Beyond Diffraction Limited Seeing Through Polarization Diversity

Steven P. James
Air Force Institute of Technology, Student
Dr. Stephen C. Cain
Air Force Institute of Technology, Professor

1. Abstract

The purpose of the algorithm developed in this thesis was to create a post processing method that could resolve objects at low signal levels using polarization diversity and no knowledge of the atmospheric seeing conditions. The process uses a two-channel system, one unpolarized image and one linearly polarized image, in a GEM algorithm to reconstruct the object. Previous work done by Strong showed that a two-channel system using polarization diversity on short exposure imagery could produce images up to twice the diffraction limit. In this research, long exposure images were simulated and a simple Kolmogorov model used. This allowed for the atmosphere to be characterized by single parameter, the Fried Parameter. Introducing a novel polarization prior that restricts the polarization parameter, it was possible to determine the Fried Parameter to within half a centimeter without any addition knowledge or processes. It was also found that when high polarization diversity was present in the image could be reconstructed with significantly better resolution and signal level did not affect this resolving capability. At very low signal levels, imagery with low to no diversity could not be resolved at all whereas high diversity resolved equally as well as if there was a high signal level. Current algorithms being used do not include polarization diversity but can substantially improve resolution. Application of this algorithm could be used in dim-object detection around satellites as well as solar surface imagery.

2. Introduction

In 2001 the Department of Defense released a comprehensive report on the United States Space Capabilities. In that report, it was said that we are ripe for a "Space Pearl Harbor." [1] Since then, there has been a concerted effort to mitigate this possibly with the advancement of Space Superiority. This is broken down into three categories: Offensive Counterspace, Defensive Counterspace, and Space Situational Awareness (SSA). In orbit around the earth it is very difficult to identify and characterize anomalies that may occur with "blue" spacecraft or the functions and purpose of "red" spacecraft. That is where SSA comes in. It is the attempt to have complete awareness of the battlespace in orbit.

Advanced sensors designed to inspect the orbital battlespace or ground-based telescope systems are required. The design and launch of satellites are very costly, especially at geosynchronous (GEO) orbit. At Geo, there is so much distance between satellites that space-based optical systems need to be maneuvered close to each Resident Space Object (RSO) of interest. This greatly limits lifetime due to finite fuel, and also restricts the response time kill chain after an event occurs. On the other hand, ground based optical systems have a comparably low cost, can be easily repaired or upgraded, and can respond quickly when an event occurs. The drawback is that observing must take place anywhere from hundreds of kilometers, for low earth, to thousands of kilometers for GEO. On top of that, resolution is reduced considerable by atmospheric seeing conditions. If large enough telescopes or telescope arrays are constructed, atmospheric distortion is the main thing that needs to be mitigated. Adaptive Optics can help significantly with this but do not correct for all atmospheric distortion. Post processing algorithms can be used to further reconstruct the RSO. Combinations of Adaptive Optics and post processing are currently being used operationally to characterize satellites and anomalies in space.

3. Previous Work

The worked developed in this thesis is built primarily from the research done by Major David Strong [2] and Lieutenant Colonel Adam McDonald [3]. Strong's dissertation created a two-channel, one unpolarized and one linearly polarized, blind deconvolution algorithm for passively illuminated objects as is done in this thesis but with several key differences. His algorithm was created for use in short exposure images as opposed to long exposure. Details of the object will not be blurred out as much from averaging over a longer period of time. The drawback to the short exposure case is that signal levels are significantly reduced and therefore the SNR is much lower. The other downside to this is that the point spread function (psf) of the atmosphere is not as well known. As an image is integrated, the psf will tend toward a well known and easily modeled transfer function, such as the Kolmogorov spectrum. In contrast, if the integration time is small, fluctuations in the atmosphere can cause the psf to vary greatly. This makes it difficult to estimate and characterize [2].

Another major difference is in the development of the two-channel algorithm derivation. In this thesis, a prior density for the polarization term is included to restrict the possible values that the polarization parameter can take. This allows for the polarization state and the object to be estimated simultaneously. The addition of the prior gives significantly increased information when polarization diversity is present and when calculating the correct seeing parameter of the atmosphere [2].

