Indian Institute of Technology, Bombay

Project Report

Digital Photography using Flash and No-flash Image Pair

Course: Digital Image Processing (CS663)

Submitted By

Himanshu Meena (Roll No. : 150070027)
Raj Gite (Roll No. : 203050092)
Shreyas Narahari (Roll No. : 203050037)

(Academic Year: 2020-2021)

Contents

1 Introduction

Digital photography has made it possible to quickly and easily take a pair of images of low-light environments: one with flash to capture detail and one without flash to capture ambient illumina- tion. We present a variety of applications that analyze and combine the strengths of such flash/no-flash image pairs.

Our applications include,

- Denoising and detail transfer (to merge the ambient qualities of the no-flash image with the high-frequency flash detail)
- White-balancing (to change the color tone of the ambient image)
- Continuous flash (to interactively adjust flash intensity)

We demonstrate how these applications can synthesize new images that are of higher quality than either of the originals.

2 Denoising and Detail Transfer

.

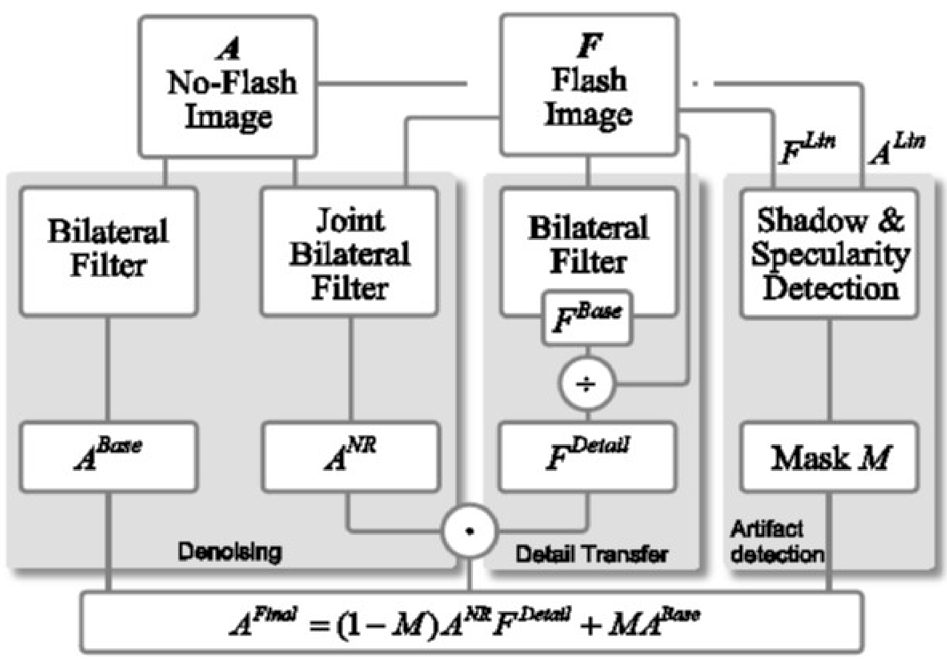


Figure 1: Overview of denoising and detail transfer

Consider following conventions:-

A : No-flash image or Ambient image
F : Flash image
A^{Base} : Basic bilateral filter applied to no-flash image
A^{NR} : Joint bilateral filter applied to no-flash image
F^{Base} : Basic bilateral filter applied to flash image
F^{Detail} : Contains the details obtained from flash image
M : Mask

For denoising ambient image Joint Bilateral Filtering is used and its output A^{NR} is multiplied by F^{Detail} to get a detailed ambient image. But there are some artifacts introduced in the regions of shadow and specularity present in flash image in this detailed ambient image. To remove them a mask built from F and A representing those regions is used. We replace those regions by respective regions present in the A^{Base} by first subtracting them from detail ambient image using (1 - M)$A^{NR}F^{Detail}$ and then adding MA^{Base}. Thus we get the final detailed denoised ambient image A^{Final}.

$$A^{Final} = (1 - M)A^{NR}F^{Detail} + MA^{Base}$$

2.1 Denoising

Applying Basic Bilateral Filtering to ambient image over blur and under blur some regions of the image where lighting conditions are not good and thus weak edges gets smoothen out as seen in Fig. **??**. To prevent this we use Joint Bilateral Filter, wherein flash image is used to give weight to the gaussian and thus where ever in the image there are weak edges highlighted by the flash light gets detected correctly and no over blurring happens as seen in Fig. **??**. From the Fig. **??** its clear that Joint Bilateral Filter is capturing many weak edges that Basic didn't captured.

