An investigation into directional statistics and the illumination estimation calculations necessary to the performance of plausible three-dimensional manipulations of an object in a photograph

Table of Contents

Introduction

The objective of this investigation is to explore directional statistics with the ultimate goal of analyzing how one can calculate the illumination estimation necessary to the performance of plausible three-dimensional manipulations of an object in a photograph. In order to accomplish this, I will first be examining the different elements of directional statistics. Then, I will be analyzing circular distribution and providing a two-dimensional example. Finally, I will be exploring von Mises-Fisher distribution and presenting a three-dimensional example.

The topic is of significance because one of the factors necessary to the creation of a conceivable photograph manipulation is the correction of illumination discrepancies; to complete this task, one must first estimate the illumination of the object. I find it extremely interesting that if any images, including those with historical significance, can be edited with plausible results, photography itself is rendered as unreliable as the spoken or written word; the common idiom "a picture is worth a thousand words" thus becomes completely void.

According to an examination conducted by researchers at Clarkson University, Carnegie Mellon University, and the University of California, Berkeley, the photograph I is a function of the object geometry $\bar{X}$, the object appearance $\bar{T}$, and the environmental illumination $\bar{L}$ (Kholgade Banerjee). I will be concentrating on the environmental illumination $\bar{L}$. I will be examining directional statics as a whole and the use of von Mises-Fisher distributions, the most common type of circular distribution, in the calculation of the estimation necessary to the performance of adjustments of an object in a photograph.

Directional Statistics

Directional statistics, or circular statistics, is a subset of statistics that analyzes unit vectors in $\mathbb{R}^n$. Essentially, $\mathbb{R}^n$ represents real coordinate space of *n* dimensions. $\mathbb{R}^n$ is literally defined as "a set of all n-tuples of real numbers" (Rootmath). A tuple is simply an ordered list of elements. For example, the vector

$$\vec{v} = [v_1, v_2]$$

is a 2-tuple, as it contains two components. Therefore,

$$\vec{v} \in \mathbb{R}^2 .$$

The vector

$$\vec{u} = [u_1, u_2, u_3]$$

is a 3-tuple, as it contains three components. Therefore,

$$\vec{u} \in \mathbb{R}^3 .$$

The vector

$$\vec{w} = [w_1, w_2, w_3, ...w_n]$$

is a *n*-tuple, as it contains *n* components. Therefore,

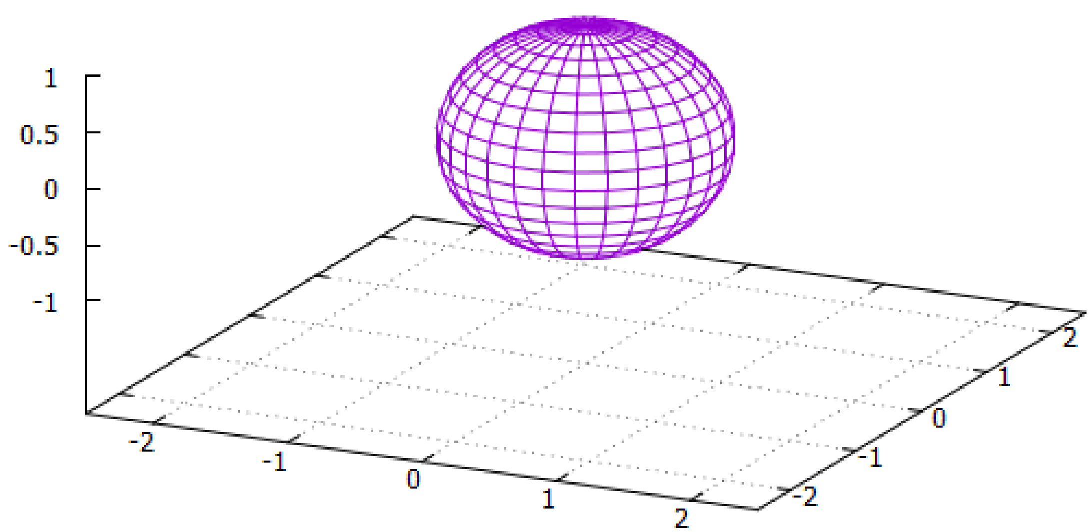
$$\vec{w} \in \mathbb{R}^n$$

(Rootmath). Thus, directional statistics essentially studies a group of unit vectors that contain a *n* number of components and that have a defined direction. For the purposes of this investigation, both two- dimensional vectors such as $\vec{v}$ and three-dimensional vectors such as $\vec{u}$ will be analyzed, as the objective is both to provide an understanding of circular distribution and to estimate the illumination of an object inside a two-dimensional photograph.

