Computer Vision Lab 1

Mark Alence 13771272, Juno Prent 11915307, Mathieu Coenegracht 11323205

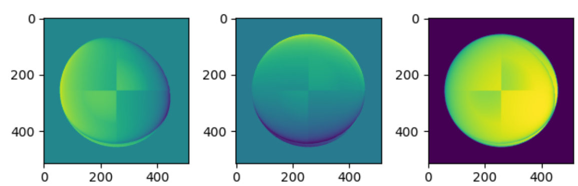
September 19, 2021

1 Photometric Stereo

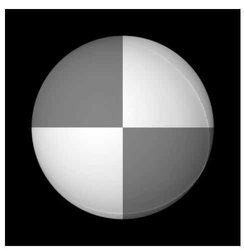
1.1

1) Since the image is of a sphere, we expect the albedo image to resemble the centre most image of the sphere. In this case, there is an image of the sphere with a camera direction of [0,0] (centred), which very closely resembles the albedo image, as expected.

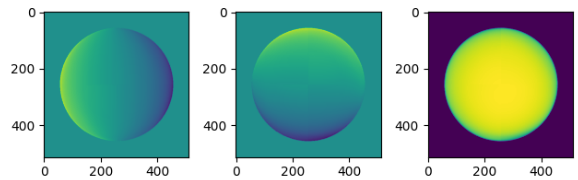
2) In principle, a minimum of three images is needed to estimate albedo and the surface normal map.



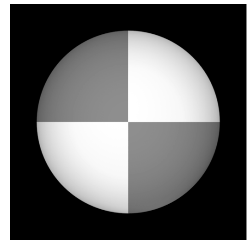
(a) The 3 channels of norms of a sphere. (X left, Y middle Z, right)



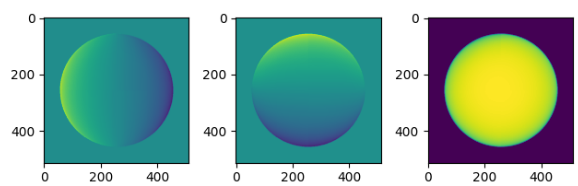
(b) Albedo of sphere with 3 images



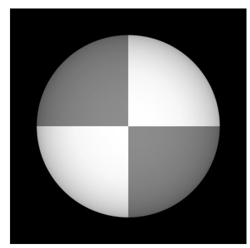
(a) The 3 channels of norms of a sphere. (X left, Y middle Z, right)



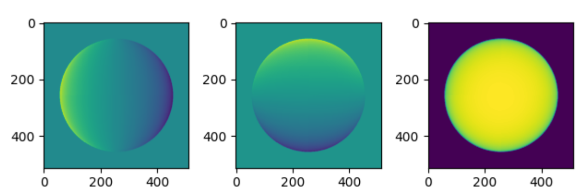
(b) Albedo of sphere with 6 images



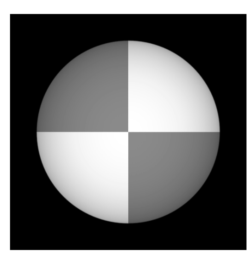
(a) The 3 channels of norms of a sphere. (X left, Y middle Z, right)



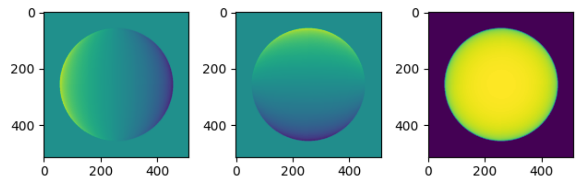
(b) Albedo of sphere with 9 images



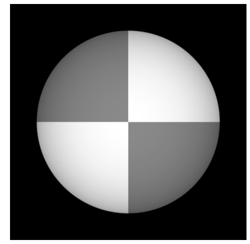
(a) The 3 channels of norms of a sphere. (X left, Y middle Z, right)



(b) Albedo of sphere with 15 images

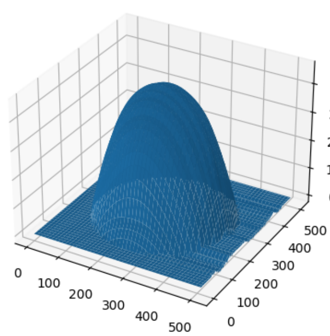


(a) The 3 channels of norms of a sphere. (X left, Y middle Z, right)

