MIT Thesis Template in Overleaf

by

Tim Beaver

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Bachelor of Science in Computer Science and Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 1990

Author . . . .
Department of Electrical Engineering and Computer Science
May 18, 1990

Certified by . . . .
William J. Supervisor
Associate Professor
Thesis Supervisor

Accepted by . . . .
Arthur C. Chairman
Chairman, Department Committee on Graduate Theses

MIT Thesis Template in Overleaf

by

Tim Beaver

Submitted to the Department of Electrical Engineering and Computer Science
on May 18, 1990, in partial fulfillment of the
requirements for the degree of
Bachelor of Science in Computer Science and Engineering

Abstract

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nam quis neque et erat laoreet finibus at ac leo. Curabitur pellentesque, diam quis dignissim finibus, enim dui feugiat leo, nec porttitor sapien mi ac felis. Nam aliquam pretium nibh, quis dapibus dolor gravida sit amet. Cras porttitor dui quis elementum pulvinar. Nulla id pulvinar massa. Nullam ut diam non lorem venenatis faucibus. Vivamus lacus ante, pellentesque vitae nisl sit amet, bibendum facilisis purus.

Thesis Supervisor: William J. Supervisor
Title: Associate Professor

Acknowledgments

This is the acknowledgements section. You should replace this with your own acknowledgements.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In this thesis, I will provide a complete treatment of occluder-based imaging. The thesis is split into two main sections. The first section will be focus on the question of designing *optimal* occluders. The natural application of this kind of research is generally within context of designer mask-based cameras. The second section will be about methods for exploiting *accidental cameras*, which means using occlusion provided by objects that happen to be in the world to image hidden scenes.

Although there is obviously a close relationship between these two areas of research, and I will use a common framework for analyzing both they are nevertheless distinct, and which one the reader is interested in will depend on the application they are interested in.

I will begin my thesis by describing in detail the analytic framework that I use for both, and introducing the assumptions I use and the optical model that underlies them. My hope is that even without much or any background in optics, all my readers will have no trouble understanding the model and assumptions I use, and my mathematically-savvy readers will have no trouble understanding the conclusions that follow from them.

I highly recommend that *any* reader begins by reading the chapter that follows before proceeding to either of the two main sections (on optimal occluders and on accidental cameras). Without this background, I expect that it will be difficult to understand what I'm saying.

Chapter 2

The Model and Assumptions

2.1 Ray Optics and BRDFs

Throughout my thesis, unless I say otherwise, I'll be using the *ray optics model* of light. This means that, as a convenient abstraction, I'll be assuming that light moves in a straight line through the air, can be bent when it hits a different light-propagating medium (like a lens), and may be absorbed or reflected by the materials it hits (like a wall). Moreover, I'll be generally assuming that light intensity is additive, meaning that two rays of light hitting the same point will generate an intensity equal to the intensity that would be generated the sum of each individual ray. This corresponds to assuming that the light we care about is incoherent and as such won't interfere with itself—a reasonable assumption when the light in question is coming from the sun or from a commercial electric light. Finally, using the ray optics model means that I'll be ignoring the effects of diffraction, which is reasonable when modeling the behavior of visible-wavelength incoherent light hitting macroscopic structures. If this paragraph was hard to understand, don't worry—just take it to mean that light travels in straight lines, bounces off stuff it hits, and generally does what you intuitively think it should.

What happens when the light hits an opaque surface, according to this model? Some of it will be absorbed, and some reflected. How much of it is reflected, and in what directions, is described by the *bidirectional reflectance distribution function*, or

BRDF, of the surface. The BRDF is a function from *incident*, or incoming, angle of the light to the outgoing angle of the light. It's best explained with two example BRDFs, which happen to describe the many of the surfaces we care about.

The first example BRDF is called the *specular* BRDF, and surfaces that have this BRDF are called specular surfaces. The typical example of a specular surface is a mirror. The specular BRDF takes the incident light and flips it across the surface normal. How can we describe this mathematically? We can describe it using a function of two arguments, the first being the angle of the incident light, θ_{in} and the second being the angle of the outgoing light, θ_{out}. Each angle is given from the surface normal. [1]Here, the specular BRDF we want is:

$$f_{\text{specular}}(\theta_{\text{in}}, \theta_{\text{out}}) = \begin{cases} \rho & \text{if} \theta_{\text{in}} = \pi - \theta_{\text{out}} \\ 0 & \text{otherwise} \end{cases}$$

The ρ here is a constant that determines the overall brightness of the surface in question—how much of the light is actually reflected rather than absorbed or transmitted. For example, for a mirror, ρ might be almost 1, meaning that almost all the light is reflected, but for a window where most of the light is transmitted rather than reflected, ρ might be much smaller, like 0.01 or 0.001.

The second example BRDF is called the *Lambertian* BRDF, after the 18th-century physicist Johann Heinrich Lambert. Surfaces that have this BRDF are often called *Lambertian*, *matte*, or *diffuse* surfaces, and I will use these terms interchangeably in this thesis. Intuitively, this BRDF takes the incoming light and scatters it "equally in all directions." Formally, here is the BRDF in question:

$$f_{\text{Lambertian}}(\theta_{\text{in}}, \theta_{\text{out}}) = \rho \cos(\theta_{\text{out}})$$

[1]If you're already familiar with the concept of BRDFs, you might be confused by this, since BRDFs are often described as functions of four real values. This confusion comes from the fact that for a 2D surface that lives in three dimensions, the angle of a light-ray from that surface is given by a 2D angle, which requires two real numbers to describe (one for the azimuth angle and one for the zenith angle). This is a detail that becomes unimportant if you treat each of the two arguments I'm describing as 2D angles, with equality between 2D angles achieved when both their azimuth angle and their zenith angle match. Here, to keep things simple, I'm implicitly assuming a 1D surface that lives in a 2D world, so angles are described by a single real number.

Once again ρ is a constant that determines the overall surface brightess. Note that the Lambertian BRDF is completely independent of the angle of the incident light—it scatters the light it reflects in exactly the same way no matter where the light came from.[2]

Now at this point, a careful reader may object: why did I claim that Lambertian surfaces scatter light "equally in all directions," when in reality, they scatter light in directions proportionally to that direction's cosine? Indeed, this is a major source of confusion when it comes to Lambertian surfaces. Google "Lambertian surface" or "Lambertian BRDF" and you will find about half your results defining it as I do, and half of them defining it instead as a perfectly constant function, depending neither on θ_{in} nor θ_{out}. This is an important confusion to resolve; I hope now to convince you beyond a doubt that my definition is the right one, and that the alternate definition of a Lambertian surface, while the objects it describes might exist in principle, in practice I've never seen one—whereas the objects my definition describes are all over the place. The walls and ceiling of the room you're sitting in, the paper you may be reading these words on, the clothes you're reading: all of these are nearly perfectly described by the Lambertian BRDF as I have defined it.

Here's where the confusion comes from. Find a sheet of paper. Lay it flat against a desk, then look at it from a few different angles. No matter what angle you look at it from, it looks equally bright.

At first, this seems it argues for the alternate definition of a Lambertian surface. After all, if you see the same amount of light coming from the sheet of paper no matter what direction you observe it from, doesn't that it mean that the amount of light it transmits is equal in all directions—that is, it doesn't depend on θ_{out}? In fact, no. Consider: depending on what angle you are looking at the sheet of paper from, the paper will take up a larger or smaller part of your field of vision. Look at the paper head-on, and it takes up a relatively large part of your field of vision; look at the paper from a very glancing angle, however, and it takes up just a sliver. And yet,

[2] For a 2D surface living in a 3D world, the Lambertian surface BRDF is exactly the same, but the cosine is of the outgoing zenith angle, and the outgoing azimuth angle doesn't matter.

no matter what the angle you observe it from is, you can still see the entire sheet of paper.

What's the upshot of this? What this means is that when you are looking at the sheet of paper from a very glancing angle, you are actually getting less total light from the paper, since the amount of light per amount-of-your-field-of-vision (sometimes called a "steradian") remains fixed, but the amount of your field of vision filled by the paper has decreased. And in fact, it has decreased by a factor of $\cos(\theta)$, where θ is your angle from the paper's normal. This is where the factor of $\cos(\theta)$ in the definition of the Lambertian BRDF comes from. Indeed, if the Lambertian BRDF sent light truly equally in all directions, as you looked at the paper from an increasingly glancing angle, the paper would appear to get *brighter*, in order to keep the total amount of light you were receiving from the paper constant. Some objects behave the opposite way, like backlit LCD screens; if you tilt them away from you, their apparent brightness will usually decrease (depending on the screen), which means that their BRDF is attenuated *faster* than $\cos(\theta_{\text{out}})$. But no object I've ever seen has a BRDF that is attenuated *slower* than $\cos(\theta_{\text{out}})$.

Most real-world surfaces lie on a spectrum somewhere between Lambertian and specular. This isn't to say that most real-world BRDFs are a linear combination of the Lambertian and specular BRDFs; an example of an object that *does* behave that way is a dirty or smudged mirror. But most objects aren't like that; they may look "shiny" or "glossy," but they don't give you a sharp-but-faint reflection, like a dirty mirror might. Rather, many glossy objects have BRDFs that send some light in all directions, but more light in directions where the outgoing angle be relatively close to the incident angle reflected across the surface normal. These BRDFs are often modeled using the "Phong" model, after the model described by Bui Tuong Phong in his PhD thesis. According to the Phong model, the extra light in the reflected directions (also called the "specular highlights") fall off polynomially with the the dot product of the outgoing angle with the reflected incident angle. The degree of that polynomial depends on the how shiny or dull the surfaces, with higher-degree polynomials yielding a smaller and more focused specular highlight.

In this thesis, the main focus will be on Lambertian surfaces. When I do consider the possibility of specular of Phong surfaces, I'll go into more detail about what exactly the BRDF model I'm using is at that time. So for the time being, let's consider what can happen using the simplest model that nevertheless describes much of reality very well: a 2D world of Lambertian surfaces.

2.2 The Far-Field Assumption

2.2.1 A point light source and a nearby surface

Let's suppose we live in a 2D world of Lambertian surfaces and diffuse light sources. (When I say a "diffuse" light source, I mean that the light source scatters light equally in all directions. Confusingly, this isn't quite the same thing as a "diffuse" surface—which is another way of saying a Lambertian surface, which actually doesn't quite scatter light equally in all directions, as I explained in the previous section—but that's how these terms are used.) Consider a point light source suspended at $(0, y_p)$, with a Lambertian surface at $y = 0$ (see Figure 2-1). What pattern of illumination can we expect to see on the surface?