.

Figure 2: Flash Image of Carpet

.

Figure 3: Ambient Image of Carpet

.

Figure 4: Basic Bilateral Filtering of Ambient Image of Carpet

.

Figure 5: Joint Bilateral Filtering of Ambient Image of Carpet

.

Figure 6: Difference between Basic and Joint Bilateral Filtered Ambient Images

2.2 Flash-To-Ambient Detail Transfer

We transfer high-frequency detail from the flash image to the denoised ambient image, since this detail may not exist in the original ambient image. While the joint bilateral filter can reduce noise, it cannot add detail that may be present in the flash image. Moreover, the flash typically provides strong directional lighting that can reveal additional surface detail that is not visible in more uniform ambient lighting.

To transfer this detail we begin by computing a detail layer F^{Detail} from the flash image as the following ratio:

$$F^{Detail} = \frac{F + \epsilon}{F^{Base} + \epsilon}$$

where F and F^{Base} are Flash Image and Bilateral Filtered Flash Image, respectively.

At low signal values, the flash image contains noise that can generate spurious detail. We add ϵ to both the numerator and denominator of the ratio to reject these low signal values and thereby reduce such artifacts (and also avoid division by zero). In practice we use $\epsilon = 0.02$ across all our results.

.

Figure 7: Basic Bilateral Filtering of Flash Image of Carpet, F^{Base}

.

Figure 8: Flash Details of carpet, F^{Detail}

F, flash image of carpet is shown in Figure 2. F^{Base} of carpet is obtained from F and

is shown in Figure 11. Using the above two, flash details of carpet, F^{Detail} is obtained as shown in Figure 12.
To transfer the detail, we simply multiply the noise-reduced ambient image A^{NR} by the ratio F^{Detail}. So, the final detection algorithm would be,

$$A^{Final} = (1 - M)A^{NR}F^{Detail} + MA^{Base}$$

Here, M is the Shadow and Specular Mask which identifies region where our $A^{NR}F^{Detail}$ gives poor results. Also, A^{Base} is Bilateral Filtered Ambient Image as before. For the ambient image of carpet as shown in Figure 3 we get the final filtered image, A^{Final} as shown in Figure 9.

.

Figure 9: Final carpet filtered ambient image with detail transfer, A^{Final}

2.3 Detecting Flash Shadows and Specularities

Light from the flash can introduce shadows and specularities into the flash image. Within flash shadows, the image may be as dim as the ambient image and therefore suffer from noise. Similarly, within specular reflections, the flash image may be saturated and lose detail. Moreover, the boundaries of both these regions may form high-frequency edges that do not exist in the ambient image. To avoid using information from the flash image in these regions, we first detect the flash shadows and specularities.

- **Flash Shadows**
  Since a point in a flash shadow is not illuminated by the flash, it should appear exactly as it appears in the ambient image. To compare flash and ambient image, we use the luminance measure of them. The luminance of an image, let's say for ambient image A, luminance ambient image A^{Lin} is given by,

  $$A^{Lin} = 0.2989A_R + 0.587A_G + 0.114A_B$$

  Ideally, shadow mask would be pixels where $F^{Lin} - A^{Lin} = 0$ but noise causes nonzero values within shadows and inter-reflection of light from the flash causes non-zero values within the shadow. To deal with this, we add a threshold when computing the shadow mask by looking for pixels in which the difference between the linearized flash and ambient images is small,

  $$M^{Shad} = \begin{cases} 1 & \text{if } F^{Lin} - A^{Lin} \le \tau_{Shad} \\ 0 & \text{otherwise} \end{cases}$$

  τ_{Shad} is adjusted such that all the flash shadow regions are properly captured.

- **Flash Specularities**
  Specular regions should be bright in luminance flash image, F^{Lin} and should therefore saturate the image sensor. Hence, we look for luminance values in the flash image that are greater than 95% of the range of sensor output values.

We form our final mask M by taking the union of the shadow and specular masks. We then blur the mask to feather its edges and prevent visible seams when the mask is used to combine regions from different images. .

Figure 10: Detail Transfer without Mask

.

Figure 11: Mask identifying the shadow

.

Figure 12: Detail Transfer using Mask

3 White Balancing

In this the author's propose a method where we estimate and correct the illumination using flash/no-flash image pairs. For some intuition behind this the flash image being used for white balancing is that the flash image can be taken as the ambient(no-flash) version of the image with a point source of white light to illuminate the scene.