In MacDonald's work, he developed a method for estimating the seeing parameter of the atmosphere (r_0) for a single channel system of laser light, conforming to a negative binomial distribution. The method involved representing the r_0 probability density function in some distribution. In that case, based on the observation that good seeing (high r_0) is much less likely than bad seeing (low r_0), an exponentially decreasing function was chosen

$$f(r_0) = \frac{N^2}{r_{avg}} e^{-N^2 \frac{r_0}{r_{avg}}} \tag{1}$$

where N is the number of pixels on a side and r_{avg} is some parameter that must be iteratively calculated. Once the average seeing parameter is found it produces a graph similar to that in Fig. 5. When equation (1) is left out when deriving the likelihood, there is no decrease around the actual r_0 value and the likelihood continues to increase forever [3].

It was assumed that MacDonald's method would be needed to find the correct seeing parameter for the algorithm described in this document. Surprisingly, this was not the case. By using the polarization diversity algorithm with a polarization prior, the likelihood curve naturally produces a maximum near the actual r_0 value. The curve looks very similar to what is produced by Adam MacDonald's iterative method.

4. Development

The experiment design model for the development of this algorithm is that of a 2 channel imaging system. From one aperture, light is split into two different CCD arrays with one channel unimpeded and the other with a linear polarizer in from of it. The orientation of the polarizer is arbitrary and will not affect the algorithm. From this point forward, the polarization parameter, P, will be defined as the ratio of the polarized intensity over the unpolarized intensity of a given pixel in the images. The CCD array photocounts are mathematically described by standard Poisson distribution given by,

$$G = \prod_{y_1}^{Y_1} \left[\left(\frac{i_{UP}(y_1)^{\tilde{d}_{UP}(y_1)} e^{-i_{UP}(y_1)}}{\tilde{d}_{UP}(y_1)!} \right) \right] \prod_{y_2}^{Y_2} \left[\left(\frac{i_P(y_2)^{\tilde{d}_P(y_2)} e^{-i_P(y_2)}}{\tilde{d}_P(y_2)!} \right) \right] \tag{2}$$

where y_1 and y_2 are general coordinates of the image planes, Y_1 and Y_2 are the total number of pixels, and $\tilde{d}(y)$ is the collected photon counts from the detectors. This data will be referred to as the incomplete data. The subscript UP denotes the unpolarized channel and the subscript P denotes the polarized channel. $i_{UP}(y)$ and $i_P(y)$ are the average intensity of the images given by the convolutions

$$i_{UP}(y_1) = \sum_x^X O(x)h(y_1 - x) \tag{3}$$

$$i_P(y_2) = \sum_x^X O(x)P(x)h(y_2 - x) \tag{4}$$

$O(x)$ is the object, $h(y-x)$ is the point spread function, $P(x)$ is the set of polarization states with values between 0 and 1, and *x* is a general coordinate for the object plane [4].

Taking the log of equation (2) gives and removing constant terms gives the incomplete data log-likelihood

$$\begin{aligned} L_{ID} = \sum_y^Y \{\tilde{d}_{UP}(y) \ln\left(\sum_x^X O(x)h(y-x)\right) - \sum_x^X O(x)h(y-x) + \\ \tilde{d}_P(y) \ln\left(\sum_x^X O(x)P(x)h(y-x)\right) - \sum_x^X O(x)P(x)h(y-x)\} \end{aligned} \tag{5}$$

In order to obtain the desired information of the object and polarization state, the complete data model must be used. This model and data relationships are described in Shultz [4]. Using those definitions and applying them to equation (5) results in the complete data log-likelihood is

$$\begin{aligned} L_{CD} = \sum_y^Y \sum_x^X \{d_{UP}(y) \ln\big(O(x)h(y-x)\big) - O(x)h(y-x) + \\ d_P(y) \ln\big(O(x)P(x)h(y-x)\big) - O(x)P(x)h(y-x)\} \end{aligned} \tag{6}$$

Here, $d_P(y)$ and $d_{UP}(y)$ are the complete data for the polarized and unpolarized channel respectively.