The way we proceed with this analysis is to discretize the surface into many small chunks, and then to consider what fraction of the light radiating out of the point light source is hitting any single given small chunk of the surface. We assume each chunk is small enough that its luminance is constant across the chunk. Asking what fraction of light radiating out of the point light source hitting any given chunk is equivalent to asking what angle over the light source is subtended by that chunk, and then dividing that angle by 2π.

Supposing that the chunk extends from $(x_c, 0)$ to $(x_c + dx, 0)$, trigonometry tells us that θ_c, the angle subtended by the chunk, is given by:

$$\theta_c = \tan^{-1}\left(\frac{x_c + dx}{y_p}\right) - \tan^{-1}\left(\frac{x_c}{y_p}\right)$$

What happens as we consider increasingly smaller and smaller chunks dx? The def-

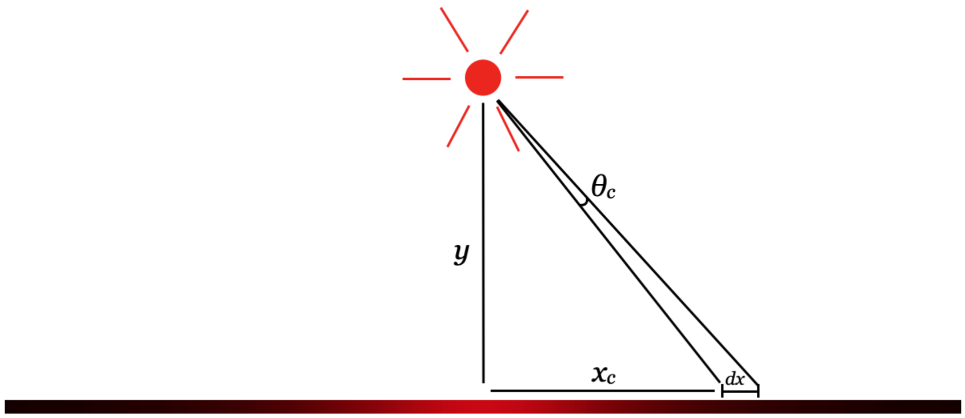


Figure 2-1: A diagram illustrating the following setup: a point light source at $(0, y_p)$, with a Lambertian surface at $y = 0$. We are interested in the resulting illumination pattern on the Lambertian surface; to investigate it, we measure the illumination of a small chunk on the surface that extends from $(x_c, 0)$ to $(x_c + dx, 0)$. The illumination of that chunk will be proportional to the angle θ_c of the point source's light subtended by the chunk.

inition of the derivative tells us that $\lim_{dx \to 0} \theta_c = dx \cdot \frac{d}{dx_c}(\tan^{-1}(\frac{x_c+dx}{y_p})) = dx(y_p/(x^2 + y_p^2))$. Thus the luminance of a chunk on the surface, assuming that the point source had a luminance of 1, would be $dx(y_p/(2\pi(x^2 + y_p^2)))$.[3]We can say that the continuous illumination function of the surface $I(x)$ is the following:

$$I(x) = \frac{y_p}{2\pi(x^2 + y_p^2)}$$

This simple formula captures a lot of interesting phenomena. Consider for instance that we take $x = 0$, meaning we consider the illumination only of the closest point on the surface to the point source. The formula tells us then that the illumination of that point goes as $1/y_p$, meaning that it scales inversely with that point's distance from the point source. Now consider fixing $y_p = 1$ and varying x. This gives us a illumination pattern that scales with $1/(1 + x^2)$, a nice "hump" pattern. The closer the surface is to the point source (meaning a smaller y), the narrower the hump will be. (See Fig. 2-2) Also note that no matter what y_p is, we have:

[3]Readers who are unfamiliar with terms like "luminance" may be confused by a subtle distinction between what I mean by "luminance" versus "illumination pattern" or "brightness." When I talk about the "luminance" of something, I'm referring to the absolute amount of light that thing emits.

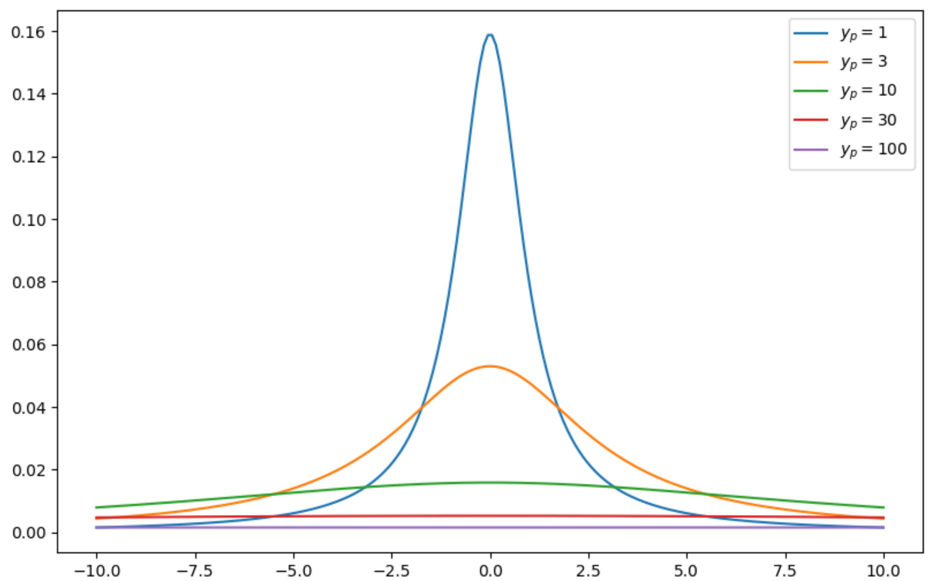


Figure 2-2: Different illumination patterns depending on different possible values of y_p, with the Lambertian surface extending from $x = -10$ to $x = 10$. As this plot shows, the far-field assumption starts to become reasonable for $y_p = 30$.

$$\int_{-\infty}^{\infty} \frac{y_p}{2\pi(x^2 + y_p^2)} dx = 1/2$$

It stands to reason that this is true, because no matter how far the surface is from the point source, if the surface is infinitely broad, exactly half the light from the point source will hit the surface. Additionally, for reference, I'll provide here the illumination function for the equivalent situation in three dimensions: a point source of luminance 1 suspended at $(0, 0, z_p)$, and a plane at $z = 0$. Then, the illumination function $I(x, y)$ can be derived in much the same way as in the two-dimensional case. This function is given by:

$$I(x, y) = \frac{z_p}{4\pi(x^2 + y^2 + z_p^2)^{3/2}}$$

In any case, the important thing to note at this point is that, as shown in Figure 2-2, the illumination pattern becomes flatter and broader the further the point source is from the surface. This phemonenon is what we rely on when we make the "far-field assumption." The far-field assumption is that the assumption that the contribution of a point light source to a faraway surface is approximately constant across that surface. As you can see, this assumption holds as long as the size of the surface in question is much smaller than the distance of the point source to the surface; that is, if, for all relevant values of x, $x^2 \ll y_p^2$, then it follows that $I(x)$ holds a constant value of approximately $1/(2\pi y_p)$ ($1/(4\pi z_p^2)$ in three dimensions), assuming the point source has a luminance of 1.

Because of the quadratic dependence on x and y_p in Eq. **??**, the far-field assumption yields a reasonable approximation even when the difference between x and y_p isn't enormous; for example, if you hold a diffuse light source three meters away from the center of a flat surface two meters in diameter, the brightness of that surface

In contrast, when I use the term "brightness" or "illumination pattern," I'm referring to the light emission *density* of that thing. When you look at a surface, that surface's apparent brightness is proportional to how much light it emits per area of your vision it occupies. So whereas the *luminance* of a small surface chunk in the example given above would be $dx(y/(2\pi(x^2 + y^2)))$, to get the *brightness* of that same surface chunk we'd want to divide by its area; hence its brightness would be $y/(2\pi(x^2 + y^2))$. This makes sense: the apparent brightness of a surface shouldn't depend on how finely you choose to discretize it!

won't vary by more than about 16% (compare $1/9^{3/2}$ to $1/10^{3/2}$). The far-field assumption gets relied on very heavily, both in my research and in work by others, and admittedly the reason for that isn't that it's always a hugely robust assumption to real-world situations (after all, depending on the application, sometimes 16% can matter a lot!). The reason, rather, is that it's an extremely *convenient* assumption. For the time being I'll leave it at that, but in later sections we will see that tolerating the far-field assumption allows us to solve quite a few different optics problems in closed form, or reduce them to easy rather than difficult problems of linear algebra. When I can I will extend my analysis to cases where the far-field assumption cannot be made.

2.3 The Standard Setup

In this section, I will briefly describe what I call "the standard setup," and introduce some terminology that I will use throughout the dissertation. The simplest version of the standard setup is shown in Fig **??**: three parallel frames in flatland, with the "intermediate frame" halfway in between the scene and the observation plane. The presumption is that the observation is a known quantity, and we'd like to infer what's in the scene. Depending on the details of the problem, the intermediate frame may also be a known quantity, or its form may be unknown. In any case, we'd like to see how much we're able to infer about the scene from the observation thanks to (or despite!) the presence of the intermediate frame.

The term "intermediate frame" is left deliberately vague. In an ordinary camera, the intermediate frame would be a lens. In most of this dissertation, I'll be considering intermediate frames that don't directly focus the light from the scene like a lens would, but partially occlude the scene. In principle, there are any number of other realistic intermediate frames.

Of course, there are many other ways to relax the standard setup to make it richer or more realistic. The intermediate frame need not be halfway in between the observation plane; the three frames need not be parallel to each other; the scene need not be planar. And, of course, the real world isn't flatland! But the standard setup is a great starting point for any optical analysis. Colloquially, the form that analysis might take is that an intermediate frame is better for imaging with if, given its presence, the observation tells us more about the scene.

2.3.1 The Transfer Matrix

Another critical concept in my dissertation is the *transfer matrix.* The transfer matrix is a matrix that describes the action of an intermediate frame on the scene to create the observation. To be more precise, suppose we approximate the scene by a vector $\vec{x}$, where each entry of that vector gives the illumination of a single chunk of the scene. Suppose that we approximate the observation plane in the same way with a vector $\vec{y}$. Then, the transfer matrix, A, will be whichever matrix satisfies $\vec{y} = A\vec{x}$ for all possible pairs $(\vec{x}, \vec{y})$.