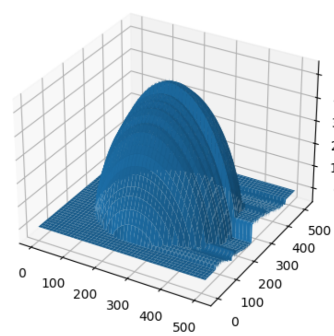


(b) Albedo of sphere with 25 images

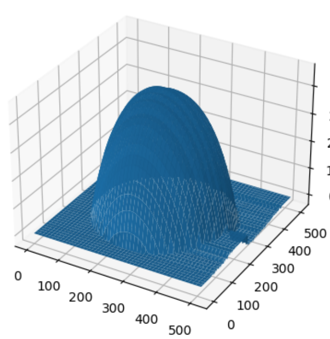
3) The problem with shadows is that it leads to ambiguity when trying to solve the linear system of equations for each point in the image. To remove this ambiguity, we zero out the contributions of shadowed regions. Without zeroing out the contributions of shadowed regions, the linear system to acquire the albedo, $g(x, y)$ for point (x, y) in the image is given by:



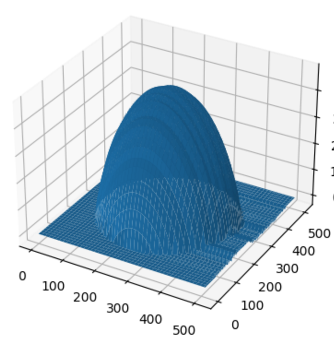
(a) Surface of a sphere using 5 images with no shadow trick



(b) Surface of a sphere using 5 images with shadow trick



(a) Surface of a sphere using 5 images with no shadow trick



(b) Surface of a sphere using 25 images with shadow trick

$$\boldsymbol{i} = \boldsymbol{V}\boldsymbol{g}(x, y) \tag{1}$$

where $\boldsymbol{i}$ is the image vector and $\boldsymbol{V}$ is the vector of light sources. To ensure that shadowed regions do not contribute to the linear system, we multiply both sides of the equation by $diag(i)$.

1.2

Errors could arise due to the light shining close to directly back into the camera, resulting in spikes in the normals. In general, the trend was that the more images we used, the less outliers we had.

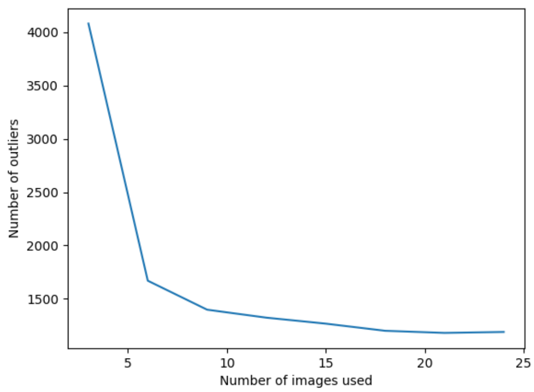
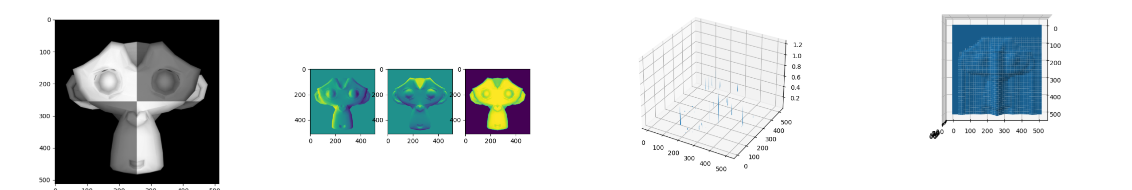
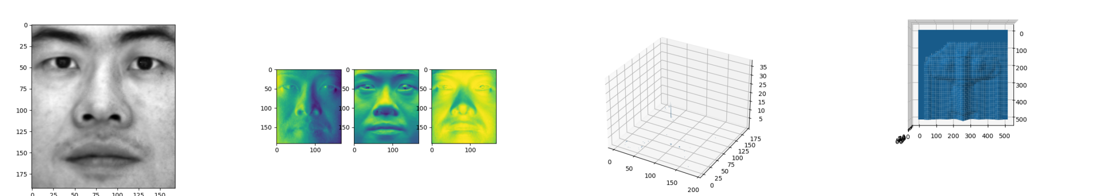


Figure 8: Graph showing number of images used vs outliers, given a threshold of 0.005

1.3



(a) Monkey gray albedo (b) Monkey gray normals (c) Monkey gray squared errors (d) Monkey gray surface (elevation=90 degrees, azimuth=0 degrees)



