CS-663 Course Project
Digital Photography with Flash and No-Flash Image Pairs

Soham Naha (193079003)
Akshay Bajpai (193079002)
Mohit Agarwala (19307R004)

December 7, 2020

1 Problem Statement

The aim is to implement a method to use joint bilateral filtering on a flash and non flash image pair for achieving the following -

- Image denoising while preserving the ambient condition details
- Detail transfer between flash and no-flash image pairs
- Red eye removal
- Continuous flash variation

2 Introduction

Lighting is an integral part of any photographic image, to capture and reproduce the visual richness of the ambience. But in low-light environment, there are hindrances. If the exposure time is kept too high, a camera shake results in motion blur, if aperture is allowed to let more light in, it reduces the depth of field information, while the other option of changing the ISO, which might result in increase in noise in images.Flash photography can be an option to use in such cases, but although it solves the problems of low-light, it has a negative impact on the lighting characteristics of the environment by destroying the ambient mood. In addition, flash photography additionally introduces speculations, unwanted harsh shadows, red-eye which renders images that seldom depict the ambient settings of the scenes. The paper [1] takes both the flash and ambient (no-flash) images to mitigate the short-comings of each other, by combining features from both images.
The intuition behind several of the algorithms followed in [1] is that while the illumination from a flash may change the appearance of the scene, it also increases the signal-to-noise ratio (SNR), and thus can help in better estimate of high frequency detail, however due to power law, the noise in higher frequencies is considerably more.

3 Block Diagram

The block diagram for noise removal is given below:

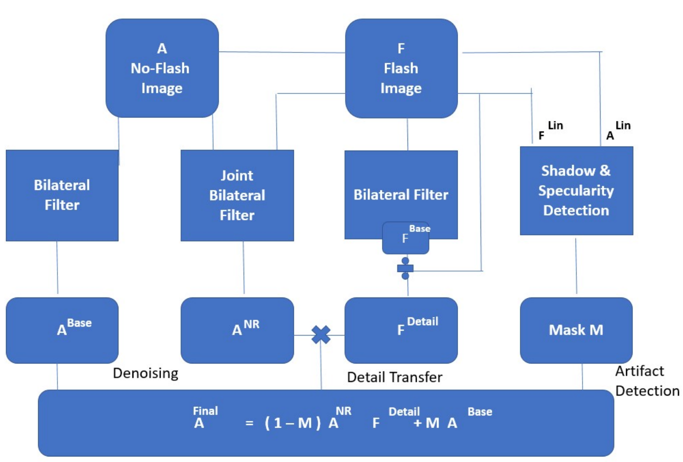


Figure 1: Block diagram for the noise removal algorithm followed

4 Algorithm

4.1 Algorithm for Denoising and Detail Transfer

4.1.1 Denoising

Although there are lot of techniques available for de-noising images, the authors of [1] have resorted to bilateral filtering and a variant of it that utilizes both the flash and ambient images for noise removal.

Bilateral Filtering is a fast, non-iterative algorithm used to determine the present pixel value depending on the neighbouring pixels having the same intensities by combining a low-pass filter with a edge stopping function, given by:

$$A_p^{base]} = \frac{1}{k(p)} \sum_{p' \in \Sigma} g_d(p' - p) g_r(A_p - A_{p'}) A_{p'} \tag{1}$$

where $A_{p'}$ is the pixel at point p' neighbourhood Σ and $k(p)$ is a normalization term:

$$k(p) = \sum_{p' \in \Sigma} g_d(p' - p) g_r(A_p - A_{p'}) \tag{2}$$

g_d and g_r are respectively the spatial domain Gaussian filter and the edge-stopping intensity domain Gaussian kernels controlled by their standard deviations σ_d and σ_r respectively.