How do we know that such a matrix even exists for all intermediate frames? Well, if we accept the assumptions implicit the ray-optics model described in Sec. 2.1—that is, we ignore the effects of diffraction and assume that light is incoherent—then what you observe should be a linear function of the presence or absence of light sources. That means that we call what you see if light a is on $f(a)$, and what you see if light b is on $f(b)$, then what you see when both lights are on, $f(a+b)$, should be the sum of what you saw in either case, $f(a) + f(b)$. You can try this at home! If you have a room which is perfectly dark when the lightswitch is off, see what the room looks like when you turn the lights are on versus when you turn on a lamp or flashlight. The brightness of every part of your room when both the lightswitch and lamp are on should roughly correspond to the sum of how bright there were when each were on individually.[4]The fact that this works in most real-world settings tells us that the assumptions of the ray-optics model aren't leading us too far astray.

[4]If you do this, what you see may not "feel" like it actually correspond to the sum of the two room brightnesses. For example, if you have two lights that both illuminate your room equally well, turning both on may not once may not feel like it's giving you a room that's "twice as bright." Make no mistake, though—that's not the fault of the ray optics model, that's the fault of your lying eyes! "Perceived brightness" is a bit of a slippery concept, but it isn't linear in the actual amount of light hitting your retina. In the same way that a 70 dB sound (vacuum cleaner) doesn't sound like it's

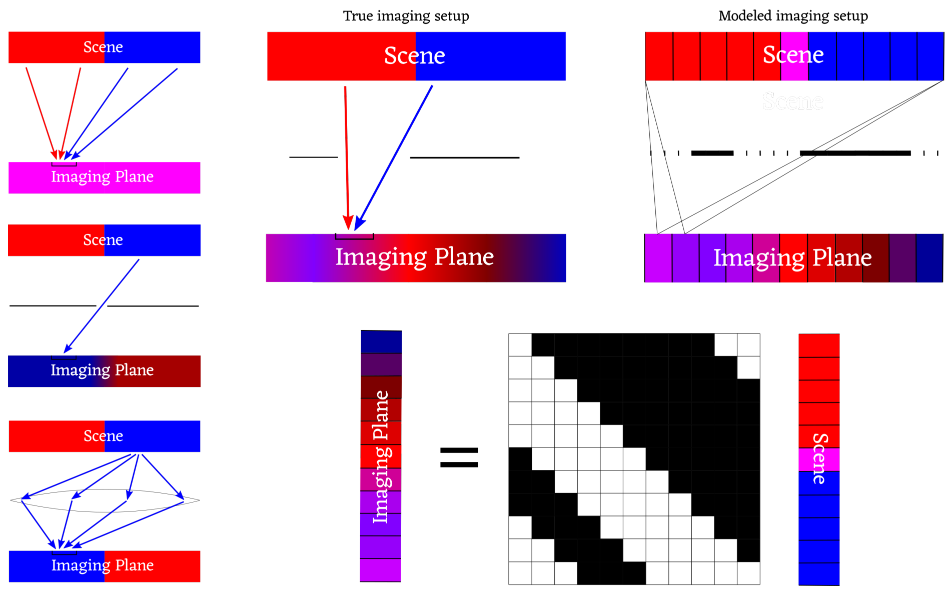


Figure 2-3: Three imaging systems (left, top-to-bottom): no aperture, a pinhole and a lens. Arrows indicate paths light from the scene takes to a particular point on the imaging plane. On the right is an arbitrary mask, an illustration of its discretization and the corresponding transfer matrix.

Because we're assuming that combining light sources behaves linearly, most real-world objects we can put in between a scene and an observation plane should be representable by a transfer matrix A. But what do matrices like these actually look like? Well, they can look like a variety of different things; Figure 2-3 shows the action of a variety of example intermediate frames, with an example transfer matrix corresponding to one of them.

So suppose we have a scene x, an observation y, and a transfer matrix A that represents how the intermediate frame distorts the scene to produce the observation. If $y = Ax$, and we know A and y, what transfer matrices A are best? In the absence of noise, any full-rank transfer matrix A should allow us in principle to perfectly reconstruct $x = A^{-1}y$. That makes the question of which transfer matrix not very interesting—it's a multi-way tie between all full-rank transfer matrices, which make up the vast majority of possible transfer matrices.

"half as loud" as an 80 dB sound (garbage disposal), 100 lumens doesn't look "half as bright" as 200 lumens.

2.3.2 Noise

Of course, it's unrealistic to expect no noise. Every real-world imaging setting will have at least some noise, and in any case it's the presence of noise that makes the problem interesting, and lets us distinguish between better and worse transfer matrices, even if both matrices have the same rank.

Adding noise, our new equation becomes:

$$y = Ax + \eta$$

where η is another vector representing random noise. Now that we have introduced a random variable into our equation, we will need to provide a probability distribution not only for η (to describe how the noise is distributed) but also for x.

Let's start by discussing the probability distribution of the scene vector, x. The simplest model to begin with is to have each entry of the scene be independent and identically distributed (IID), and drawn from a Gaussian. For example, suppose that each entry of the scene vector x was independently drawn from a Gaussian with a mean of μ and a standard deviation of σ. To describe this situation, we can write:

$$x \sim \mathcal{N}(\mu I, \sigma^2)$$

Before we can proceed with this model, there are a few problems for us to worry about. The first is possible negative entries. Real scenes don't cast negative light! To solve this problem, we take $\mu \gg \sigma$. That way, the probability of negative entries will be vanishingly small. A vanishingly small chance of a negative entry is good enough for us; it means that our model's distance from the real world due to this issue (where negative entries are impossible) is also vanishingly small.

The next problem is subtler: recall that x is a discrete vector, but it is meant to represent a continuous scene of fixed size. We haven't yet talked about the number of entries in x, which we'll call n. The variable n controls how finely we discretize the scene x. Ideally, choosing n to be larger will mean that our discrete representation of the true continuous scene will be more faithful (though perhaps at a computational cost). And we might also hope that once n gets large enough, that's a close enough to the real scene that increasing it further won't make the model noticeably better. That's not such an unrealistic expectation; after all, if you're reading these words on a laptop screen, you're probably looking at a discrete array of a couple thousand by a thousand pixels, and that's plenty enough to give you the impression of a "continuous image" on your screen. Tripling the number of pixels on your laptop without increasing the size of your screen probably won't improve your impression of how "continuous" your screen looks by much, unless you're very good at noticing this kind of thing.

We'll talk more about exactly what we mean by this concept later (i.e, how finely we need to discretize the scene before we consider that to be "good enough"). For the time being, though, we have a more serious problem. The problem is this: varying n should give us representations of the true, continuous scene that are varyingly faithful. However, choosing a different value of n shouldn't qualitatively change what the scene looks like. It should always give us the closest discrete approximation possible to the true, continuous scene.

So first, we need to make sure that the total luminance of the scene (in other words, the total amount of light the scene emits) doesn't depend on n. But we said earlier that each entry of x was IID with a mean of μ, so at the moment the total luminance of the scene is $n\mu$. This means that μ is going to have to depend on n; in particular, we'll say that $\mu = J/n$, where J is a constant that represents the total luminance of the scene.

Our difficulties don't stop there, though. If each entry of x is IID and drawn from a Gaussian, then if we want the scene to be qualitatively the same independent of n, σ must also depend on n. In this case, what we mean by "qualitatively the same" is a little bit fuzzy, but we might make it formal by asking that scenes with $n = k$ be drawn from the same distribution as scenes with $n = 2k$ that are then "pixelated" by a factor of 2. (To "pixelate" a vector by a factor of k is to average groups of k contiguous pixels together—for example, if we pixelated the vector [1,2,3,4,5,6] by a factor of 2, we would be left with the vector [1.5, 3.5, 5.5].) By this definition, we

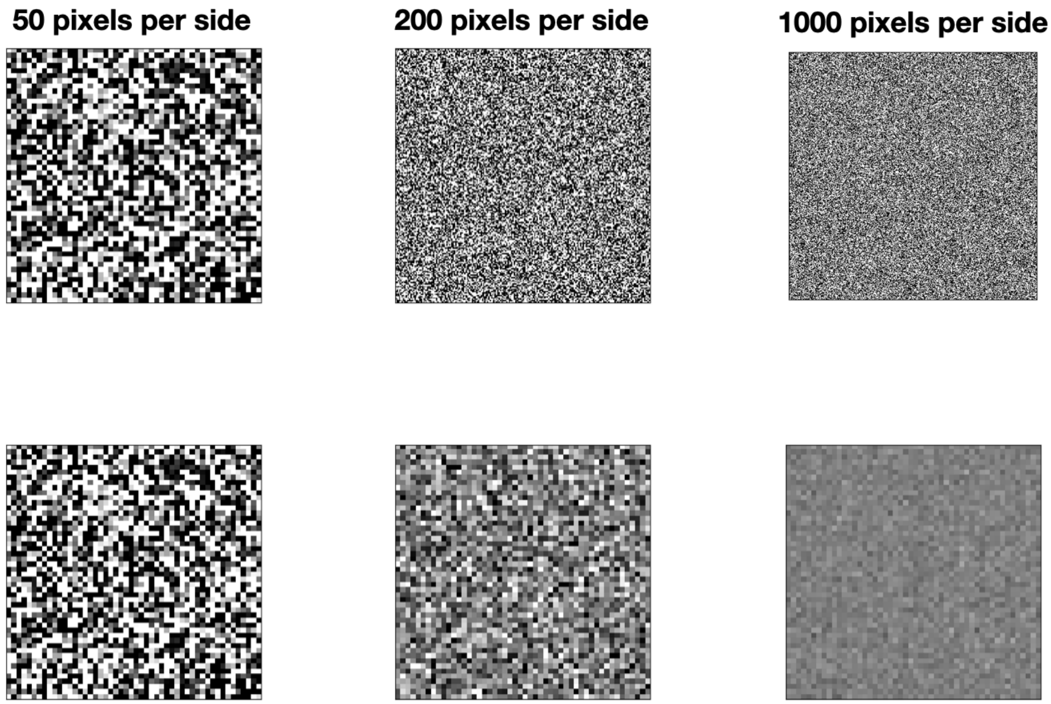


Figure 2-4: Top row: IID scenes with $\sigma = 1$, with three different levels of scene discretization ($n = 50$, $n = 200$, and $n = 1000$). Bottom row: each of those scenes, pixellated so that they have 50 pixels to a side. As you can see, the pixellated scenes look different from each other—this is a problem, because how finely we choose to discretize the scene shouldn't make a difference to what the scene looks like. This is why IID scenes of different levels of discretization are qualitatively different from each other, unlike correlated scenes whose covariance matrices are chosen carefully to scale properly with discretization level.

would need σ to actually *increase* linearly with n in order to have the finer-discretized scenes be less-pixelated versions of the coarser ones! And this is a problem because we said that $\mu \gg \sigma$, and that needs to be true for all values of n, which will be impossible if σ grows with n and μ shrinks with n. If this is hard to follow, see Figure 2-4 for an illustration of this issue.

