

# On the Theory of Imaging

September 6, 2017

This writeup discusses the relative theoretical merits of different occluders and their limitations.

## 1 General Setup

In this section, we describe the setup of the experiment. For the time being, we posit a one-dimensional scene and a one-dimensional observation plane, with a one-dimensional “intermediate frame” halfway in between. We presume that the observer sees the observation plane and knows the form of the intermediate frame, and must infer the scene. See Fig. 1 for a picture of what the scene’s setup looks like.

Once we’ve specified what is in the intermediate frame, we’ve implied a *transfer matrix*, which we call  $A$ .  $A$  describes how what is in the scene,  $x$ , determines what is on the observation plane,  $y$ . We use  $x$  and  $y$  as vectors representing the amount of intensity at each point in the scene and the observation plane, respectively. In a noiseless setting, we would be able to write:

$$Ax = y$$

There are a few possible different scenarios worth discussing, each one corresponding to a different material in the intermediate frame. For example, the intermediate frame could be a simple occlusion pattern, or some sort of pattern of glass and mirrors, or something else altogether. We will consider each of these cases in turn.

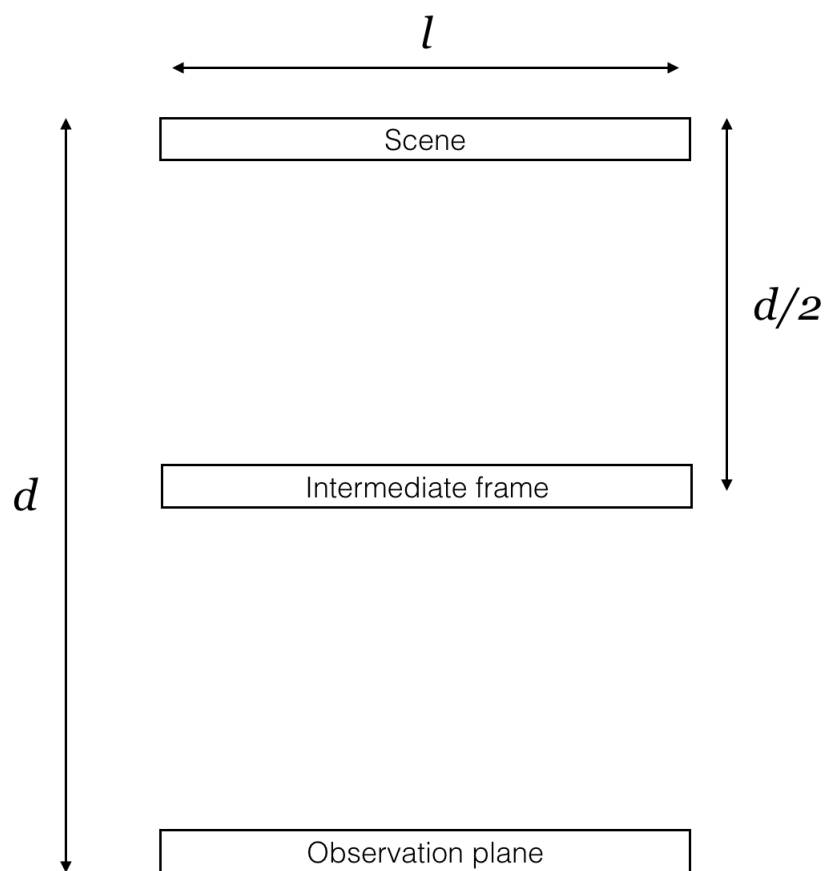


Figure 1: The setup of the scene.

## 2 Occlusion-based patterns

### 2.1 The Setup

If the intermediate frame contains a simple occluder—meaning a pattern of material that is somewhere between completely opaque and completely transparent at each point, that does not change the direction of the incoming light—then that implies a lot of constraints on the form of the transfer matrix.

In order to explore the nature of these constraints, let us begin by discretizing the scene  $x$  and the observation plane  $y$  into  $n$  little patches of size  $l/n$ . Each of the  $x_i$  has some intensity given by the native intensity of the scene; as a result, each of the  $y_i$  gains some intensity from light originating in the scene and reflecting off of the observation plane. For these  $x_i$  and  $y_i$ , we take integer  $i$  with  $0 \leq i < n$ . We presume for the time being that there is no light originating from outside the scene.

We also discretize the intermediate frame into  $2n - 1$  little patches of size  $\frac{l}{2n-1}$ . Each of these patches has some permittivity  $a_i$  between 0 and 1: 0 for completely opaque, and 1 for completely transparent. For these  $a_i$ , we take integer  $i$  with  $|i| < n$ . See Fig. 2 for what the scene setup looks like now.