Joint Bilateral Filter is a modification of the vanilla bilateral filter that the intensity domain edge-stopping Gaussian is computed using the flash image F, as given by:

$$A_p^{NR} = \frac{1}{k(p)} \sum_{p' \in \Sigma} g_d(p' - p) g_r(F_p - F_{p'}) A_{p'} \tag{3}$$

where $k(p)$ is also changed accordingly. Here, $A^{NR]}$ is the noise-reduced version of A. This method relies on the flash image as an estimate for the ambient image, and can incorporate flash shadows and specularities which are removed using the subsequent methods using the following equation:

$$A^{NR'} = (1 - M)A^{NR} + MA^{base} \tag{4}$$

4.1.2 Detail Transfer

is the method that transfers the details from the flash images that were removed by the denoising action of the bilateral filtering action carried out on the flash image. Flash images can reveal a additional surface details not visible in the ambient light setting, and can illuminate the areas that were otherwise in the shadow regions of the ambient setting. The detail layer is calculated as :

$$F^{Detail} = \frac{F + \epsilon}{F^{base} + \epsilon}, \tag{5}$$

where ϵ is a constant used to reduce artifacts (such as division by zero), and F^{base} is the bilateral filtered output of the flash image F. To transfer the details of the flash image, simple point wise multiplication is done with joint bilateral output. So, the modified equation for the denoising output is:

$$A^{NR'} = (1 - M)A^{NR}F^{Detail} + MA^{base} \tag{6}$$

4.1.3 Detecting Flash Shadows and Specularities

is an intermediate step for noise removal that creates a mask to remove the shadows and other specularities(showing parts of image having high reflectivity) introduced due to the flash setting. There are four issues that arises due to flash:

- surfaces that do not refelct any light are to be considered as a shadow region in both the flash and ambient settings.

- distant surfaces where the flash light was unable to reach is always a shadow in both the setting.
- noise introduced causes non-zero values in the shadows, which can be removed by thresholding (using τ_{shad}, a parameter) the difference image between flash and ambient settings.
- inter-reflections due to flash also causes non-xero values in shadows,which also removed using thresholding.

The mask may contain some of the noisy elements of the difference image, which are then removed using the methods of erosion and dilation and to smoothen the mask to feather the edges using a Gaussian Blur.

4.2 Continuos Flash Adjustment

As the amount of flash in an image is of aesthetic nature, depending on the person, the paper [1] explores a simple method to interpolate between flash and no-flash images using the formula:

$$F^{Adjusted} = (1-\alpha)A + \alpha F \tag{7}$$

where A is the ambience image, F is the flash image and the parameter α can be adjusted by first converting the image to YCbCr space.

4.3 Red-Eye Correction

Normal red-eye detection methods incorporates heuristic and machine learning methods to localize the red-eye region and then darken these pixels. The paper [1] that suggests an algorithm using the change of pupil color in flash and ambient images by converting the images to YCbCr color space. First, the luminance information is decorrelated from the chrominance to infer the relative redness measure given by:

$$R = F_{C_r} - A_{C_r} \tag{8}$$

segment the regions where R¿τ_{eye}, a chosen threshold and in order to limit the consideration of noise in the mask, seed pixels which satisfy the following are chosen:

$$R > max(0.6, \mu_R + 3\sigma_R) \text{ and } A_Y < \tau_{Dark}, \text{ where } \tau_{Dark} = 0.6 \tag{9}$$

On the basis of the above thresholds a mask is obtained, which undergoes morphological transformation to remove noise and then *"glints"* are removed using the following steps:

- Consider all the non-zero pixels in our mask
- around each such non-zero pixel we consider a window of size (100,100)
- Within each of these windows the center region, a window of (47,47) is assumed to be the pupil and set as zero
- After making them zero, if in the (100,100) window, there are values that are still one, we consider that pixel not to be a part of the red-eye region
- else, we substitute the pixel-value to 80% of its luminance value.