So what's the solution? The solution is for our model of the scene to include correlations between nearby pixels. We can do this by supposing that the covariance matrix of the scene includes off-diagonal elements:

$$x \sim \mathcal{N}(Q, \sigma^2)$$

The covariance matrix Q captures the correlations between nearby pixels. In real scenes, the closer together two pixels are, the more correlated they'll become. Our model should be faithful to this as well—and in doing so, we will simultaneously create the situation we wanted before, in which scenes with different values of n look qualitatively similar to each other, just at different levels of fidelity.

To make things concrete, I'll provide an example of a scene covariance matrix, which I'll call the *exponential-decay prior.* Recall that an IID covariance matrix would just be a multiple of the identity, $Q = \frac{\theta}{n} I$, where θ is a constant in n. The exponential-decay prior is given by $\mathbf{Q} = \mathbf{F}_n^* \mathbf{D}^\star \mathbf{F}_n$, where $\mathbf{F}_n$ is the normalized DFT matrix of size n and $\mathbf{D}^\star$ is a diagonal matrix with the following entries: $d_1 = 1$, $d_i^\star = d_{n-i+1}^\star = \frac{\theta}{n} \beta^{\frac{i-1}{\lceil (n-1)/2 \rceil}}, i = 2, \ldots, \lceil (n+1)/2 \rceil$, for some frequency decay rate parameter $0 < \beta < 1$. A lower β implies a more strongly correlated scene.

It will be easier to make sense of all this if you look at Figure 2-5 to see for yourself what these covariance matrices look like, and what scenes generated from them look like.

2.3.3 Noise

Now that we have a model of the probability distribution over scenes, we need a model for the probability distribution over the noise. This crucially depends on what application it is we care about. We'll try and make the noise model general enough to apply well to all cases.

We distinguish between two different types of noise.

(Thermal noise): This includes noise sources that are independent of the contribution to the measurements due to the scene of interest. That means that the thing causing the noise isn't light coming from the scene; it's light coming from somewhere else (i.e. "glare"). We model it as additive Gaussian with variance W/n, where W denotes the constant net noise power and each pixel absorbs power proportional to its size, giving

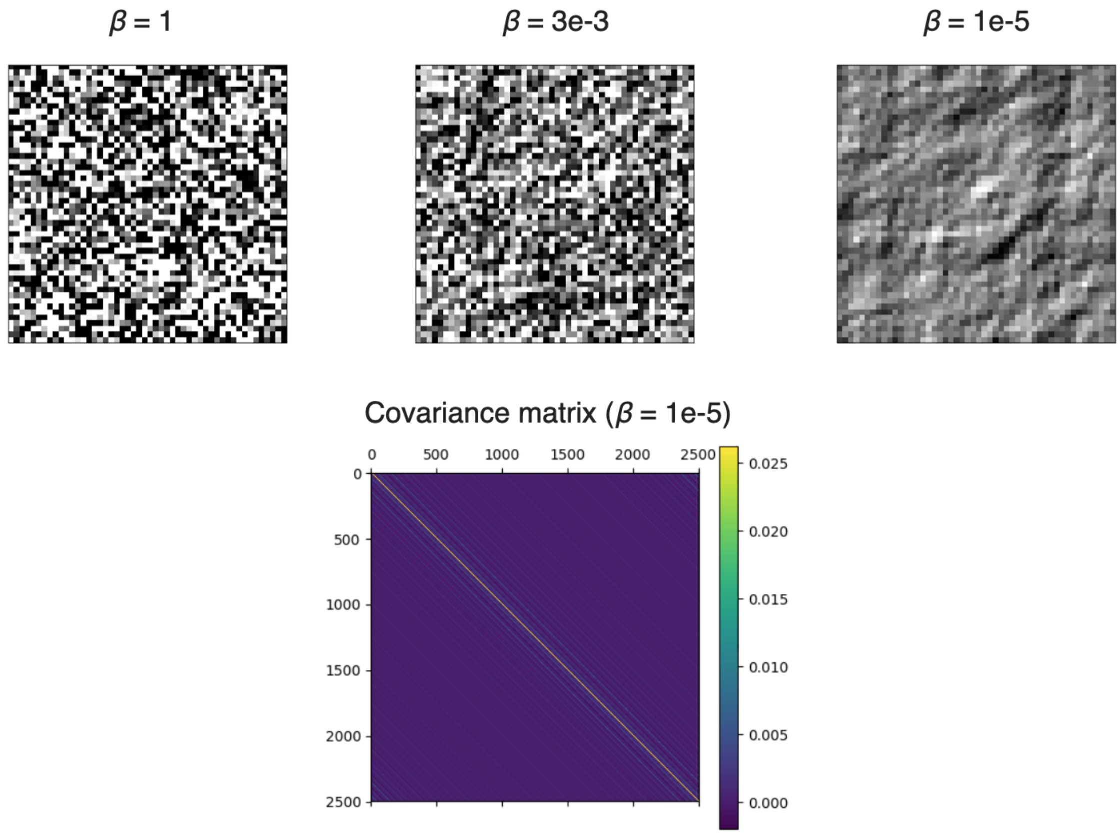


Figure 2-5: Top row: Correlated 50×50 scenes with three different values of β. $\beta = 1$ implies an IID scene, with increasing correlation as β approaches 0. On bottom: the covariance matrix with $\beta = 10^{-5}$. Note that the covariance matrix is 2500×2500, since it describes the covariance of each of the 2500 scene pixels with each other pixel. The pattern of banding that you see depends on how we choose to flatten the 50×50 scene array into a 2500-entry vector; here, the scene is flattened in reading order.

rise to the $1/n$ factor.

(Shot noise): This includes measurement noise that depends on the contribution due to the scene of interest. This results in additive Gaussian noise of variance $\rho \cdot \frac{J}{n}$ (proportional to the net power of light that goes through the aperture).

Thermal noise should be more important in passive non-line-of-sight imaging applications using accidental cameras (in other words, the applications described in Chapter 3) because in that application, the bulk of the light reaching the observation will generally come from sources that aren't the scene of interest, like overhead lighting or sunlight.

Shot noise should be more important in designer-camera applications using coded apertures (in other words, the applications most relevant to the considerations described in Chapter 2) because in those applications, the camera will presumably be designed in such a way as to prevent glare.

Both of these sources of noise are in reality captured by a Poisson distribution, since that's the way light behaves in real-world scenarios. Fortunately, the limit of a Poisson distribution as the number of photons gets very large is a Gaussian, which is much more easily modeled. This approximation is extremely close to reality in the applications we're interested in, which are passive-imaging applications with plenty of light. In super-low-light applications, however, such as active-imaging scenarios where the scene is in perfect darkness except for light introduced by the experimenter through a laser, this modeling issue can become important.

Now that we've established the noise model, we can finally write down the equation relating the scene to the observation.

$$y = Ax + \eta$$

$$x \sim \mathcal{N}(\mu \mathbf{1}, Q)$$

$$\eta \sim \mathcal{N}(0, (W + \rho \cdot J)/n)$$

We'll leave Q ambiguous for now, and talk about it on a case-by-case basis depending on the application. $\mu = J/n$ as described earlier. J is the total radiance of the scene, whereas W is a parameter that describes the level of glare (the scene-independent noise).

2.3.4 Mutual Information

Mutual information. The mutual information (MI) between the measurements $y_j, j \in [n]$ and the unknowns $f_i, i \in [1, n]$ of the imaging problem is given as $\mathcal{I} = \log\det\left(\frac{1}{\sigma^2}\widetilde{\mathbf{A}}\mathbf{Q}\widetilde{\mathbf{A}}^T + \mathbf{I}\right)$, where the noise variance $\sigma^2 = (W + \rho \cdot J)/n$.

2.4 High SNR, IID scene, Occluding mask

In this section, I will go into great detail about a regime that appears unphysical, but gives us a lot of insight into a variety of different important regimes. It also has important mathematical implications.

Consider the mutual information equation from the previous section:
$\mathcal{I} = \log\det\left(\frac{1}{\sigma^2}\widetilde{\mathbf{A}}\mathbf{Q}\widetilde{\mathbf{A}}^T + \mathbf{I}\right)$

Suppose we make the following assumptions:

1. The scene is IID, $Q = kI$.
2. The SNR is very good, $\sigma \ll k$, so we can ignore the identity term.
3. The intermediate frame only occludes light or lets it through; it doesn't redirect the light.
4. The far-field assumption applies. This point and the previous one let us assume that the transfer matrix $\widetilde{\mathbf{A}}$ is a multiple of a binary-valued Toeplitz matrix.

Suppose now that we want to find the best possible intermediate frame under these conditions, i.e. we want to find the best possible transfer matrix that satisfies the constraints above.

What transfer matrix maximizes the mutual information? Naturally, it will be whichever transfer matrix maximizes $\log\det\left(\frac{1}{\sigma^2}\widetilde{\mathbf{A}}\mathbf{Q}\widetilde{\mathbf{A}}^T + \mathbf{I}\right)$, which given our assumptions is equivalent to maximizing the determinant of $\widetilde{\mathbf{A}}^T\widetilde{\mathbf{A}}$. This corresponds to maximizing the product of the norms of the transfer matrix's singular values, since each eigenvalue of $\widetilde{\mathbf{A}}^T\widetilde{\mathbf{A}}$ corresponds to the norm squared of one of the eigenvalues of $\widetilde{\mathbf{A}}$.

If we further assume that $\widetilde{\mathbf{A}}$ is circulant, not just Toeplitz, then that tells us that in fact the eigenvalues and singular values of $\widetilde{\mathbf{A}}$ have the same norm. This is a convenient assumption that doesn't necessarily match reality. Under which real-world conditions will the transfer matrix actually be circulant rather than Toeplitz? This is somewhat unintuitive, but it corresponds to the scenario in which the occluding pattern of the intermediate frame repeats itself once. Once is enough! See Figure **??** for a visual explanation of why this is.