Initially flash and ambient images are linearized using the formula given in (1) to make the images linear so that the numerical intensity values correspond to the perceived intensity when we manipulate the images.The flash and ambient image considered for processing is given in

$$Linearized\ Image = 0.2989 * R + 0.5870 * G + 0.1140 * B \tag{1}$$

Figure 13: Image on the left is with the flash on and right is the ambient image

Then the difference of images is found out from equation (2) to get the illumination due to flash. This is proportional to the surface albedo at every pixel p. This is an albedo estimate Δ has unknown scale, because both the distance and orientation of the surface are unknown.

$$\Delta = F^{Lin} - A^{Lin} \tag{2}$$

The ambient image A and the scaled albedo Δ have value of A_p and Δ_p at pixel p. We can estimate the ambient illumination at the surface with the equation (3)

$$C_p = \frac{A_p}{\Delta_p} \tag{3}$$

After this the C_p contains the estimated color which has to be reduced to a single value per channel using mean. To make this mean robust we ignore pixels corresponding to C_p for which $||A_p|| < \tau_1$ for every channel because this will contain very small values and $\Delta_p < \tau_2$ because this will lead to very large values. Here τ_1 and τ_2 are set to about 2% of range of color values. Finally, the ambient color estimate is calculated by taking the per channel average. As in [Figure 15]. The ambient image is scaled by the ambient color estimate (c) per channel to obtain the final white balanced image. [Figure 16]

$$A_p^{WB} = \frac{1}{c} A_p \tag{4}$$

Figure 14: Estimated ambient illumination colors C

Figure 15: Estimated ambient illumination colors c

Figure 16: White balanced image

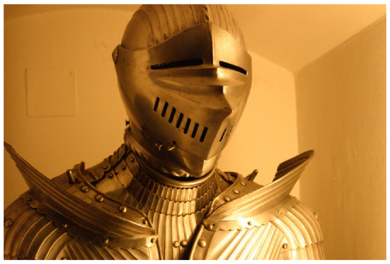
4 Continuous Flash Adjustment

When taking a flash image, the intensity of the flash can sometimes be too bright, saturating a nearby object, or it can be too dim, leaving mid-distance objects under-exposed. With a flash and non-flash image pair, we can let the user adjust the flash intensity after the picture has been taken.

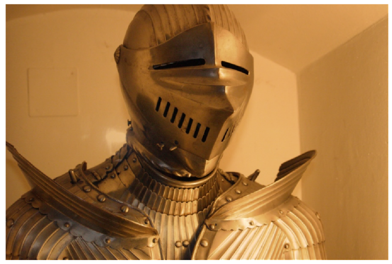
The most effective scheme is to convert the original flash/no-flash pair into *YCbCr* space and then linearly interpolate them using, Here α is from [0.0, 1.0]

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

Here $\alpha = 0.0$ is the non-flash image and $\alpha = 1.0$ is the flash image. The image has been converted to a linear color space such as *YCbCr* as it is a linear color space where numerical intensity values correspond proportionally to their perceived intensity.

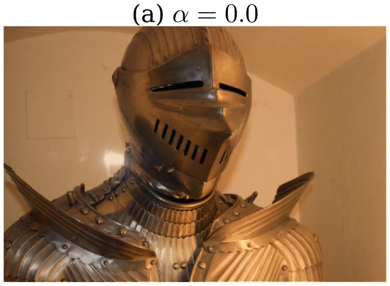


(a) $\alpha = 0.0$

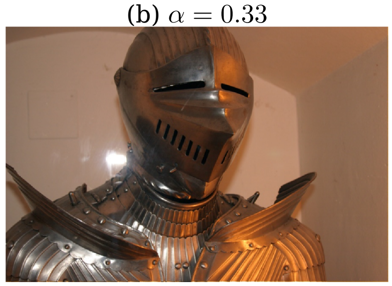


(b) $\alpha = 0.33$

(c) $\alpha = 0.5$



(d) $\alpha = 0.66$



(e) $\alpha = 1.0$

Figure 17: The helmet images for different values of α ranging from [0.0, 1.0]

5 References

1. Digital Photography with Flash and No-Flash Image Pairs, *Georg Petschnigg, Maneesh Agrawala, Hugues Hoppe. Richard Szeliski, Michael Cohen, Kentaro Toyama*