Assignment 4: Photometric Stereo

Cristian Mitroi, Kaela DeAngelis, Terne Thorn Jakobsen

January 17, 2018

1 Introduction

The objective of photometric stereo is to construct a three-dimensional shape from a set of images of the same scene under different illumination settings. In the previous assignment, we learned that a disparity map can provide an estimation of depth in an image, which is useful information for a 3D reconstruction of in image but, as explained by Professor Robert J. Woodham [8], this traditional stereo correspondence technique is rather difficult. Woodham describes photometric stereo as the technique of using light radiance values of each image location in corresponding images, rather than positions of image features, to determine surface orientation. When using the photometric technique, determining correspondence is no longer an issue, since the viewing direction is constant.[8]

In this assignment, we perform photometric stereo by analyzing the information given by the reflection of light. With two sets of images of the same scene but with different shading, we show that this information is adequate for computing a depth map measuring the shape of the surface, and thereby provide 3D reconstructions.

This report is structured in similar fashion to the previous one, with the following sections:

1. introduction
2. overview of methods and theory
3. results
4. discussion and concluding remarks

2 Methods and Theory

Basic photometric stereo was applied to two datasets, "Beethoven" and "Buddha". The Beethoven dataset contained three, synthetically rendered images while the Buddha dataset contained ten naturally captured images. Natural images are accompanied by levels of uncertainty in directional distribution and orientation in light and in surface reflection [4]. Therefore synthetic images are useful for assessing image processing algorithms prior to using natural images because they control for uncertainties in photometry.

2.1 Lambert's Law

One important aspect pertaining to computing the depth map in photometric stereo is Lambert's Law (full name "Lambert's cosine law" [7]). It states that light is reflected by a surface equally in every direction. More precisely: "the radiant intensity or luminous intensity observed from an ideal diffusely reflecting surface or ideal diffuse radiator is directly proportional to the cosine of the angle θ between the direction of the incident light and the surface normal" [7]. An essential effect of this is that a surface will have the same "apparent brightness" [6] no matter the direction from which it is perceived. A material is referred to as "Lambertian" if it follows this law.

2.2 Albedo

The concept of albedo describes a surface that has the property of diffuse reflection; specifically, the albedo parameter is the quantity of light waves that uniformly bounce off a convoluted surface of an object, independent of the direction of the incident light [2].

The albedo, or surface reflectance, is used in implementing photometric stereo because image intensity values from illumination can be extracted to help solve for the height of particular surface points.

2.3 Surface normal and depth

An imaged object's surface, such as the objects of Beethoven and Buddha within the images, exhibits the feature of *surface normal* or *orientation* of a scene element at a pixel [5]. The concept of orientation encompasses a change, whether abrupt or continuous, from one pixel to the next, indicating curvature or smoothness. Pixels of a surface appear darker as rays of illumination become perpendicular to the normal.

Therefore surface normals can be computed from a set of image intensities and light vectors and help stimulate a 3D interpretation of a surface.

To compute the surface normals, one must obtain the albedo modulated normal field M and then extract the normal field by normalizing M. The procedure hereof is briefly explained here in the case of our Beethoven set (using inverse instead of pseudo-inverse):

For each pixel $I(p)$ of image I then:

$$I(p) = S \cdot (n(p)\rho(p))$$

Where S denotes light source, $n(p)$ denotes the normal of the pixel and $\rho(p)$ denotes the albedo. Furthermore we know that:

$$n(p)\rho(p) = S^{-1}I(p) = m(p)$$

and we can solve for the surface normal of the pixel p as:

$$n(p) = \frac{m(p)}{||m(p)||}$$

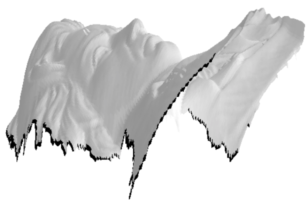
3 Results

Below are the results from implementing a photometric stereo algorithm on datasets Beethoven (section 3.1) and Buddha (section 3.2)

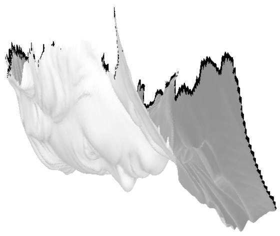
3.1 Beethoven

Figure 1: 2D image display of the albedo within the mask

(a)



(b)



(c)

Figure 2: 3D image display of Beethoven

The "Beethoven" dataset provided clear and satisfying results. This is in part due to the dataset containing synthetic images. As previously mentioned, synthetic data is useful for determining the efficacy of an image processing algorithm [4]. In using the Beethoven dataset, we can make assumptions on the application and its outputs. We can also make use of the results by comparing them to real image datasets and adjust our expectations.