Assuming that the transfer matrix is circulant, not just Toeplitz, is tremendously convenient. It means that we can compute the mutual information between the scene and the observation extremely efficiently, since the eigenvalues λ_i are given by the Fourier transform of the first row of the transfer matrix, for which the time to compute is log-linear in n (as opposed to $O(n^3)$ for a general determinant). And since $|\lambda_i| = |\sigma_i|$, that gives us all the information we need to compute $\log\det(\widetilde{\mathbf{A}}\widetilde{\mathbf{A}}^T)$.

But realistically speaking, can we restrict ourselves just to circulant transfer matrices rather than Toeplitz ones? After all, occluders that give rise to that kind of transfer matrix make up only a very small fraction of all possible occluders. The answer is that it depends on what you want, but if you're only concerned with finding the *best* possible occluder, restricting your attention to circulant transfer matrices costs you very little. Here's why.

2.4.1 Hadamard's Bound

The point of this section is to justify the following bound on all $\{0, 1\}$ matrices B. If this doesn't interest you, skip to the next section.

$$|\det B| =\leqslant 2^{-n}(n+1)^{(n+1)/2}, \tag{2.1}$$

By a *binary* matrix we mean a matrix whose elements are in one of the sets

$S_{01} := \{0, 1\}$ or $S_{\pm 1} := \{-1, 1\}$. It will be clear from the context which of these two cases is being considered. A *binary circulant* is a circulant matrix whose elements are in S_{01} or $S_{\pm 1}$.

There is a natural correspondence between the integers $\{0, 1, \ldots 2^n - 1\}$ and the binary circulant matrices of order n. If $N \in \{0, 1, \ldots, 2^n - 1\}$ has the representation

$$N = \sum_{j=0}^{n-1} 2^{n-1-j} b_j,$$

so may be written in binary as $b_0 \ldots b_{n-1}$, we associate N with $\mathrm{circ}(a_0, \ldots, a_{n-1})$, where $a_j = b_j$ in the case of S_{01}, and $a_j = 2b_j - 1$ in the case of $S_{\pm 1}$.

The *maximal determinant problem* is concerned with the maximal value of $|\det A|$ for an $n \times n$ binary matrix A. The *Hadamard bound* [**?**] states that, in the case of binary matrices A over $\{\pm 1\}$, we have

$$|\det A| \leqslant n^{n/2}. \tag{2.2}$$

Moreover, Hadamard's inequality is sharp for infinitely many n, for example powers of two or n of the form $q+1$ where q is a prime power and $q \equiv 3 \pmod 4$ (Paley [**?**]).

There is a well-known connection between the determinants of $\{0, 1\}$-matrices of order n and $\{\pm 1\}$-matrices of order $n + 1$. This implies that an $(n + 1) \times (n + 1)$ $\{\pm 1\}$-matrix always has determinant divisible by 2^n. See [**?**] for details. We give an example with $n = 3$, starting with an $n \times n$ binary matrix B and ending with an $(n + 1) \times (n + 1)$ $\{\pm 1\}$-matrix A, with $\det A = 2^n \det(B)$.

$$B = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \xrightarrow{\text{double}} \begin{pmatrix} 2 & 0 & 2 \\ 2 & 2 & 0 \\ 0 & 2 & 2 \end{pmatrix}$$

$$\xrightarrow{\text{border}} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 0 & 2 \\ 0 & 2 & 2 & 0 \\ 0 & 0 & 2 & 2 \end{pmatrix} \xrightarrow[\text{first row}]{\text{subtract}} \begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & -1 & 1 \\ -1 & 1 & 1 & -1 \\ -1 & -1 & 1 & 1 \end{pmatrix} = A.$$

The doubling step is the only step where the determinant changes, and there it is multiplied by 2^n.

Thus, Hadamard's bound (2.2) gives the bound

$$|\det B| = 2^{-n}|\det A| \leqslant 2^{-n}(n+1)^{(n+1)/2}, \tag{2.3}$$

which applies for all $\{0, 1\}$-matrices B of order n. We shall refer to both (2.2) and (2.3) as *Hadamard's inequality*, since it will be clear from the context which inequality is intended.[5]

2.4.2 Binary Circulants Achieving Hadamard's Bound

Now we get to the reason that, if you're only concerned with finding the *best* possible occluder, restricting your attention to circulant transfer matrices costs you very little. The reason is that there are several constructions for binary circulant matrices that achieve Hadamard's bound—this *despite* the fact that Hadamard's bound is general for all binary matrices, not just circulant ones! Please take a moment to reflect on how lucky it is that among the binary circulant matrices, which comprise a tiny subset of all binary matrices that, some of those binary circulant matrices achieve determinants as large as any from *all* binary matrices of the same size. I'll call these remarkable binary circulants "determinant-maximizing binary circulants," or DMBCs.

There are four known constructions for DMBCs. All four of these constructions yield a matrix whose eigenvalues are all equal, save for the first; the first eigenvalue has a value of $(n + 1)/2$, and every other eigenvalue has a value of $\sqrt{(}n + 1)/2$. The

[5] In fact, Hadamard in [?] proved a more general inequality than (2.2), and as far as we are aware he never stated (2.3) explicitly. A simple proof of (2.2) is given by Cameron [?].

dot product of any pair of rows in a DMBC from one of these four constructions is $(n+1)/4$. In each construction, there are $(n+1)/2$ 1s and $(n-1)/2$ -1s. It follows from that last sentence that if you take one of these constructions and replace all the 0s with -1s to yield a $\{1,-1\}$ matrix, the dot product of any pair of rows will be -1.

Thanks to this last point, DMBCs have a close relationship to Hadamard matrices. Hadamard matrices are $\{1,-1\}$ matrices (not necessarily circulant) for which every pair of rows is orthogonal, that is, the dot product of any pair of rows is 0. (The same is true of pairs of columns.) Any size-n DMBC from one of these four constructions can be easily adapted to create a size-$(n+1)$ Hadamard matrix as follows: replace all the 0s with -1s in the DMBC, and then add a first row and a first column of all -1s. The dot product of the first row with any other row will be 0 (because every other but the first is exactly half 1s and half -1s). The dot product of every row other than the first with any other row other than the first will also be 0 (because the dot product of each pair of rows before adding the extra row and column was -1, and then adding the extra column adds an extra 1 to the dot product). Hence each size-n DMBC yields a size-$(n+1)$ Hadamard matrix.

Hadamard matrices of this kind are known in the literature as "Hadamard matrices with circulant core," for obvious reasons. Now I will go into a little bit more detail about the various constructions.

Theorem 1 (Hadamard circulant core construction)**.** *A Hadamard matrix of order $n+1$ with circulant core of order n exists if*

(1) $n \equiv 3 \pmod 4$ is a prime;

(2) $n = p(p+2)$, where p and $p+2$ are prime;

(3) $n = 2^k - 1$, where k is a positive integer; or

(4) $n = 4k^2 + 27$, where k is a positive integer and n is a prime.

Proof. Case (1) is due to Paley [**?**]; case (2) is due to Stanton and Sprott [**?**] and also Whiteman [**?**]; case (3) is due to Singer [**?**]; and case (4) is due to Hall [**?**, Theorem 2.2]. □

Hall [**?**, p. 980] remarks that case (4) is subsumed by case (1), since $4k^2 + 27 \equiv 3 \pmod 4$, but we mention case (4) since Hall's construction is different from that of Paley.

We do not know if the list given by Theorem 1 is exhaustive. The computational results given in Tables **??**–**??** show that, for $1 \leqslant n \leqslant 52$, only those n given by Theorem 1 can provide a Hadamard matrix of order $n+1$ with a circulant core. Also, a circulant $\{0,1\}$-matrix of order $n \leqslant 52$ can achieve the upper bound (**??**) if and only if $n \leqslant 4$ or n satisfies condition (1), (2) or (3) of Theorem 1.

This gives us the four known constructions for DMBCs. Note that the fourth construction is completely redundant with the first, since any prime n such that $n = 4k^2 + 27$ where k is a positive integer is guaranteed to also be prime and congruent to 3 mod 4! It is only considered a separate construction because it yields an additional DMBC beyond the one given by the first construction (that isn't a trivial transformation). The fourth construction is therefore of no additional practical value to us: we can't use it to create a mask that images at a given level of resolution that we couldn't already. It's quite interesting mathematically, but doesn't help us solve an imaging problem.

It's worth giving more detail about the first construction, since it's by far the most common one over the real numbers; there are many more primes congruent to 3 mod 4 than there are powers of 2 or products of twin primes! Indeed, it's common enough that no matter what level of resolution you need, there will be a reasonably suitable mask at a nearby resolution level, thanks to the first construction.

The first construction, due to Paley, is as follows:

If n is prime and congruent to 3 mod 4:

$$x_i = \begin{cases} 1 & \text{if } \left(\frac{i}{n}\right) = 0 \text{ or } 1 \\ 0 & \text{otherwise} \end{cases}$$

Here, $\left(\frac{i}{n}\right)$ is the Legendre symbol. It's equal to 0 is a multiple of n, and is otherwise equal to 1 if i is a quadratic residue (meaning a perfect square) modulo p and -1 if not. It's very easily computed, since by Euler's criterion, we have, for any a and any prime p:

$$\left(\frac{a}{p}\right) \equiv a^{(p-1)/2} \mod p$$

Figure 2-6 shows the first few sequences of the Paley construction. As you can see, it has a very "random-looking" appearance. Of course, the construction is very much non-random; what gives it that appearance, though, is its non-self-repeating character. All four constructions, in fact, try to repeat themselves as little as possible. That's not surprising from the point of view of wanting to keep the sequence's Fourier spectrum flat, of course. But the way I like to think of it from an imaging point of view is that these sequences try and make the different possible shadows cast by a light source in a variety of different locations as different as possible from each other. A light source at each possible point in the (implicitly planar) scene will yield a different rotation of the occluding sequence, so the best sequences for distinguishing light sources at different locations will be sequences that are orthogonal to their own rotations. See Figure 2-7 for a visual explanation of this phenomenon.

In any event, because DMBCs achieve the maximal possible mutual information of any binary matrix of their size, the fact that we restricted our attention to circulant matrices rather than considering all possible Toeplitz matrices doesn't matter, assuming the value of n we're using admits the existence of a DMBC. We therefore know that the once-repeating occluder suggested by that DMBC outperforms all other possible occluding intermediate frames, including ones that don't repeat themslves.