Directional statistics also analyzes manifolds. A manifold is simply a collection of points that form a set. For example, one-dimensional manifold is a curve; a two-dimensional manifold is a surface. It is at this point that it is necessary to define "sphere". A *n*-sphere (where *n* represents any natural number) is defined as a series of points in (n+1)-dimensional Euclidean space that are located a distance *r* (radius), which can be any positive real number, from the center. A *n*-sphere is denoted as S^n. Thus, a *n*-sphere may be generally defined as

$$S^n = \left\{ x \in \mathbb{R}^{n+1} : \|x\| = r \right\}.$$

For, a 0-sphere (denoted S^0) is a pair of points at the ends of a one-dimensional line segment. This investigation will primarily analyze both 1-spheres and 2-spheres. A 1-sphere (denoted S^1) is a circle. A 2-sphere (denoted S^2) is "the set of all points in three-dimensional Euclidean space $\mathbb{R}^3$" (Weisstein, "Sphere") that are located a distance *r* (radius) from the center. Essentially, this definition states that a sphere consists of a two-dimensional manifold (the surface) embedded within a three-dimensional space. The following image shows a 2-sphere, which I graphed using gnuplot, a command-line driven graphing program.



The following is the command line that I used.

```
set nokey
set parametric
set hidden3d
set view 60
set isosamples 30,20
set xrange[-2.5 : 2.5]
set yrange[-2.5 : 2.5]
set zrange[-1 : 1]
splot [-pi:pi][-pi/2:pi/2] cos(u)*cos(v), sin(u)*cos(v), sin(v)
```

Circular Distribution

Directional data can be analyzed and graphed in terms of circular distributions. Circular distribution, or polar distribution, is a probability distribution in which the values of a random variable are "defined in terms of angles, generally lying in the range [0, 2π)" ("Circular Distribution").

In order to analyze the data in question, it first must be "transformed into rectangular polar coordinates" (Marr). To do so, a unit circle, which has a radius of one, is utilized. Thus, the polar location is defined as "the angular measurement and its intersection with the unit circle" (Marr). This location is then placed "into a standardized Cartesian space" by "cosine and sine functions" (Marr). The cosine of angle *a* is defined as

$$\cos a = \frac{x}{r} = \frac{x}{1} = x \, ,$$

(Marr) while the sine of angle *a* is defined as

$$\sin a = \frac{y}{r} = \frac{y}{1} = y \, .$$

(Marr). For example, one may consider the following data:

Angle *a*	$\cos a$ (azimuth)	$\sin a$ (azimuth)
353	0.99255	-0.1219
165	-0.9659	0.25882
118	-0.4695	0.88295
73	0.29237	0.9563
13	0.97437	0.22495
275	0.08716	-0.9962
126	-0.5878	0.80902
247	-0.3907	-0.9205
302	0.52992	-0.848
42	0.74314	0.66913

In order to calculate the above angle measures, I marked a spot on my desk and assigned it the coordinate (0,0); next, I drew the axes through this origin, aligning them with the cardinal directions. I then created vectors from the origin to the central points of ten objects on my desk and determined angle measurements between these vectors and the positive *y*-axis. The mean angle can be calculated with the equation

$$\theta_r = \arctan(\frac{\sin \bar{a}}{\cos \bar{a}}),$$

where $\sin \bar{a}$ is defined as

$$\bar{y} = \frac{\sum_{i=1}^{n} \sin_a}{n}$$

and $\cos \bar{a}$ is defined as

$$\bar{x} = \frac{\sum_{i=1}^{n} \cos_a}{n} \;;$$

The bar above the variables *a*, *x*, and *y* simply represents the mean.