Note that because we have posited that the intermediate frame contains only a simple occluder, if we want to measure the intensity contribution to a part of the observation plane  $y_j$  from a part of the scene  $x_i$ , we need only consider the straight-line path that the light would take from  $x_i$  to  $y_j$ . This path passes through  $a_{i-j}$  on its way there.

### 2.2 The Transfer Matrix

What fraction  $F_{ij}$  of the light from emitted from  $x_i$ , then, reaches  $y_j$ ? Well, assuming each patch in the scene emits light uniformly across the semi-circle pointed directly downward, and assuming a vertical offset of  $d$  between  $x_i$  and  $y_j$  and a horizontal offset of  $f$ , the fraction of the emitted by  $x_i$  that follows this exact path is given by:

$$F_{ij} = \frac{\Delta x d}{\pi(d^2 + f^2)}. \quad (1)$$

This is the angle subtended by the patch  $d$ . In the equation above,  $\Delta x$  represents the size of the patch  $x_i$ . We can fill in some of these values in terms of the parameters of the setup. In particular,  $\Delta x = \frac{l}{n}$  and  $f = |i+j-n+1|\frac{l}{n}$ . Once we plug in these values, the new expression becomes:

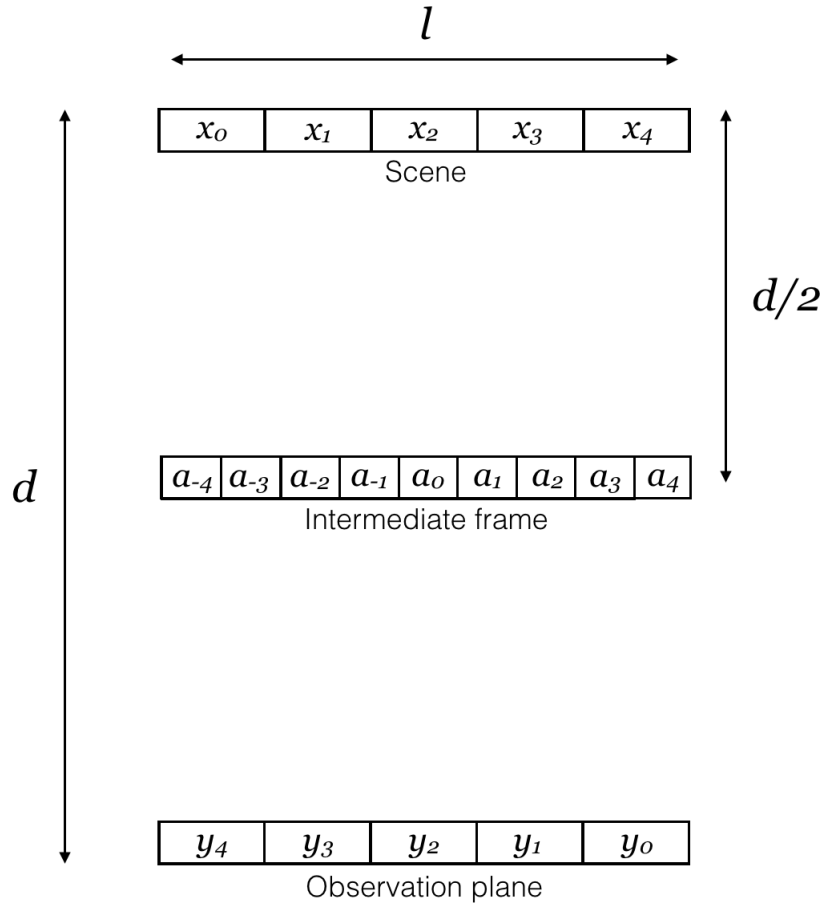


Figure 2: The setup of the scene when the intermediate frame is a simple occluder.

$$F_{ij} = \frac{\frac{l}{n}d}{\pi(d^2 + (i+j-n+1)^2 \frac{l^2}{n^2})} \quad (2)$$

Of course, this is just the fraction of the light that would follow this path. But we know that the intermediate frame blocks some fraction of the light. So the true intensity contribution of a unit of intensity at  $x_i$  to the intensity at  $y_j$  is given by:

$$A_{ij} = a_{i-j} F_{ij}. \quad (3)$$

This is how we can determine the entries of the transfer matrix  $A$ .