We can see that all the edges of the profile are quite sharp. This is more obvious in image 2b. This implies the detection of lot of fine details in the final 3D visualization, a sign of a good dataset and and of a good implementation of the algorithm.

3.2 Buddha

Figure 3: 2D image display of the albedo within the mask

(a) (b) (c)

Figure 4: 3D image display of Buddha

It is visible in image 4a that some characteristics of the Buddha has been captured (e.g. wide nose and eyes in a somewhat chubbier face, and the forehead showing the beginning of his braided hairstyle). However, when viewing from different angles as in 4b and 4c, Buddha becomes unrecognizable, and we conclude that the algorithm could not construct any discernible heights. This is due to the dataset being a naturally-captured and thus "messy".

4 Discussion

In conclusion, we think our assignment is a success. We have implemented a simple photometric stereo algorithm to reconstruct the depth maps of an image from different sets of image intensities (I), light vectors (S), and masks. The computation was made possible by the fact that each set contained multiple images of the same object (I) under different light settings (S). We then displayed these at different angles to examine the results.

We can definitely notice a difference in quality and precision between the two datasets ("Beethoven" and "Buddha"). As mentioned earlier, we believe this is due to "Beethoven" being a synthetic (rendered on a computer) set. Also, in general, problems might arise from faults in measurement or "substantial regions of the surface [being] in shadow for one or the other light" [2].

Shadowing is unavoidable despite potential caution in arranging the light source; when not appropriately accounted for, it can cause distortions in 3D depth maps [1]. A traditional approach is to include additional images with greater variation in position of light sources to help obtain better surface point coverage. However, the Buddha dataset contained ten images, which is seven more than the recommended baseline. Another comparable approach would then be to use multiple light sources per image [1]. Multiple light sources can improve signal strength, enhancing signal to noise ratio, while also creating a better condition for the light source matrix. The latter positively effects estimates of surface normal.

Another possible consequence of using real images is the effect of ambient light on the resulting depth map. Ambient light can cause unintended illumination and shadowing of surfaces [3]. Additionally, it is difficult to control for ambient light because it may only appear when artificial lights are in use. Otherwise, a suggested approach is to take an image without using artificial light and compute the difference between that output and those illuminated by artificial light [3]. Unfortunately, we are not aware of how the shadows are cast and are therefore unable to adjust to the appropriate attributes.

References

[1] M. Chandraker, S. Agarwal, and D. Kriegman. Shadowcuts: Photometric stereo with shadows. pages 1–8, June 2007.

[2] D. A. Forsyth and J. Ponce. *Computer Vision*. Pearson Education Limited, 2011.

[3] F. Logothetis, R. Mecca, Y. Quéau, and R. Cipolla. Code for: Near-field photometric stereo in ambient light. 09 2016.

[4] H. Rushmeier, G. Ward, C. Piatko, P. Sanders, and B. Rust. *Comparing Real and Synthetic Images: Some Ideas About Metrics*, pages 82–91. Springer Vienna, Vienna, 1995.

[5] L. Shapiro and G. Stockman. *Computer Vision*. Prentice Hall, 2001.

[6] WikiLectures Contributors. Lambert's law - wikilectures. http://www.wikilectures.eu/w/Lambert%27s_law, 2018.

[7] Wikipedia contributors. Lambert's cosine law — wikipedia, the free encyclopedia. https://en.wikipedia.org/w/index.php?title=Lambert%27s_cosine_law&oldid=811111044, 2017. [Online; accessed 16-January-2018].

[8] R. J. Woodham. Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 19(1), 1980.