(a) Face albedo (b) Face normals (c) Facesquared errors (d) Face surface (elevation=90 degrees, azimuth=10 degrees)

2 Colour spaces

RGB Colour model

Colour space conversion

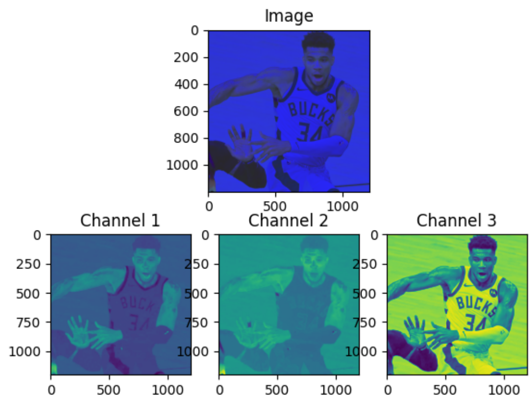


Figure 11: Image transformed to the opponent colour space, including the separate RGB channels

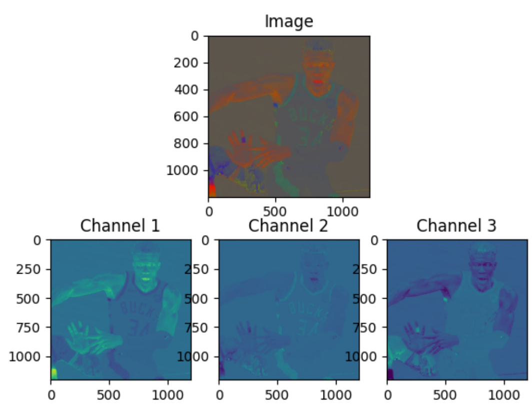


Figure 12: Normalized RGB image, including the separate rgb channels

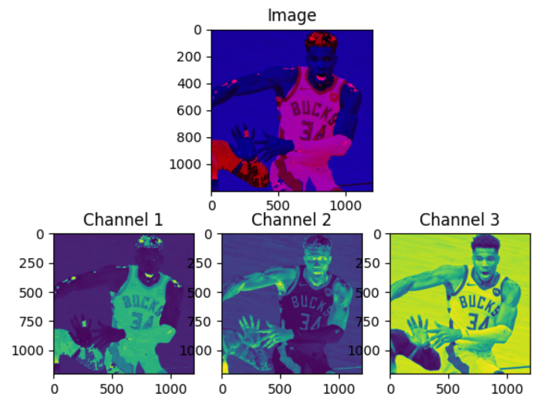


Figure 13: Image transformed to hsv colour space, and the separate RGB channels

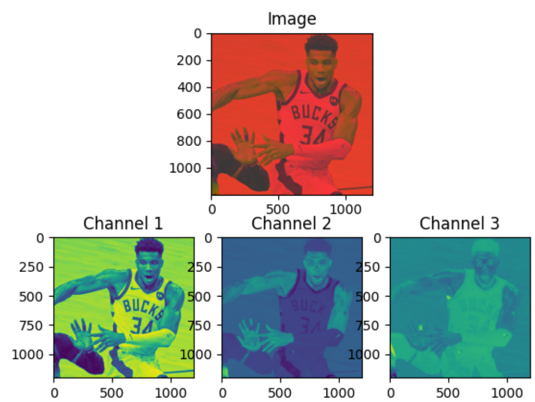


Figure 14: Image transformed to the YCbCr colour space, and the separate RGB channels

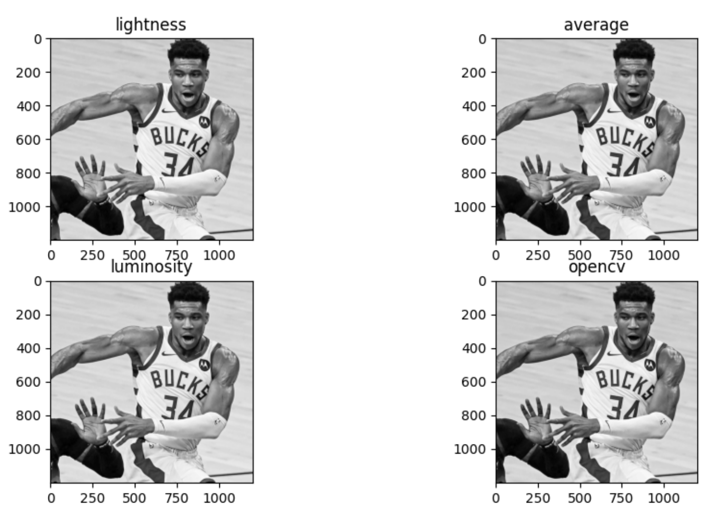


Figure 15: The original image and recolored image

Colour Space Properties

Opponent

RGB normalized

hsv

YCbCr

Grayscale

2.0.1 More on Colour Spaces

Another colour space is the LMS colour space, with LMS standing for long, medium and short. This is named after the sensitivity of the three types of cones in the human eye and can for instance be used to detect colour blindness caused by a defective cone.