The complete data log-likelihood has now been defined however, because the parameter P is Poisson, there is nothing restricting the algorithm from choosing values greater than 1, which are non-physical solutions. In order to mitigate the possibility of this a constraining term is added for P. Ideally the constraint should be a step probability function that goes from 1 to 0 at a P value of 1. This discontinuity would create problems when taking the derivative so the alternative used is a "super-Gaussian,"

$$f(P(x)) = Ae^{-P(x)^n}, P(x) \geq 0 \tag{7}$$

where n is a positive even integer. The greater n is, the better a step function is modeled. Inserting this into the complete data log likelihood results in a final equation of

$$L_{CD} = \sum_{y}^{Y}\sum_{x}^{X}\{d_{UP}(y)\ln\big(O(x)h(y-x)\big) - O(x)h(y-x) + d_P(y)\ln\big(O(x)P(x)h(y-x)\big) - O(x)P(x)h(y-x) - P(x)^n\} \tag{8}$$

Now, in order to solve the GEM (General Expectation Maximization) algorithm, the first step is to take the conditional expectation of the complete data log-likelihood, which is called the Q function. The only random variable in equation (8) is the complete data so after some manipulation the equation becomes

$$Q = \sum_{y}^{Y}\sum_{x}^{X}\{E[d_{UP}(y)\,]\ln\big(O(x)h(y-x)\big) - O(x)h(y-x) + E[d_P(y)\,]\ln\big(O(x)P(x)h(y-x)\big) - O(x)P(x)h(y-x) - P(x)^n\} \tag{9}$$

Yet again, taking definitions for Shultz [4] the conditional expectations in equation (9) are,

$$E[d_{UP}(y)] = \frac{h(y-x)O(x)}{i_{UP}(y|O,h)}\tilde{d}_{UP}(y) \tag{10}$$

$$E[d_P(y)] = \frac{h(y-x)P(x)O(x)}{i_P(y|O,P,h)}\tilde{d}_P(y) \tag{11}$$

The final step is to maximize equation (9) for both parameters O(x) and P(x). This is done simply by taking the derivative and setting equal to zero. The solution for these are,

$$O(x_0) = \frac{\sum_y^Y E[d_{UP}(y)] + \sum_y^Y E[d_P(y)]}{1 + P(x_0)} \tag{12}$$

$$0 = nP(x)^{n+1} + nP(x)^n + P(x)\sum_{y}^{Y}E[d_{UP}(y)] - \sum_{y}^{Y}E[d_P(y)] \tag{13}$$

In the case of equation (13), $P(x)$ is the real root of the polynomial expression given. The total number of roots to the equation is $n+1$, which is reason not to choose an n that is too large. $P(x)$ only having one real root has not been proven but was verified in all simulation tests.

All of the pieces of the algorithm are in place and for visual clarity the algorithm flow is shown in Fig. 1.