Why is this useful, if the notion of an IID scene doesn't make sense, as we explained in the previous section? After all, these sequences are only optimal assuming an IID scene, and each different value of n yields a qualitatively difference scene model. How can we know which of these non-repeating sequences to apply in real life?

The answer is that even if "IID scenes" with different values of n are qualitatively different from each other, that doesn't mean that each one doesn't describe a reasonable approximation of reality. Each different value of n can be thought of as describing

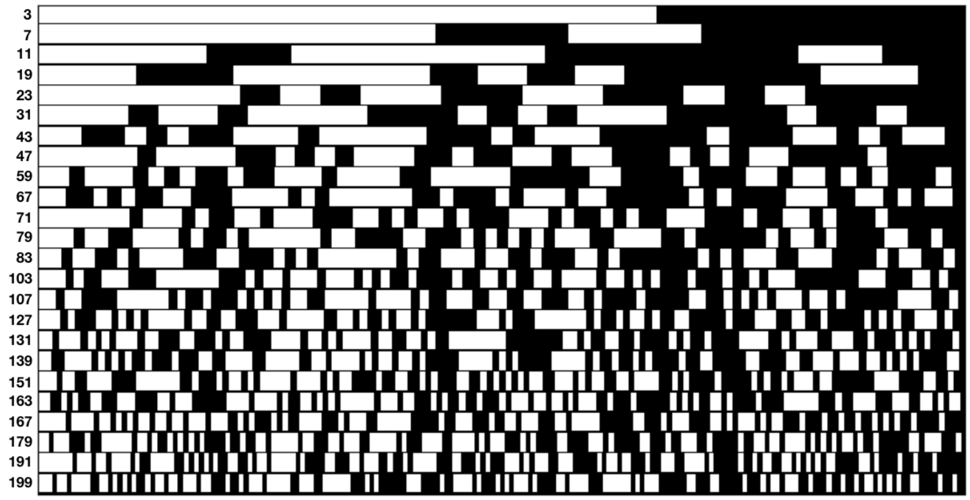


Figure 2-6: These are the first few Paley sequences, i.e. on-off patterns whose spectrum is flat and that yield a DMBC (determinant-maximizing binary circulant) when used as the first row of a circulant matrix. On the left is n, the number of on/off chunks used. Note that for a Paley sequence to exist, n must be prime and congruent to 3 mod 4.

a different model, with lower values of n describing scenes with fewer "effective pixels" and higher values describing scenes with more "effective pixels." What do I mean by the number of "effective pixels" that a scene has? Well, roughly speaking, it's the number of pixels you need to get a reasonable view of the scene; you can think of a 144p video as having 144 effective pixels, even if more pixels than that are used to display it on your screen. More correlated scenes will have fewer effective pixels, and less correlated scenes will have more; having a lower SNR will also mean you have fewer effective pixels, and a higher SNR will mean you have more.

Figure 2-9 is a plot of the approximate number of effective pixels in the system, as a function of the SNR and β (i.e. the level of correlation in the scene).

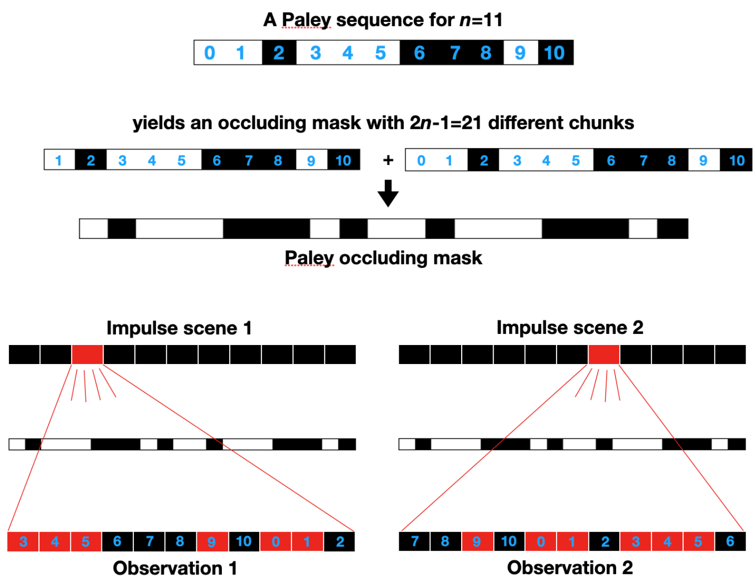
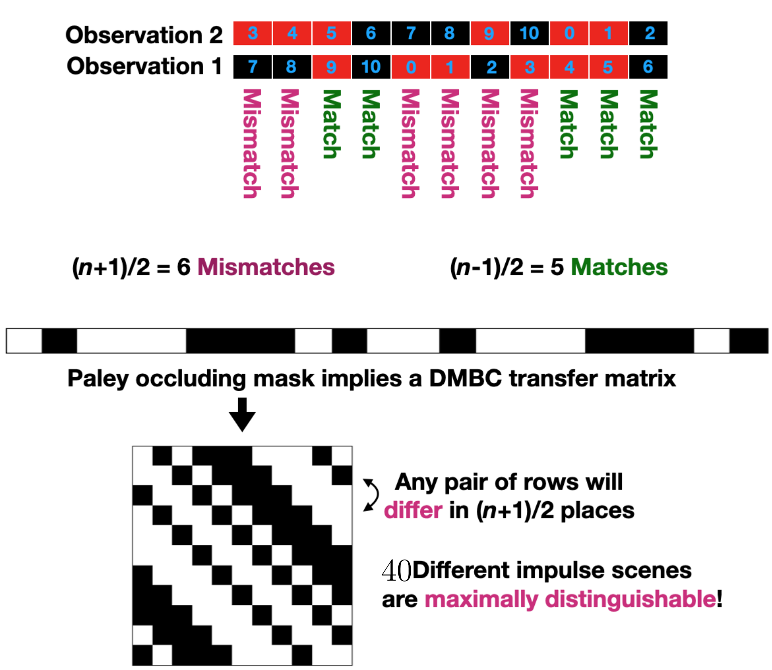


Figure 2-7: A visual explanation of Paley sequences, the "optimal" occluding masks that be constructed from them, and the transfer matrices associated with those masks. Top: we consider the Paley construction applied to the case $n = 11$. The paley construction gives us a binary sequence, with 1's (white, or "on") element i of the sequence if i is 0 or a quadratic residue modulo n, and 0's (black, or "off") for elements i of the sequence if i is not a quadratic residue modulo n. From this sequence we get an occluding mask, which consists of the sequence twice; every element is repeated exactly once except for the 0 element, which is in the center. Given the far-field assumption and assuming the occluder is halfway in between observation and scene, an impulse scene will cast a shadow that corresponds to half the occluder, which is some rotation of the original Paley sequence. Bottom: sequences given by the Paley construction (as well as any other flat sequence) are as different as possible from their own rotations; this means that depending on where the impulse light source is in the scene, the cast shadows will be as different as possible. This makes reconstructing the location of a point light source, given a cast shadow, as easy as possible.



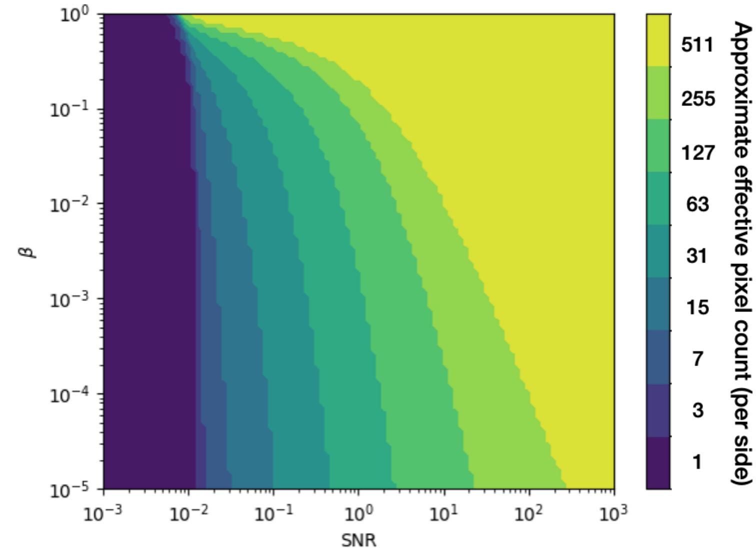
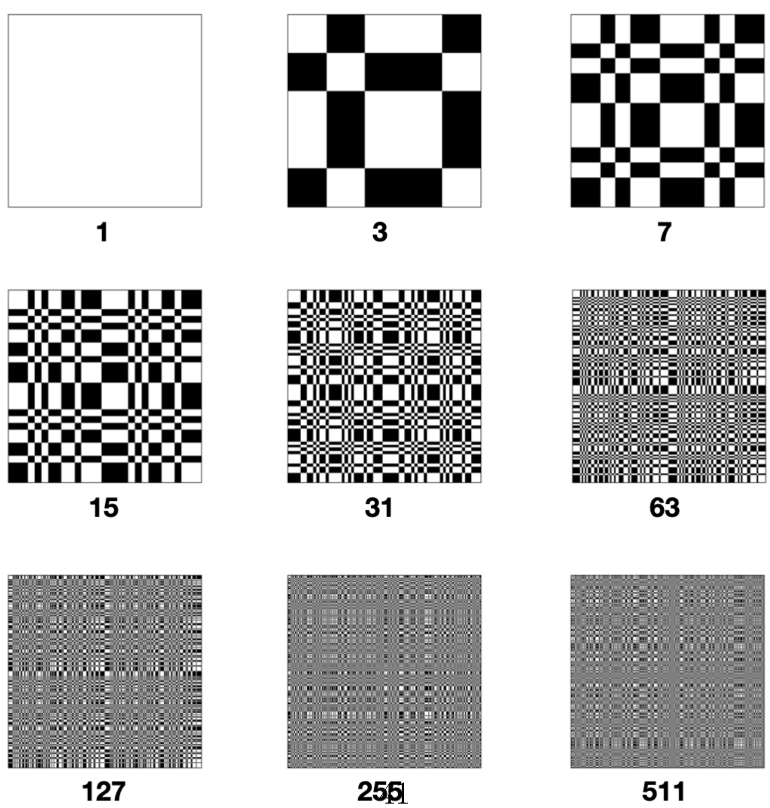


Figure 2-9: Top: the approximate effective pixel count of scenes generated with a given level of correlation (β) and under a given signal-to-noise ratio (SNR). As expected, more correlated scenes and noisier scenes both have fewer effective pixels. Note, however, that even highly correlated scenes can have high effective pixel counts if the SNR is high enough, but if the SNR is low enough the scene will always have a low effective pixel count. Effective pixel count was estimated by choosing which of the nine spectrally-flat masks shown on bottom yielded the highest mutual information. Bottom: masks corresponding to each of the effective scene pixel counts. Note that these masks repeat themselves once in each dimension, so each mask is $2n-1 \times 2n-1$ if the effective pixel count is n. This is due to the phenomenon described in Figure **??**.