First, an interesting observation about this setup: if  $d \gg l$ —that is, if the scene and observation plane are very far away from each other—then in the expression for  $F_{ij}$  shown in Eq. 2, the second term in the denominator,  $(i+j-n+1)^2 \frac{l^2}{n^2} < l^2$ , and so must be much less than  $d^2$ , so we can ignore it entirely. This simplifies the expression for  $F_{ij}$  greatly, and it becomes:

$$F_{ij} = \frac{l}{\pi n d}. \quad (4)$$

The assumption that  $d \gg l$  is used frequently in imaging and is called the *far-field assumption* or *far-field approximation*. In this case, it gives us something quite useful. Note that in the simplified expression for  $F_{ij}$  under the far-field assumption,  $F_{ij}$  is constant in  $i$  and  $j$ . Because  $A_{ij}$  depends only on  $a_{i-j}$  and on  $F_{ij}$ , this tells us that  $A_{ij} = A_{km}$  for any  $i, j, k$ , and  $m$  such that  $i-j = k-m$ . This is the definition of a Toeplitz matrix; as such, we know that  $A$  is Toeplitz in this case.

### 2.3 The Mutual Information Upper Bound

Suppose we have a Gaussian channel, with  $x$  being a vector of random variables whose entries are drawn i.i.d. from a Gaussian distribution with mean 0 and variance 1. Suppose that we have  $y = Ax + \eta$ , where the noise vector  $\eta$  is also drawn i.i.d from a Gaussian distribution, this one with mean 0 and variance  $\sigma^2$ . What is the mutual information between  $x$  and  $y$ ; in other words, how many bits of information do we learn about  $x$  when we learn  $y$  (or vice-versa)?

We know from information theory that in this case, the mutual information  $I(x, y)$  is given by the following expression:

$$I(x, y) = \log(\det(\frac{AA^T}{\sigma^2} + I_n)), \quad (5)$$

where  $I_n$  is the  $n \times n$  identity matrix. This expression is useful for two purposes. The first is that by calculating this expression for various specific  $A$  matrices, we can compare the reconstruction quality of different occluders under varying levels of SNR. The second is that we can use this expression, combined with known bounds on the determinants of matrices, to bound the overall mutual information possible with *any* occluder.

Let's talk about doing that now. Recall from the previous subsection that we have:

$$A_{ij} = a_{i-j} F_{ij} \quad (6)$$

$$0 \leq a_i \leq 1 \quad (7)$$

$$F_{ij} = \frac{\frac{l}{n}d}{\pi(d^2 + (i+j-n+1)^2 \frac{l^2}{n^2})} \quad (8)$$

Using Eq. 8, we can write an upper bound on  $F_{ij}$ :

$$F_{ij} \leq \frac{l}{\pi n d}. \quad (9)$$

This bound we get from the fact that the denominator in the exact expression for  $F_{ij}$  can be no smaller than  $d^2$ . The bound is tight when  $i+j = n-1$ . Importantly, this bound holds *even without the far-field assumption*. Moreover, the expression above shows up so often in this problem that it deserves its own variable. We define:

$$c_f = \frac{l}{\pi n d} \quad (10)$$

Using Eqs. 6 and 7, and the fact that  $F_{ij}$  is positive, we can infer that  $A_{ij} \leq F_{ij}$ . Therefore, we can write:

$$A_{ij} \leq c_f. \quad (11)$$

We know from Hadamard's bound on the determinant that no  $n \times n$  matrix with entries between 0 and 1 can have a determinant whose absolute value exceeds  $\frac{n^{\frac{n}{2}}}{2^{n-1}}$ . And more generally (as a direct extension of the previous bound), no matrix with entries between 0 and  $k$  can have a determinant whose absolute value exceeds  $k^n \frac{n^{\frac{n}{2}}}{2^{n-1}}$ .

We know that  $A$  must have entries between 0 and  $\frac{l}{\pi n d}$ . Therefore,  $AA^T$  has entries between 0 and  $\frac{l^2}{\pi^2 n d^2}$  (which is the result of the dot product between two vectors of length  $n$  and all entries equal to  $\frac{l}{\pi n d}$ ).  $AA^T + I$  must therefore have entries between 0 and  $1 + \frac{l^2}{\pi^2 n d^2}$ .

Using Hadamard's bound of the determinants of matrices, we can write:

$$\det\left(\frac{AA^T}{\sigma^2} + I\right) \leq \left(1 + \frac{l^2}{\pi^2 n d^2 \sigma^2}\right)^n \frac{n^{\frac{n}{2}}}{2^{n-1}} \quad (12)$$

$$I(x, y) = \log \det\left(\frac{AA^T}{\sigma^2} + I\right) \leq n \log\left(1 + \frac{l^2}{\pi^2 n d^2 \sigma^2}\right) + \frac{n}{2} \log(n) - n + 1 \quad (13)$$

## 2.4 The Mutual Information of the Pinhole

We can find the mutual information of a pinhole camera using this same basic approach. Let us suppose that the intermediate frame is composed of a simple pinhole camera (see Figure 3). Then, our expression becomes:

$$\det\left(\frac{AA^T}{\sigma^2} + I\right) = \prod_{i=0}^{n-1} \left(1 + \frac{F_{ii}^2}{\sigma^2}\right) \quad (14)$$

$$I(x, y) = \log\left(\sum_{i=0}^{n-1} \left(1 + \frac{F_{ii}^2}{\sigma^2}\right)\right) \quad (15)$$

where  $F_{ii}$  is defined in Equation 8. In the far-field limit, of course, the mutual information goes as  $\log(1 + \frac{c_f^2}{\sigma^2})^n = n \log(1 + \frac{c_f^2}{\sigma^2})$ .

