include features such as night scene and high spatial frequency objects to simulate what the user would see on their daily photos taken with the phone. (Fig. 11, 12)

VI. Discussion

Experimental data showed promising results from the joint bilateral filter technique where one could reconstruct images that contains more high frequency details and at the same time retain the noise filtering of the bilateral filter.

Figure 11: Original image (left) and Special effect implementation using the dynamic edge preservation filtering (right).

The study uncovered the issue of subtle misalignment during the capture of the image pair. This places constraint on the lower limit of the shutter speed that can potentially benefit from this method. Ideally, the image pair should be captured when the user push the capture button once where the camera immediately take the no-flash photo and then the flash photo to prevent unwanted camera movement.

White balancing results demonstrated the usefulness of this process. From the image pair, one could fine tune any 'white' illuminance in-between as well as outside the original range. The technique can be used to restore image balance due to distortion of specific light sources.

Regarding the dynamic edge preservation modification to the bilateral filter, the intended special effects were achieved where it simulate a 'portraitfication' of the original image but still retrains the edges of the individual objects within the scene. The idea is inspired by similar art styles which were extensively used in *Boarderlands* the video game. Since the goal is to provide the cell phone users with a way to manipulate their images, the same special effect may not appeal to all. Thus, it is imperative to include a bundle of different algorithms that can produce drastically different special effects.

VII.Future Work

One process not implemented in this project is the ghost shadow removal that is from the flash image. Without this step, the final image may retain a weak but still noticeable artifact.

The inherent processing steps must be able to run with available hardware resources on the cell phone. This means that finding multiple solutions that can shared one set of hardware implementation is most desirable.

Figure 12: A closer look at the original image (right) vs. dynamic edge preservation filtering (left).

References

[1] Apple Inc.. (2016, Feb 26). Compare iPhone Models [Online]. Available: http://www.apple.com/iphone/compare

[2] G. Petschnigg, M. Agrawala, H. Hoppe, R. Szeliski, M.Cohen, and K. Toyama, 2004. Digital Photography with Flash and No-Flash Image Pairs. ACM SIGGRAPH

[3] J. M. Dicarlo, F. Xiao, and B. A. Wandell, 2001. Illuminating Illumination. Ninth Color Imaging Conference, pp. 27-34.

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

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