Thus, $\sin \bar{a}$ is equivalent to

$$\bar{y} = \frac{\sum_{i=1}^{10} \sin_a}{10} = 0.091455$$

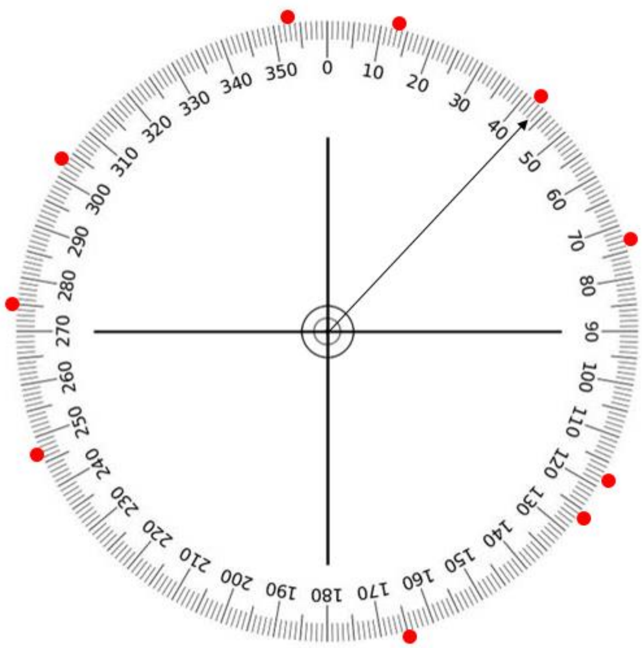
and $\cos \bar{a}$ is equivalent to

$$\bar{x} = \frac{\sum_{i=1}^{10} \cos_a}{10} = 0.120559 \,.$$

Thus,

$$\theta_r = \arctan(\frac{0.091455}{0.120559}) = 0.758591(\frac{180}{\pi}) = 43.464076$$

The following is a graphical representation of this data.



If more data was to be calculated, it would be more likely to be close to the mean.

Circular distributions have a standard deviation that is defined as

$$R = \sqrt{\bar{x}^2 + \bar{y}^2}$$

(Marr), where R is essentially the equivalent to standard deviation except for the fact that it ranges from 0, which denotes "uniform dispersion" (Marr), and 1, which denotes "complete concentration" (Marr) in a singular direction. Thus,

$$R = \sqrt{(0.091455)^2 + (0.120559)^2} = \sqrt{0.008364 + 0.014534} = \sqrt{0.022898} = 0.151322$$

A central idea of this investigation is to apply the concept of two-dimensional circular distributions to the third dimension. To accomplish this, it is necessary to examine von Mises-Fisher distributions.

Von Mises-Fisher Distributions

The von-Mises Fisher distribution is a probability distribution on a (n-1)-dimensional sphere in $\mathbb{R}^n$. It is often known as the circular normal distribution.

One can perform realistic photographic manipulations quite effortlessly with programs such as Photoshop, as they allow one to execute a variety of image operations. However, despite the fact that photographs depict three dimensions, these programs are restricted, particularly in terms of geometric correction, to two-dimensional manipulations. In other words, "the photograph 'knows' only the pixels" (Kholgade Banerjee) of the two-dimensional projection of the object, not its three-dimensional structure. In order to perform three-dimensional manipulation of objects in a photograph, one must edit "'what we know' about the scene beside the photograph" (Kholgade Banerjee) as opposed to "simply editing 'what we see'" (Kholgade Banerjee). One of the elements necessary to the performance of realistic manipulations in three dimensions is the generation of accurate lighting effects. It is necessary to calculate the original illumination of the object, as one must then project it onto the corrected object. Assuming “light sources are distant…, illumination can be approximated as a mixture of light distributions on a unit sphere of light directions” (Panagopoulos). Essentially, this means that the illumination of an object can be represented as a set of vectors of *N* random three-dimensional unit vectors that “produce *N* samples of the illumination environment” (Panagopoulos). The von Mises-Fisher distribution can be used as a method to calculate these distributions.