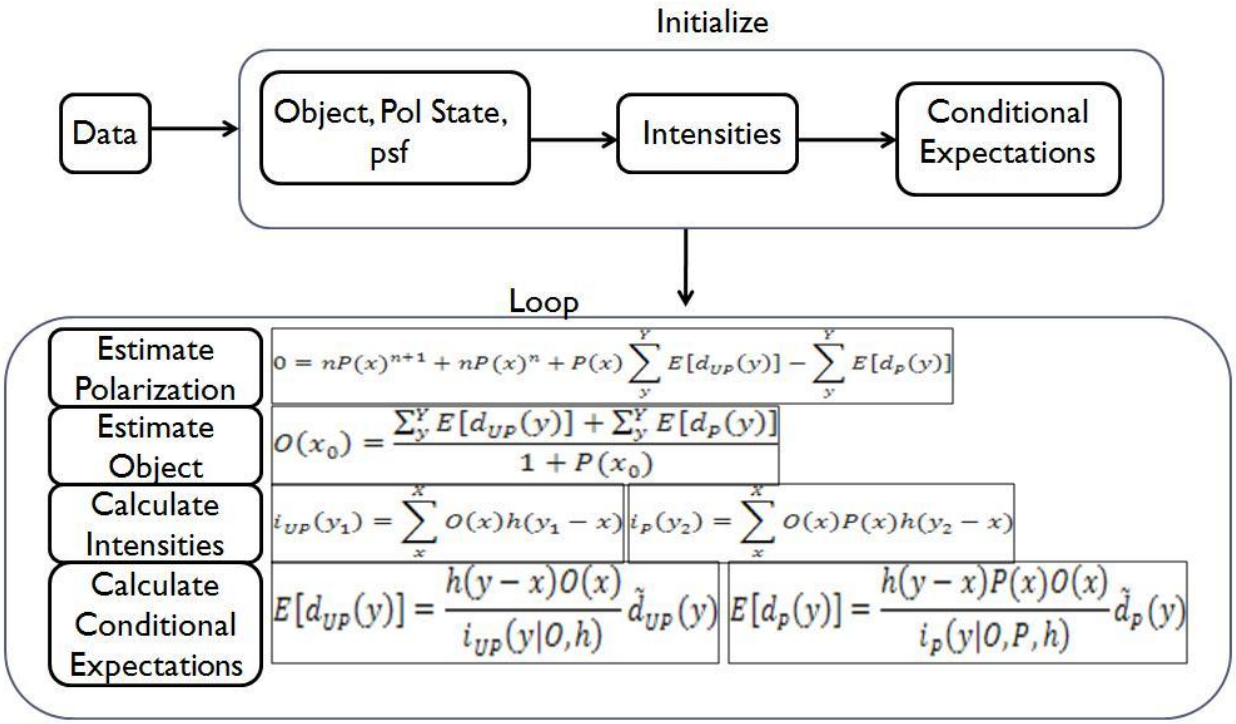

Fig. 1: This flow chart shows the equations and process that the algorithm uses to enhance image resolution

5. Simulation Testing

To test the algorithm, a simulation was designed to measure how well two distinct points could be resolved. This was done by creating a polarized an unpolarized image with two separated bar targets, Fig. 2. Each of the two bars was given a different polarization state for the polarized image. Using this simulated data, the algorithm attempts to reconstruct the original object. It should be noted that for this test, only the correct r_0 value is used; no r_0 estimation is done. Since the object is symmetric about the y-axis and to decrease the data size, a cross section is taken through the corrected image, Fig. 3. The goal is not to reconstruct the object perfectly but rather to simply discern that there are two different bars. To characterize how well the algorithm preformed, the ratio of the minimum between bar targets and the peak is used. This will always be a number between 0, perfect distinction, and 1, no distinction. Given this characterization ratio, the signal strength and polarization state of the bar targets were repeated changed to see how they would affect the reconstruction. Fig. 4 shows colored contour plots of these results. Each plot is a different signal strength and the axes of the plots are the polarization given each bar target. For clarification, each plot shows ratio results from multiple runs of the algorithm. As would be expected, there is symmetry in the plots since swapping the polarization on the bars should not affect the reconstruction. At the edges of the plot, where polarization diversity is high, there is noticeably better resolution. However as one approaches the center diagonal where there is little diversity, resolution significantly decreases. This shows that including the second polarization channel and incorporating that information does increase the ability to reconstruct higher spatial frequencies. An unanticipated effect of the two-channel algorithm is that the reconstruction is relatively invariant to signal strength. In comparing the plots in Fig. 4, as the signal decreases from 50000 to 500 photons the resolving capability along the low diversity diagonal gets progressively worse, from about 0.7 up to 1.0. However at the edges, the ratio stays consistent at approximately 0.3.

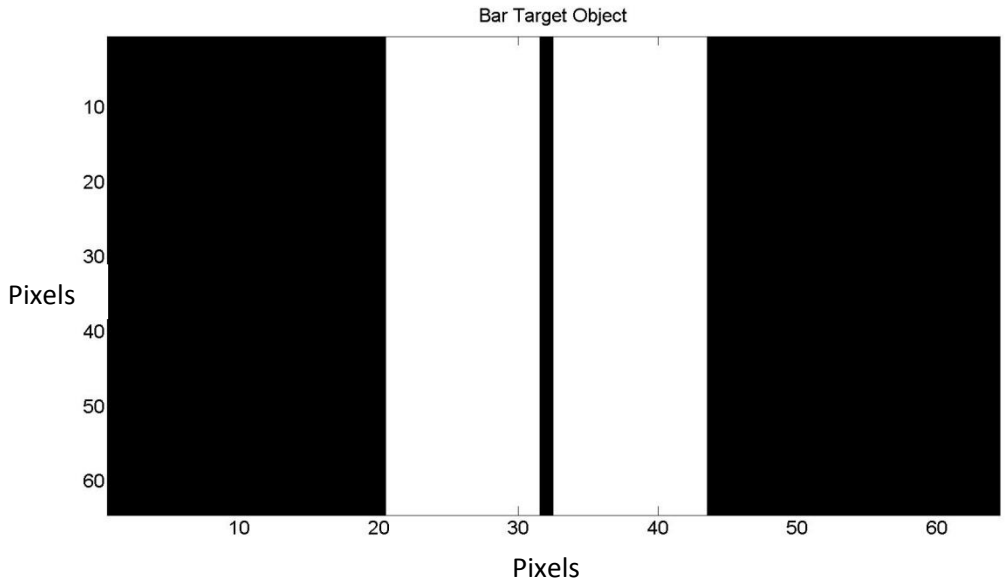

Fig. 2: Two bars to be propagated through a simulated atmosphere and telescope aperture. Each bar is given a different polarization parameter.