2.5 Varying the distance between observation, occluder, and scene

We continue examining each of the idealized model's assumptions one after the other. Next on the docket is the assumption that the occluder lies exactly halfway in between the scene and observation plane. This was a tremenously convenient assumption because it allowed us to assume that the occluder's transfer matrix had Toeplitz structure. But in the real world, the assumption is completely unrealistic. In a designer-mask camera application, the occluder will presumably be much closer to the camera's photosensitive material than to the scene, and even in an accidental-camera application, we can't assume that the occluder will be exactly halfway between the wall we're looking at and whatever it is we're trying to image. So let's try removing the assumption and seeing what happens. Note that we'll still be assuming that scene and occluder are both planar—we'll get to that eventually, but not yet. And we're still using the far-field assumption—meaning that regardless of what we are taking the *relative* distances of the occluder to the scene and observation to be, we are always assuming that the distance between scene and observation to be much bigger than the size of the scene or observation.

What exactly is it about an occluder halfway between the scene and observation that gives us Toeplitz transfer matrices? The answer is that when the occluder is halfway bewteen the scene and observation, the shadow cast by a moving light source will move at exactly the speed the light source is moving, but in the opposite direction. Try holding a flashlight (such as one from a smartphone) with your right hand, illuminating a table or a wall, and then hold your left hand halfway in between the flashlight and the table. (I encourage you to actually do this!) Keep your left hand steady, and then move the flashlight around. You can see that your hand's shadow moves at the same speed your flashlight does, and in the opposite direction.

Now try varying the height of your left hand relative to the table. What happens to the speed of your hand's shadow relative to the speed at which you move the flashlight? The answer is that when your hand is closer to the table than to the

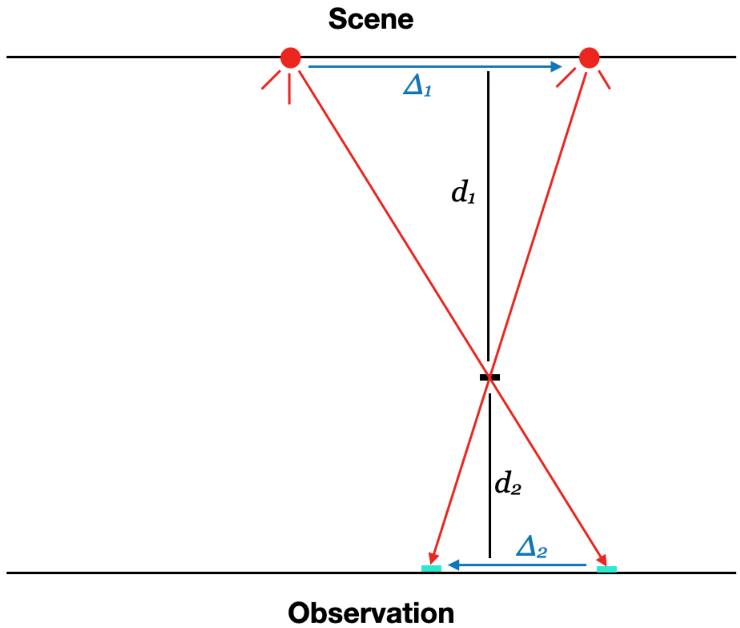


Figure 2-11: A simple layout that explains the phenomenon whereby the relative speeds of a light source and its shadow are given by the relative distances of the scene and the observation to the occluder. Suppose we have a pinspeck occluder, and a light source that moves by an amount Δ_1 to the right. If we suppose that its shadow moves by an amount Δ_2 to the left, and that the occluder has a perpendicular distance d_1 from the scene and d_2 from the observation, then the fact that the top and bottom triangles are similar tells us that $\Delta_1/\Delta_2 = d_1/d_2$.

flashlight, your hand's shadow will move slower than the flashlight; and when your hand is closer to the flashlight than to the table, your hand's shadow will move faster than the flashlight. (Of course, your hand's shadow will always in the opposite direction from the shadow—that part won't change.)

In fact, to be more precise, the "speed multiplier" that your hand's shadow gets relative to the flashlight—that is, your hand's shadow's speed divided by the flashlight's speed—is the same as the distance between your hand and the table divided by the distance between your hand and the flashlight. Figure **??** gives a visual explanation of this phenomenon.

This "speed multiplier" conceit is crucial to understanding how varying the occluder's depth warps the resulting transfer matrix. Remember that each column of the transfer matrix tells us what the observation will look like in response to a point light source at each different location in the scene. If we imagine, then, a point light source moving at a constant speed of 1 space-unit per time-unit across the scene, then we can imagine the transfer matrix as a movie of the observation plane while that happens, with each column of the transfer matrix being one frame of that movie.

When the occluder is halfway in between the scene and observation plane, we know exactly what that movie should look like: the shadow should move at the same speed as the point light source. That is, it should move at a speed of 1 space-unit (1 "bin," or $1/n$) per time unit (1 "frame," or column of the transfer matrix), assuming we discretize the scene and the observation plane equally finely.

It's for this reason that the occluder being halfway between the scene and observation gives us the perfect, constant diagonals that characterize a Toeplitz matrix. Compare a column of the transfer matrix to the column adjacent to it, and you should see a copy of that column, but shifted by one row.

What if we continue to imagine that the transfer matrix is a movie of the shadow cast by a point light source moving a constant speed of 1 space-unit per time-unit—but now we supposed that the occluder was twice as close to the scene as it was to the observation plane? We know from Figure **??** that that means that the shadow must move at a speed of 2 space-units per time-unit. Therefore, on the transfer matrix, moving one column (time-unit) to the right will cause the shadow to shift two rows (space-units) down. (Remember that we are sticking to our convention of labeling the observation plane right-to-left instead of left-to-right, as explained in Section **??**—if we weren't, that would cause the shadow to shift two rows *up*!)

Similarly, if the occluder was twice as close to the observation plane as to the scene, the shadow would move at a speed of 0.5 space-units per time-unit. And if the occluder was right up against the observation plane, the shadow wouldn't move at all—and if the occluder was right up against the scene, there would be no shadow! Figure **??** shows some example transfer matrices for each of these scenarios.

The Lambertian BRDF is an excellent approximation for the reflectance properties of most objects.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aliquam quis sollicitudin

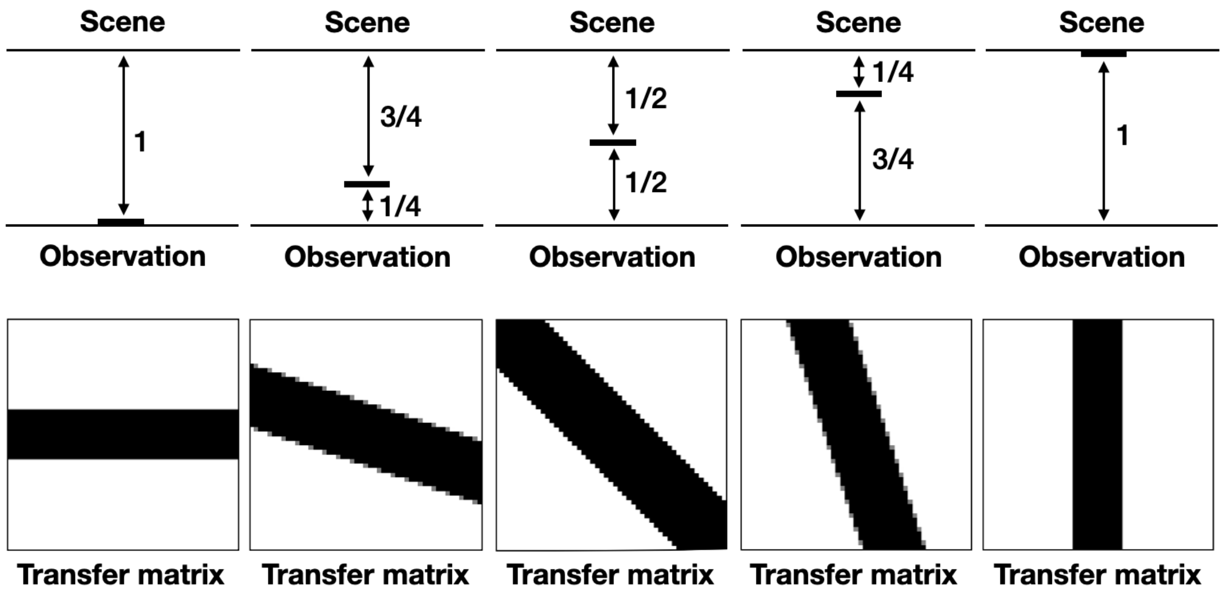


Figure 2-12: Top row: five different scenarios with the occluder at five different depths. Bottom row: the transfer matrices corresponding to each different scenario.

metus. Quisque quam ex, tincidunt et porttitor quis, tincidunt faucibus quam. Nulla facilisi. Nam a libero posuere, mattis leo ac, ultrices est. Nullam auctor lacus eu metus venenatis, gravida consectetur felis laoreet. Nam non ante felis. Maecenas id dignissim turpis, eget pulvinar nisl. Cras nec mauris feugiat, aliquam elit ac, blandit ex [?].

Nunc vehicula velit nibh, ut facilisis magna blandit semper. Fusce varius odio in velit mattis porttitor. Vivamus aliquam vulputate quam, non auctor augue ornare ut. Vivamus ac pretium sapien, sed gravida elit. Etiam facilisis risus libero, sed vulputate velit lobortis sed. Aliquam erat volutpat. Sed est orci, dignissim nec varius et, auctor non purus [?, ?]. Morbi feugiat gravida ipsum. Sed maximus nibh eget feugiat tempor. Pellentesque nec urna varius, volutpat dolor at, mattis leo. Vivamus elit eros, pretium sed magna nec, faucibus egestas enim. Vivamus ipsum ex, condimentum eu felis ac, tempus feugiat metus discussed in section ??.