The von Mises- Fisher distribution is defined as

$$f_p(x,\mu,\kappa) = C_p(\kappa)e^{(\kappa\mu^T x)},$$

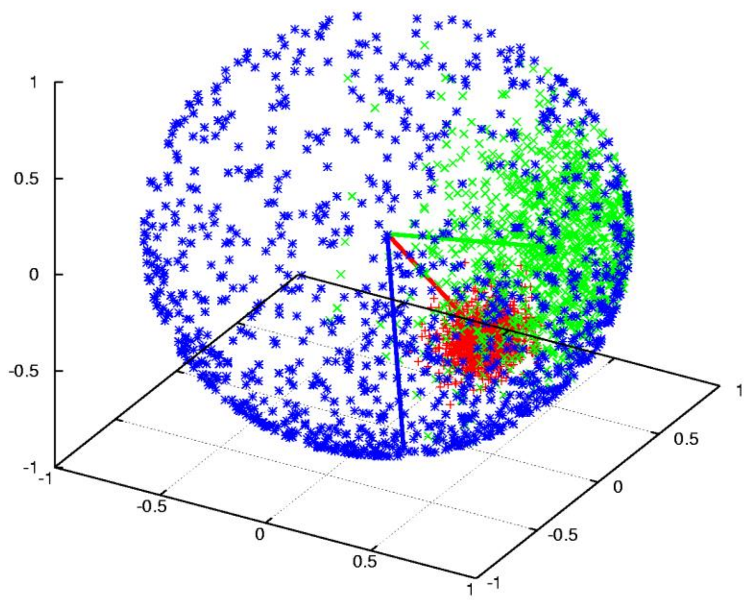
(Dhillon) where *x* represents a vector. $C_p(\kappa)$ is the normalization constant, which is defined as a constant by which a function must be multiplied so that the area under the graph is equivalent to one. Essentially, this allows for the function in question to become a probability density function. The normalization constant is defined as

$$C_p(\kappa) = \frac{\kappa^{\frac{p}{2}-1}}{(2\pi)^{\frac{p}{2}} I_{\frac{p}{2}-1}(\kappa)};$$

(Dhillon) *I* denotes a modified Bessel function of the first kind, which behave like cosine graph. Bessel functions form the basis for series expansion (Weisstein, “Modified Bessel Function of the First Kind”). However, it is not necessary to examine Bessel functions for the purposes of this investigation, as when $p=3$, $C_3(\kappa)$ is defined as

$$C_3(\kappa) = \frac{\kappa}{4\pi \sinh\kappa} = \frac{\kappa}{2\pi(e^{\kappa} - e^{-\kappa})},$$

(Kholgade Banerjee) where $\sinh$ is the hyperbolic sine function, which can also be expressed with the constant *e*. κ is the concentration parameter. Essentially, the greater the value of κ, the higher the concentration of the distribution around the mean direction μ. The following is a graphic example of three von Mises-Fisher distributions.



("Pattern Recognition")

$\kappa = 1$ for the data in blue, $\kappa = 10$ for data in green, and $\kappa = 100$ for the data in red. The mean direction μ is represented by the blue, green, and red arrows respectively. Thus, one can observe the aforementioned description of κ. When $\kappa = 0$, the distribution is uniform on the sphere. For $\kappa > 0$, the distribution is unimodal, which means that the data has a singular, clear peak. κ is defined as

$$\hat{\kappa} = \frac{\bar{R}(p - \bar{R}^2)}{1 - \bar{R}^2}.$$

(Dhillon) $\hat{\kappa}$ denotes an estimator of κ ; $\bar{R}$ is the mean resultant length, which is defined as

$$\bar{R} = \frac{r}{n},$$

in which *r* represents the resultant length and *n* represents the sample size.

The following is a side view of *Cloud Gate* (or as it is more commonly known, The Bean), which is a public outdoor work in Chicago, Illinois that was created by British artist Anish Kapoor ("Art and Architecture").

(Hess)

The following is a front view.

("Art and Architecture")

I chose to examine The Bean due to its stainless steel surface, as I knew that it would be easier to estimate illumination from a highly reflective surface. I then asked myself what the result would be if I estimated the illumination of The Bean; I thought it would be artistically interesting to examine the outcome of a rotation or a flip of The Bean, as this would be a manipulation of the artist's original objective (which was the theme of distorted perception).