## 2.5 The Mutual Information of a Lens

What if the intermediate frame is allowed to do more than just block light or let it through? What if it is allowed to redirect the light, like a lens? Then we lose the element-wise restriction on the entries of  $A$  that we formulated in Eq. 11—after all, the intensity contribution of one spot in the scene to another spot on the observation plane could now be much bigger! It could be collecting the light over a wide area and sending it all to the same spot.

We still can, however, come up with a restriction on the entries of  $A$ , but a looser one. In particular, we know that the total amount of light hitting the observation plane can't be more than the total amount of light hitting the glass—glass can only bend light, not create it! This corresponds to a constraint that  $A$  be left stochastic matrix, scaled by a factor of the amount of light hitting the lens.

More precisely, suppose there is a point in the scene with intensity  $I$ . Its *total* intensity contribution to the entire observation plane can't possibly exceed  $I$  times the fraction of its emitted light the lens—in the far-field limit, that's

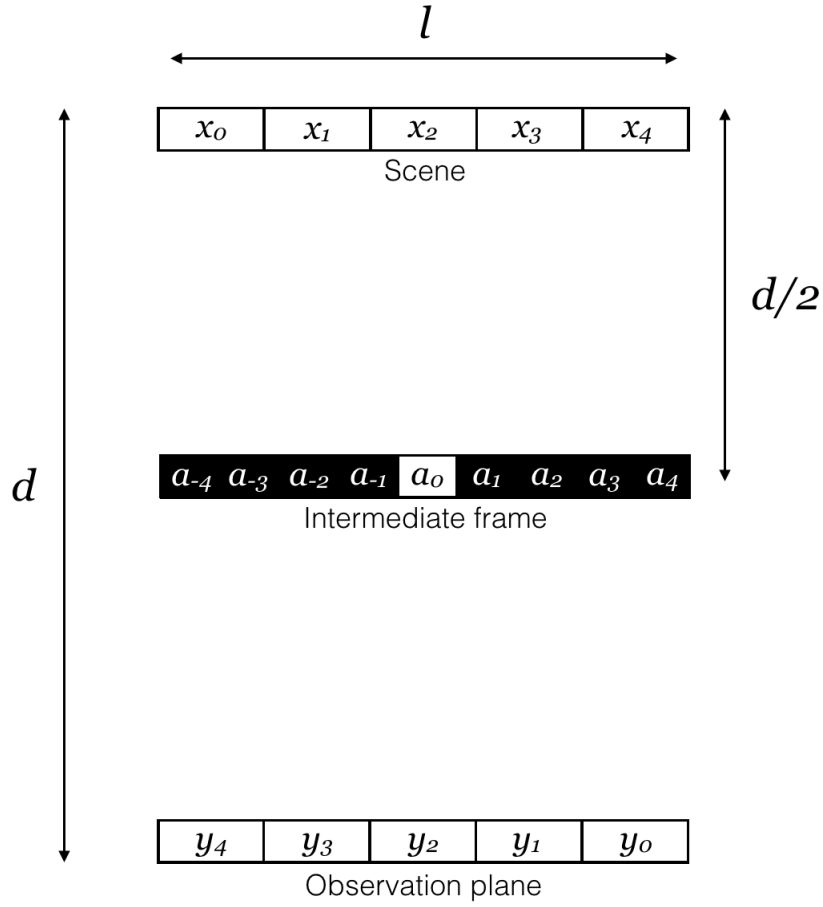


Figure 3: The setup of the scene when the intermediate frame is a simple occluder.



$$c_l = \frac{2l}{\pi d}$$

The determinant of any stochastic matrix is at most 1, with the bound being tight for the identity matrix. Thus, in the realm where we want to maximize the mutual information, and we are subject to the constraint that our transfer matrix  $A$  has its columns sum to at most  $c_l$ , we have:

$$I(x, y) = \log \det \left( \frac{AA^T}{\sigma^2} + I \right) = \log \det \frac{(c_l^2 + 1)}{\sigma^2} I = \log (c_l^2 + 1)^n = n \log (c_l^2 + 1) \quad (16)$$