5 Results

5.1 Noise Removal

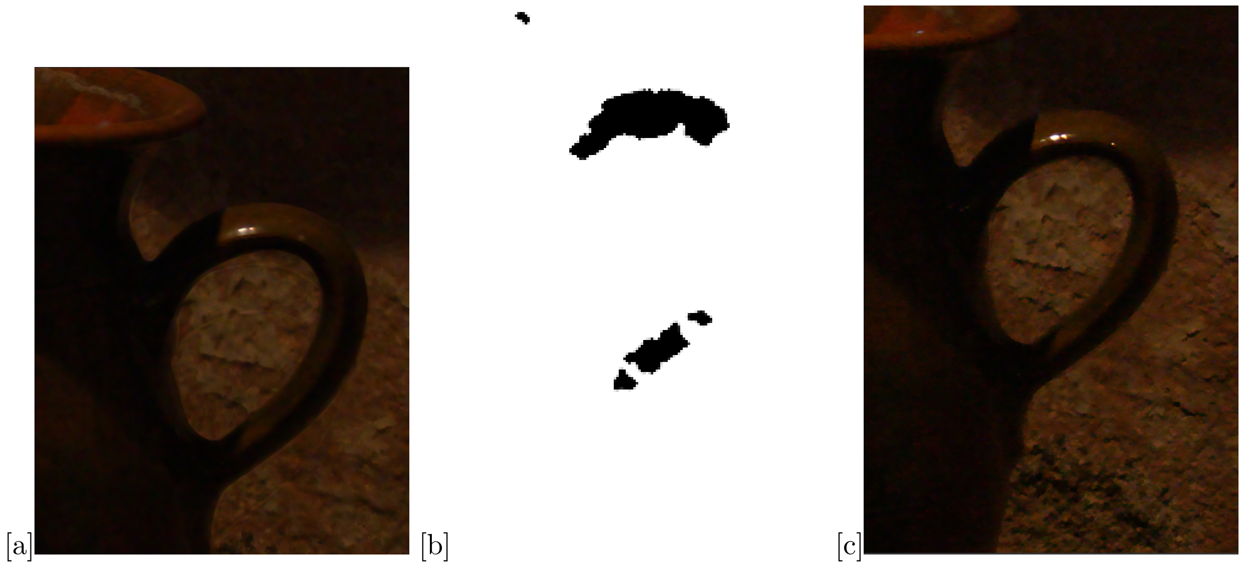


Figure 2: (a) Shadowing artifacts present in denoised image without using the mask (b) Shadowing and Specularity mask calculated (c) Shadowing artifacts present in denoised image using the mask

Figure 3: (a) Noisy ambient image (b) Denoised image

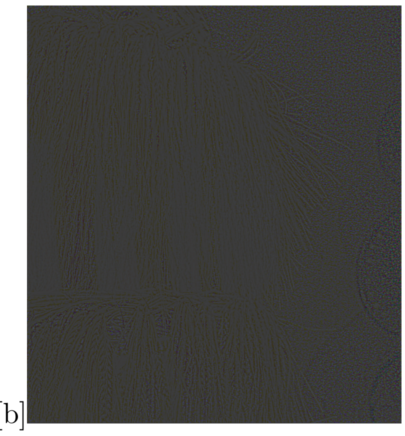
[a]  [b] 

Figure 4: (a) Detail Transfer of lamp image, to the left the joint bilateral filtered image, to the right after detail transfer (b) Detail image

5.2 Red-Eye Removal

Figure 5: Red Eye Removal

6 Conclusion

The algorithms specified to utilize the pros of both the flash and no-flash images can be used to create an ambient image that is denoised and captures the details of the ambient scene. But there are shortcomings that the parameters of the algorithms are to be tweaked manually, depending on the effects desired by the user.
The bilateral and joint bilateral filters, although non-iterative, they are non-linear and hence inseparable and cannot be used as convolution to make the operations fast. Processing each of the color channels separately also adds to the computational overhead.
The luminance is the only thing that changes between the flash and ambient images, keeping the environment setting the same. Motion blur cannot be handled and requires de-blurring.

References

[1] Georg Petschnigg et al. "Digital Photography with Flash and No-Flash Image Pairs". In: *ACM Trans. Graph.* 23.3 (Aug. 2004), pp. 664–672. ISSN: 0730-0301. DOI: `10.1145/1015706.1015777`. URL: `https://doi.org/10.1145/1015706.1015777`.