Using the first image, I estimated three values for the concentration parameter κ: $\kappa = 4$ for the blue data, $\kappa = 40$ for the red data, and $\kappa = 400$ for the green data. If one looks at the image, one will see that the data is extremely concentrated around the mean, but then progressively spreads out. I therefore used three different parameters for κ in order to account for all of the data. I then found $C_3(\kappa)$. For example,

$$C_3(4) = \frac{4}{4\pi \sinh 4} = \frac{4}{2\pi(e^4 - e^{-4})} = 0.011664$$

I calculated the mean direction μ of the data to be

$$\mu = [0.419643, 0.649642, 0.633929]$$

I then generated random vectors that corresponded to both κ and μ. The following are a few examples.

$$\vec{v} = [0.420173, 0.652837, 0.630284]$$

$$\vec{u} = [0.486542, 0.695302, 0.528992]$$

$$\vec{w} = [0.475483, 0.635281, 0.608548]$$

$$\vec{a} = [0.434578, 0.678532, 0.592230]$$

$$\vec{b} = [0.453876, 0.610721, 0.648858]$$

$$\vec{c} = [0.486740, 0.672452, 0.557577]$$

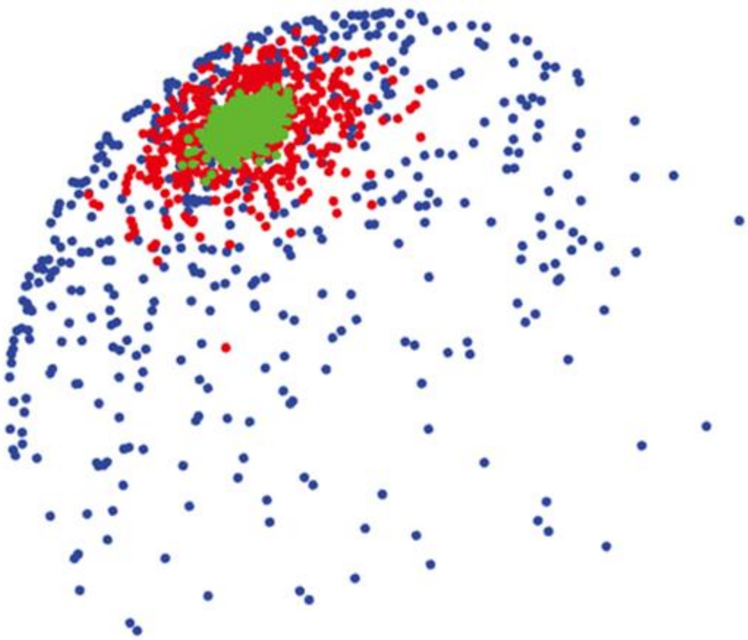
$$\vec{d} = [0.429763, 0.645987, 0.630876]$$

$$\vec{e} = [0.489532, 0.605764, 0.627223]$$

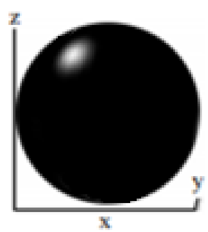
$$\vec{f} = [0.422587, 0.652824, 0.628682]$$

$$\vec{g} = [0.435847, 0.627943, 0.644767]$$

Using this data, I then created the following graphical representation, which is modeled after a graph in the article titled “Pattern Recognition”.



Thus, the following is the estimated illumination of The Bean in that specific photograph.



I modeled the above representation after a graph in the investigation titled "3D Object Manipulation in a Single Photograph Using Stock 3D Models." This is the illumination that will need to be projected onto the adjusted object.

Conclusion

As stated in the introduction, the aim of this investigation was to explore directional statistics with the ultimate goal of analyzing how one can calculate the illumination estimation necessary to the performance of plausible three-dimensional manipulations of an object in a photograph.

I began this exploration with a vague notion that I wanted to research two-dimensional and three-dimensional vectors, specifically in photography. After some research, I discovered the investigation conducted by researchers at Clarkson University, Carnegie Mellon University, and the University of California, Berkeley (the title of which is “3D Object Manipulation in a Single Photograph Using Stock 3D Models”). I was immediately interested in the concept. The idea, which I mentioned in the introduction, is that one can manipulate objects in photographs in the third dimension; this is as fascinating as it is revolutionary. It is essentially a “new and improved” version of programs such as Photoshop.

The implications of this concept, which I also briefly commented upon in the introduction, are quite monumental. As I stated, the phrase “a picture is worth a thousand words” is now essentially void. Photography has followed in the footsteps of other means of communication in that it is now even more susceptible to human interference. The advancement of technology and mathematics has allowed for the successful completion of photographic manipulations, such as the rotation of a chair in an office (as is shown in the picture below) or of a house in a suburb.