2.6 Section sample 1

Ut hendrerit risus egestas, sollicitudin mauris sit amet, fermentum ipsum. Donec vulputate enim in justo pellentesque rhoncus. Nunc a dui condimentum, egestas ipsum eu, fermentum enim. Duis condimentum iaculis luctus. Nam sodales pellentesque luctus. Aenean tristique ante mattis tellus tincidunt, vitae mattis nunc tristique. Ut nec mattis ante, eu sodales ex.

Praesent pulvinar risus in diam mollis tincidunt. Nam aliquam lacus sed eleifend mattis. Ut at blandit ex, eget molestie massa. Morbi fermentum sit amet mauris ut placerat. Ut in eros pretium, congue felis sit amet, gravida justo. Suspendisse purus leo, posuere sed odio vel, mollis tempus nisi. Aenean est tortor, tincidunt et rhoncus ut, fermentum semper leo [?]. Integer id viverra metus, a blandit neque. In in enim eu ipsum euismod hendrerit.

Cras et pellentesque sapien. Maecenas vitae sollicitudin sapien. Sed elementum feugiat ligula, eu maximus nunc porttitor in. Etiam porttitor nisi ante, vel fermentum arcu pulvinar et. Sed bibendum diam nisl, vitae dapibus elit convallis congue. Ut vehicula lectus et enim consectetur cursus. Praesent non viverra augue, id bibendum risus. Maecenas lorem massa, dignissim ac purus in, tincidunt sodales nisi. Praesent condimentum tempus mauris. Etiam vitae sem maximus, auctor orci at, rhoncus diam. Donec pretium sodales sodales. Donec sodales ultrices metus ac pharetra. Mauris non ullamcorper urna. Mauris ac faucibus tortor, eu lobortis leo are described in detail in section **??**.

2.7 Section sample 2

Phasellus sed elit vehicula, gravida odio in, vulputate quam. Quisque in elit enim. Vivamus finibus justo elit, sed semper turpis aliquam porttitor. Nulla posuere bibendum nunc sit amet consequat. Vivamus commodo lorem sed metus fermentum rhoncus. Etiam porta sodales purus, vel aliquet lacus facilisis et. Etiam ornare velit non dui auctor fermentum. In elit augue, fringilla at lacinia at, facilisis sit amet lectus. Sed et hendrerit ex. Morbi tristique felis a augue egestas commodo. Nulla porttitor ut urna nec dignissim. Fusce ac pharetra risus, id rhoncus ligula. Pellentesque euismod viverra sem, vel porttitor libero blandit quis. Phasellus orci augue, mattis nec dolor ut, cursus mattis quam. Sed tincidunt eu metus sed pulvinar. Ut a nulla at leo semper accumsan efficitur eget leo.

Sed vel lectus ut dui tempor molestie. Suspendisse blandit sapien posuere quam tempor lobortis. Duis sollicitudin tincidunt dui, at aliquam lorem dictum sit amet. Aenean congue nibh lectus, ut faucibus turpis facilisis quis. Ut aliquet magna at placerat ultricies. Mauris convallis, risus efficitur gravida dapibus, lacus lorem malesuada ligula, eget porta diam felis non turpis. Nulla sed sem finibus, vehicula quam at, vulputate tellus[6]

Quisque elit enim, molestie ac metus ut, condimentum convallis nibh:

2.7.1 Subsection sample

In tempus ex nibh, non eleifend risus iaculis ac. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Nullam in nisi eu arcu laoreet sollicitudin. Mauris consectetur venenatis arcu id finibus. Aenean pellentesque consectetur erat lacinia vulputate. Praesent tempus tempus lorem at dignissim. Proin at odio vitae tortor sollicitudin pretium. Quisque ac purus eu sem rutrum bibendum.

Pellentesque ac leo eget lorem vulputate mattis eu a nisl. Duis elit erat, consectetur vulputate ullamcorper a, finibus quis turpis. Vivamus tincidunt dui id purus bibendum malesuada. Fusce accumsan, ipsum quis feugiat sodales, enim est aliquet leo, ut ornare justo mauris quis ex. Sed eros magna, suscipit et blandit non, pretium id felis. Praesent a vehicula tortor. Donec blandit dolor a ipsum sodales, eget aliquet nisl fermentum.

[6] Here is a sample footnote referencing figures **??** and **??**.

2.7.2 Subsection with list

Ut sollicitudin, lectus eget posuere porttitor, risus dui facilisis risus, a pharetra lacus elit vel eros. Proin fermentum accumsan mauris, quis posuere nisi pharetra scelerisque.

1. Item 1.
2. Item 2.
3. Item 3.

Cras nec ullamcorper mauris. Aliquam erat volutpat. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Suspendisse sed dui ac mi auctor scelerisque. Etiam at semper nisi. Cras nec dolor ac purus feugiat auctor. Nunc eget pulvinar massa.

2.8 Section sample 3

Quisque sed ultrices leo. Donec vestibulum auctor nibh, at faucibus libero mollis in. Quisque massa lorem, feugiat a lectus in, lobortis volutpat lectus. Donec accumsan dui erat, eu tempor tortor facilisis sed. Nulla ullamcorper augue et sapien dapibus, quis bibendum velit porta. Nullam mattis vehicula tortor porttitor porta. Interdum et malesuada fames ac ante ipsum primis in faucibus. Praesent suscipit, lorem vel viverra rhoncus, turpis orci dignissim dui, bibendum pulvinar justo sem vel lorem. Nam porttitor mollis tristique. Aliquam rhoncus magna quis nisl varius mattis. Sed rhoncus, diam in gravida iaculis, mauris tellus imperdiet turpis, at porttitor est leo vel velit. Praesent faucibus ornare sodales. Sed eu lorem purus.

2.8.1 Another subsection sample

Nullam rhoncus posuere lacus, id volutpat nisi pulvinar viverra. Quisque quis ultricies ante. Duis sollicitudin sapien nec consequat vehicula. Vestibulum convallis erat in arcu aliquam eleifend. Nunc scelerisque lorem non luctus sodales. Curabitur eleifend odio et sagittis semper. Praesent sodales, diam nec vulputate iaculis, neque leo consectetur nunc, a luctus lacus purus et dui. Sed sit amet tortor ullamcorper, malesuada libero quis, imperdiet tortor. Cras tempor blandit massa, sit amet molestie sapien tincidunt quis. Nullam hendrerit venenatis massa, sed lacinia ligula tincidunt vitae.

Nam efficitur et lacus sed eleifend. Aenean quis ipsum eget leo ultrices ornare. Nullam rhoncus ante odio, at dignissim neque posuere eu. Pellentesque sodales tortor est, nec egestas sapien mollis quis. In lectus sapien, pellentesque congue erat consequat, hendrerit aliquet elit. Pellentesque eleifend purus ac diam bibendum, ac auctor ipsum posuere. Cras suscipit leo nec velit fermentum, id varius erat eleifend. Proin sagittis purus id ante lacinia, et congue eros tincidunt. Pellentesque at cursus tellus. Quisque id semper nunc. Quisque viverra a ex at ullamcorper. Morbi mollis erat at ex viverra fringilla. Proin ante dolor, dignissim sodales nisl ac, finibus egestas urna.

Nulla porta urna at pulvinar consectetur. Pellentesque suscipit, neque vitae ultricies rutrum, eros tellus iaculis dui, nec pulvinar justo nibh eu urna. Ut euismod massa nisi, et bibendum risus placerat quis. Integer pretium nulla id risus lobortis laoreet. Aenean quis quam fringilla, elementum odio non, lacinia purus. Vestibulum dui sapien, mollis sit amet massa vel, egestas faucibus velit. Phasellus non justo ut ante vestibulum dictum. Nam in nibh et libero malesuada aliquet. Donec in ex in magna luctus volutpat.

Sed quis dapibus libero. Curabitur id finibus nulla, sed semper felis. Proin dapibus nulla interdum, bibendum tortor et, blandit sapien. Etiam pretium tristique tortor non lacinia. Aliquam dapibus turpis lorem, sit amet porta ex dignissim vitae. In neque felis, sagittis sed ullamcorper lacinia, lobortis ut turpis. Nam quis aliquet justo. Nam eros mi, aliquam vel massa ac, ornare dignissim erat. This is done by using some combination of

$$
\begin{aligned}
a_i &= a_j + a_k \\
a_i &= 2a_j + a_k \\
a_i &= 4a_j + a_k \\
a_i &= 8a_j + a_k \\
a_i &= a_j - a_k \\
a_i &= a_j \ll m\text{shift}
\end{aligned}
$$

instead of the multiplication. For example, you could use:

$$
\begin{aligned}
r &= 4s + s \\
r &= r + r
\end{aligned}
$$

Or by xx:

$$
\begin{aligned}
t &= 2s + s \\
r &= 2t + s \\
r &= 8r + t
\end{aligned}
$$

Cras pharetra ligula nec lectus bibendum, euismod mattis purus cursus. Nullam ut mi molestie purus ultricies lacinia. Phasellus sed orci ac lacus convallis vestibulum. Quisque id nulla ut ipsum finibus vehicula. Curabitur scelerisque erat lobortis, dapibus purus eget, faucibus sapien. Nam enim leo, faucibus id ante sed, fringilla luctus eros. Morbi vulputate, purus at commodo aliquet, turpis dolor sollicitudin libero, id vehicula risus dui sit amet nulla. Sed auctor efficitur urna. Praesent sagittis tellus ac velit vestibulum dignissim. Vivamus justo enim, pellentesque eu posuere id, mattis vitae felis. Aliquam id tincidunt diam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas.

Chapter 3

Optimal Occluders

In this chapter, we consider

3.1 The Mutual Information Model of IF Optimality

Throughout this thesis, a recurring theme is that we'd like to be able to measure how useful an occluder (or other intermediate frame) is to do imaging with. The simplest way to do this is to measure the mutual information between the scene and the observation plane, given a particular transfer matrix (which describes how the intermediate frame transforms the scene before it reaches the observation plane) and signal-to-noise ratio.

3.1.1 The Scene Model

For the time being, we will consider a 2D "flatland" model of the world for simplicity. Extensions to the 3D world are simple, and as previously explained, the 2D model is still representative of the real world in that it accurately describes what happens when you integrate over one the two spatial dimensions in your observation.

3.1.2 The Noise Model

Let's assume that the amount of noise

Appendix A

Tables

Table 1: Armadillos

Armadillos	are
our	friends

Appendix B

Figures

Figure B-1: Armadillo slaying lawyer.

Figure B-2: Armadillo eradicating national debt.