Shadow Extraction

Karl S. Gudmundsson, Jon Atli Benediktsson, Francis S. Cagatin
University of Iceland, Hjarðarhaga 2-4, Reykjavik, Iceland 107
E-mail: karlsg@hi.is

Abstract. This paper examines the shadow extraction problem associated with satellite images; namely the fact that images taken at different time include different shadow component. In the present research we attempt to define the shadow component of an image in a controlled lab environment in terms of the phase of the Fourier spectrum of the image, devise a novel method for extracting simple shadow component from the image, and to create a meta-image suitable for change detection post processing.

1. Introduction

Shadows are natural everyday occurrences. Whenever an object blocks the path of light, a shadow is cast behind it. As such one of the most evident properties of shadows is that it darkens the surface upon which it is cast. Shadows are different from dark surfaces in the sense that it is caused by another object. Also, since shadows are formed by an object occluding the path of light, the shape of the shadow follows the projection of the silhouette of that object. This property is exploited in cases where the object may not be entirely seen as in aerial images. Irvin et al. extracted key features of buildings from the shadows that each made. According to Jiang and Ward there are two types of shadows: projected shadows and self-shadows. A cast shadow is projected by an object to a surface belonging to a different object. A self-shadow on the other hand is a region of the object which receives little or no light due to its position in relation to the light source. In this paper, we would concentrate on cast shadows for the reason that we deem self shadows as part of the object. All succeeding references to shadow would refer to cast. Figure 1 illustrates the difference between these two types of shadows.

![Figure 1. Cast and self-shadow](image-url)
As explained earlier, shadows introduce a degree of variation on the surface it falls on. It alters the color and shape of its occupied region. This poses a serious problem in computational algorithms. Straightforward object detection methods, for example, would mistakenly mark shadow boundaries as object edges. Consequently this problem has been the subject of growing interest in the image processing world.

Salvador classified shadow detection techniques into two groups: model-based and property-based techniques. Model-based techniques rely greatly on a priori knowledge of the image: the geometry of the scene and illumination direction. Usually, gray-level images are analyzed and the luminance information is used. Typically, shadows are the darkest parts of the images; therefore, shadows were determined by thresholding. The geometric relationships of the shadow candidates and their shadowcasting objects are then analyzed to establish the presence of the shadow. As these relationships can vary greatly from one subject to another, model-based techniques are custom-designed for every situation. They are often not adaptable across applications. Property-based techniques, on the other hand, take advantage of a combination of spectral and geometric properties of imaged shadows. This approach improves upon the model-based techniques and expands the use of each to more applications. Salvador et al., for instance, used the invariance of the hue to shadows in order to extract the shadows from the image. Comparing the detected edges in the original image and the extracted hue image, they were able to remove the shadows while preserving the object.

Both of the shadow detection approaches previously discussed work solely on the time-domain. Ramamoorthi et al. presented another way of analyzing shadows. They used Fourier basis functions and convolutions for their research. While their work was more concerned on representing shadows rather than detecting them, it offers a significant step in analyzing shadows in the Fourier domain.

Representing images in the Fourier domain allows manipulations in the magnitude and phase. Most operations however, focus on the magnitude. Lim and Oppenheim proved the importance of the phase in the Fourier representation of images. They showed that much of the information is stored in the phase. More surprisingly, they showed that if the phase information of an image was combined with the magnitude information from an ensemble average of a group of unrelated pictures, the reconstructed picture was almost identical to the original picture. Furthermore, Horner and Gianino illustrated the fact that a phase-only filter performs better discrimination than a classical matched filter and a magnitude-only filter for certain scenes. And while this has been applied to optical processing, not much work has been done on this for image processing.

This research would also be working on the Fourier domain. We would take the same direction as the inverse-filtering technique introduced by Karim and Awwal. They implemented the inverse-filter to recover information from a motion-degraded image. Using the velocity of the object and its direction, the degradation function is computed. This in turn is applied to the degraded image in order to retrieve the original image. Our goal is to define the variables of the shadow component in an image in terms of the phase of the Fourier spectrum of the image and subsequently devise a novel method for extracting complex shadow components.

2. Theory
Motion degraded images in coherent light imaging system have been corrected in the past by considering the Fourier transform of an image

\[
O(u_x, u_y) = I(u_x, u_y)H(u_x, u_y)
\]

where \(O(u_x, u_y)\), and \(I(u_x, u_y)\) are the Fourier transforms, respectively of the output and input images while \(H(u_x, u_y)\) is the coherent transfer function. To correct for the motion degradation a filter an inverse filter:
needs to be computed and multiplied point by point to the output spectrum. It is tempting to consider the shadows in incoherent light images in analogous way. However, the frequency spectrum of the image intensity is fundamentally different. The incoherent system is linear in intensity while the coherent system is highly nonlinear in intensity [Goodman]. The image intensity for coherent system is given by the convolution equation 3.

\[ I_i = |h \otimes U_g|^2 \]  

However, when dealing with incoherent system the image intensity is given by equation 4.

\[ I_i = |h|^2 \otimes I_g = |h|^2 \otimes |U|^2 \]

Using the analogy of motion degraded coherent light images one might therefore consider the shadow present in the incoherent images to be the degrading factor. Which is be given by the convolution equation 5.

\[ I_i = |h|^2 \otimes |U_{main} + U_{shadow}|^2 = |h|^2 \otimes |U_{main}|^2 + |h|^2 \otimes |U_{shadow}|^2 \]

Furthermore, the frequency spectra incoherent can be written directly

\[ F\{I_i\} = [H \times H]G_{main} \times G_{main} + [H \times H]G_{shadow} + G_{shadow} \]

If we consider the first term of equation (6) to be A and the second term to be B the equation for the frequency spectra can be re-written as:

\[ F\{I_i\} = A + B \]

Thus we would like to find \( B^{-1} \).