(Kholgade)

More seriously, it has facilitated the potential interference of various parties in certain critical situations, such as in journalistic or police investigations or in the creation of an unbiased collection of historical events. An example of such is the following image, as a photographic manipulation of the Grumman TBF Avengers would have changed the course of that particular World War II battle.

(Kholgade)

Naturally, this investigation has its limitations, the most substantial of which is the von Mises-Fisher distribution itself. Much of the material gathered, specifically the equations, is from a multitude of sources; this is certainly beneficial in determining the accuracy of the information. My intention throughout the investigation was to maintain the integrity of the information found in the sources. Nevertheless, the research done in this area is extremely advanced; my understanding and explanation of the concepts may thus be at best oversimplified and at worst inaccurate. Another major limitation was the difficulty of the acquisition of the programs necessary to the completion of investigation. For example, it was challenging for me to find and utilize the graphing program gnuplot. Because I have no real experience with commands, even graphing a sphere proved to be problematic. As in any investigation, there is also always a possibility of human error in the gathering of data.

In general, this project has been a chance for me to explore and a mathematical topic that I never would have otherwise; although it was quite challenging, it has also been enjoyable.

Bibliography

"Art & Architecture." *Millennium Park*. City of Chicago, 2016. Web. 22 Feb. 2016.

"Circular Distribution." *Math: Statistics*. NCS Pearson, 2016. Web. 22 Feb. 2016.

Dhillon, Inderjit, and Suvrit Sra. "Modeling Data using Directional Distributions". *Department of Computer Sciences*. University of Texas, Austin, 2003. Web. 22 Feb. 2016.

Hess, George. "June/July 2006." *Photo of the Month*. George Hess, June-July 2006. Web. 22 Feb. 2016.

Kholgade Banerjee, Natasha, Tomas Simon, Alexei Efros, and Yaser Sheikh. "3D Object Manipulation in a Single Photograph Using Stock 3D Models." *OM3D*. Carnegie Mellon University, 2014. Web. 22 Feb. 2016.

Marr, Paul. "Directional (Circular) Statistics." *Paul Marr: Professor of Geography*. Shippensburg University, 2012. Web. 22 Feb. 2016.

Panagopoulos, Alexandros, Dimitris Samaras, and Nikos Paragios. "Robust Shadow and Illumination Estimation Using a Mixture Model." *Imaging and Bioinformatics Lab*. Berkeley Lab, 2009. Web. 22 Feb. 2016.

"Pattern Recognition." *LausdeoF*. Tistory, 17 Sept. 2014. Web. 22 Feb. 2016.

Rootmath. "1.1 Definition of R^n." *YouTube*. Google, 7 Aug. 2011. Web. 22 Feb. 2016.

Weisstein, Eric W. "Modified Bessel Function of the First Kind." *Wolfram MathWorld*. Wolfram Web Resources, n.d. Web. 22 Feb. 2016.

Weisstein, Eric W. "Sphere." *Wolfram MathWorld*. Wolfram Web Resources, n.d. Web. 22 Feb. 2016.

Assessment criteria:

Criterion		Comments *	Achievement level		
			Teacher	Moderator	Senior moderator
A	Communication	The paper was well written, organized in an easy to follow manner, and incorporated a variety of visuals to aid the reader. All information was concise and relevant to the topic, leading the reader along well.	**0-4** 4		
B	Mathematical presentation	The variety of presentation components was excellent. Images were included at appropriate times so the reader did not have to backtrack. The 3D images were especially useful in following the topic.	**0-3** 3		
C	Personal engagement	The topic selection was extremely unique and creative, but also added a high degree of complexity. Ideas were presented in a unique, personal way that showed individual investment in the work.	**0-4** 4		
D	Reflection	The initial reflection of images not being reliable anymore flowed right into the end analysis that discussed links to art, wartime imagery, and accuracy of images in general. Limitation were noted and explored.	**0-3** 3		
E	Use of mathematics	The level of mathematics ranged from solid to complicated. The use of 2D,3D, and higher spaces were an excellent extension of core topics. The calculations themselves were solid, but not rigourous.	**0-6** 4		
		Total:	**0-20** 18		