3 Intrinsic Image Decomposition

Other Intrinsic Components

Structure and texture. In structure-texture image decomposition the image is decomposed into a structural part, which represents the main objects in the image, and a textural part, which contains the finer details. [1]

Synthetic Images

Synthetic image datasets can provide a large number of annotated images with a ground truth decomposition to check your algorithms against. Going through the process of manually decomposing and annotating a dataset of real images would be a very tedious and time consuming task. In addition, they lack a ground truth decomposition, as the images would have to be decomposed with an intrinsic decomposition algorithm first.

Image Formation

The **iid_image_formation.py** script was used to reconstruct the *ball.png* image from its intrinsics *ball_albedo.png* and *ball_shading.png*. The result is shown in figure 16.

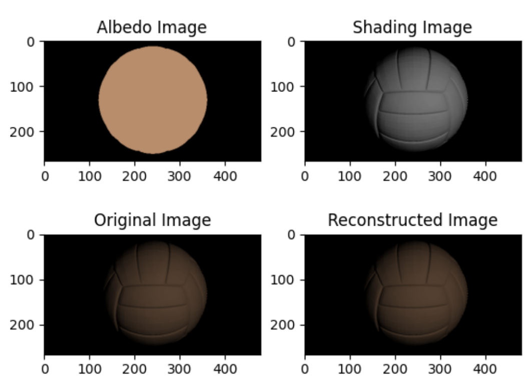


Figure 16: Ball image reconstructed from its albedo and shading intrinsics

3.1 Recoloring

3.1.1 True material color of ball

The true material color of the ball in RGB space is [184 141 108].

3.1.2 Recoloring script

The **recoloring.py** script recolors the *ball_albedo.png*. It assigns all the pixels that do not have an RGB value of [0, 0, 0] the value of [0, 255, 0]. The recolored image is then obtained by element-wise multiplication of the albedo image with the shading image. The original image and the recolored image are shown in figure 17.

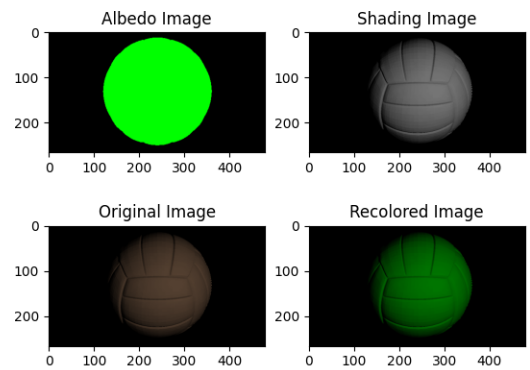


Figure 17: The original image and recolored image

3.1.3 Colour distribution

This is because of the shading intrinsic. It alters the darkness of the pixels based on the angle to the light source. This gives the image the perception of depth. Without it the image would look like it's albedo component.

4 Colour Constancy

AWB.py is a script that applies the grey-world algorithm to a given image. This is a color constancy algorithm that ensures that the average of each of the three color channels is exactly in the middle of their possible ranges of values by linearly stretching the RGB values of the image's pixels. The result of this is an image that looks to be taken under a white light source. The effect of this algorithm on an image is displayed in figure 18.

4.1 Grey-world algorithm

Figure 18: The effect of the grey-world algorithm on an image

4.2 Drawback of the grey-world algorithm

One of the downsides of the grey-world algorithm is that it fails when a scene's surface reflectances are inconsistent. As such, it cannot be used for any scene in which objects appear or disappear over time, only on static scenes [2].

4.3 Other algorithm

An alternative to the grey-world algorithm is the max-RGB algorithm. In this algorithm, the pixel values are linearly stretched in such a way, that the pixel with the most intensity gains the RGB values (255, 255, 255).

References

[1] Jean-François Aujol, Guy Gilboa, Tony Chan, and Stanley Osher. Structure-texture image decomposition—modeling, algorithms, and parameter selection. *International journal of computer vision*, 67(1):111–136, 2006.

[2] José Buenaposada and Luis Baumela. Variations of grey world for face tracking. 7, 06 2001.