3. Experiment

For this paper, we used a rock as the subject. We positioned the subject on a non-reflective sheet of cloth. It prevents any light from reflecting back to the subject. Any light falling on the subject or on the surface that is not coming directly from the light source would introduce irregularities in the acquired pixel levels. We would like the image to be as ideal as possible and therefore we address any reflections on the scene. The target is represented in Figure 2.
Directional light sources tend to produce inconsistent illumination. It is brightest at its focus, and drops as you go farther from it. In order to mimic natural lighting conditions, we used a photographic umbrella. Most commonly used in studio photography, this type of umbrella diffuses the light from a directional light source. As a result the light that reaches the subject appears soft and smooth. This would ensure that the natural umbra and penumbra of the shadows would be similar or at least close to one that would be produced in ordinary lighting.

Both the camera and light source are mounted on tripods. This makes it easier to adjust the position of the light source. The camera, on the other hand is fixed directly above the target. This creates a constant vantage point from which to compare the images from different lighting positions. It also prevents the camera from casting its shadow on the target. The position of the light source would represent the changing time of day, i.e., movement of the sun across the sky. The experimental setup is shown in figure 3.

The typical movement of the light source is shown in Figure 3. The movement of the light source is constrained to 1.1m from its origin. This limits its angle with respect to the target at 30.74°. Clearly approximately 60° movement of the light source represents only few hour changes in daylight.

For every scene one greyscale image is taken. Originally each image is 1288 x 1936, but they are scaled by 4, to 322 x 484 pixels for the purpose of speeding up the simulation. After which the images are then re-sampled to have only 128 gray-levels instead of the original 256 gray-levels. The resizing and re-sampling facilitates faster and simpler processing of data. Figure 4 depicts the images acquired.
4. Discussion

In order to better understand the relationship between the shadows and the phase of the signal an algorithm was devised which varied the phase frequency spectrum of the image incrementally in small steps from 0 to $\pi$. The algorithm was simulated on computer while collecting simple statistical information for each step. The information collected was the standard deviation, minimum, maximum, and the mean values of the pixel intensity of the resulting image in the space domain. Figure 5 shows the standard deviation for all three resulting images as the phase is varied while figure 6 shows the mean value.

![Experiment images with varying light source angle](image)

**Figure 4.** Experimental images with varying light source angle

![Standard Deviation Comparisons of the Three Images](image)

**Figure 5.** Standard Deviation of resulting images as the phase is varied
Interestingly, for all three images, varying the phase resulted in smoothing of the well illuminated parts of the image while enhancing the less illuminated portions of the image. This modulation improves significantly the contrast between the shadow boundaries and the background. The boundary between the object and the shadow is somewhat improved but not as significantly. The images presented in this paper are all scaled into the same gray scale value of 128. In figure 7, the left-most figure shows the image with the phase modulation of 8/100 * π which we considered providing the best value for the shadow extraction process. The image in the middle shows a meta image after thresholding the resulting image to define the shadow boundaries. Finally the image on the right shows a composite image where the boundaries of the original shadowed area is showed and the background texture has been built up using the background from the left and right images in figure 4.

5. Conclusions
In this simple experiment we have attempted to understand how the phase of the image in the frequency spectrum is related to the shadow component of the image in the space domain. This experiment was inspired by the fact that motion degraded images in coherent imaging system can be restored by a simple inverse filter. Shadows are somewhat similar to the smearing effect in the motion degraded images. However, for incoherent system the inverse filter, for shadow extraction, if it exists is more complicated and will probably rely on both the amplitude and the phase of the frequency spectrum of the image. Phase modulation can clearly be used to enhance the shadow extraction process in simple scenes such as we have used in this experiment. Future work in this area is to
consider the phase and the amplitude of the signal at time and to investigate if more complicated scenes show similar results.

References
[1] Salvador E 2004 Shadow Segmentation and Tracking in Real World Conditions (Universit`a di Trieste) 36-40
[2] Irvin R and Mckeown D 1989 IEEE Trans. Syst., Man, Cybernet 1564–75
[3] Jiang C and Ward M 1992 Proc. CVPR 606–12
[4] Salvador E op. Cit. 46-54
[5] Salvador E, Cavallaro and A Ebrahimi T 2004 Computer Vision and Image Understanding 95(2) 238-59
[6] Ramamoorthi R, Koudelka M and Belhumeur P 2005 IEEE Transactions on Pattern Analysis and Machine Intelligence 27(2) 288-95
[7] Oppenheim A and Lim J 1981 Proc. of the IEEE The Importance of Phase in Signals 69(5) 529-41
[8] Horner J and Gianino P 1984 Applied Optics 23(6) 812- 6
[9] Karim M and Awwal A 1992 Introduction to Optical Computing (WileyBlackwell) 46-56