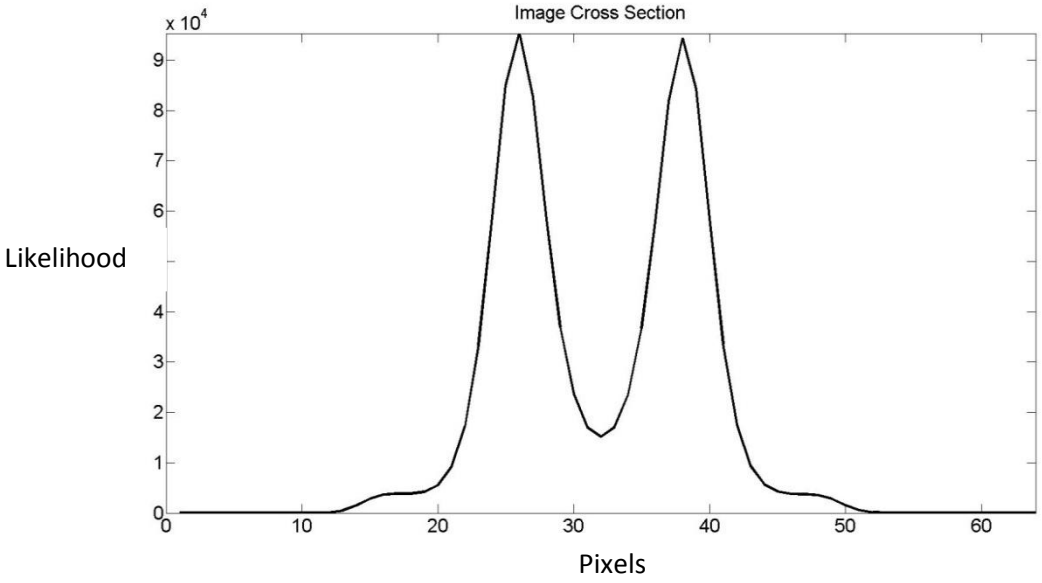

Fig. 3: An example of the reconstructed bar target. In this case, the algorithm was easily able to distinguish two different objects.

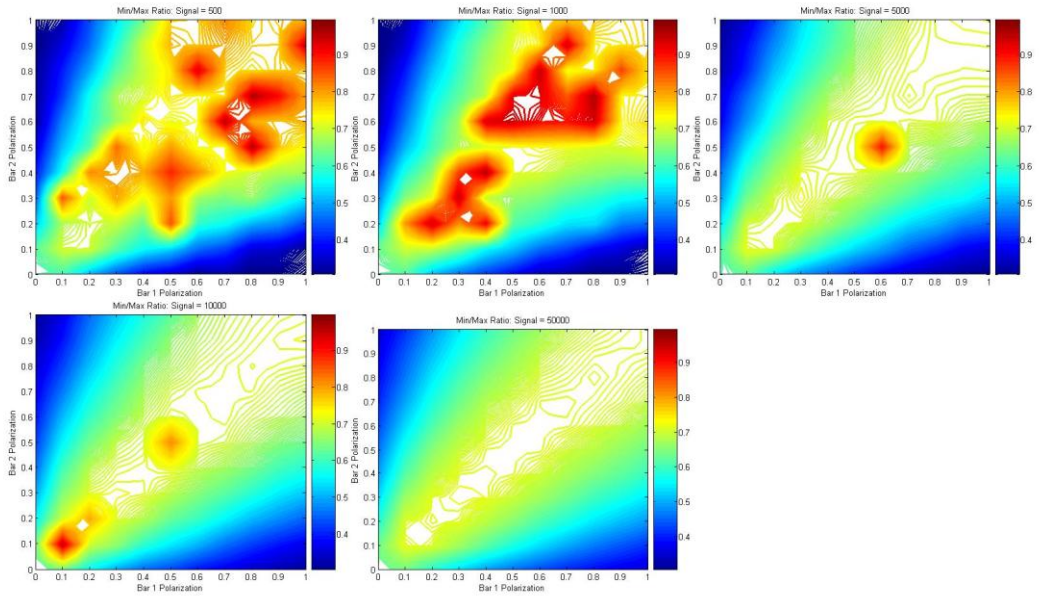

Fig. 4: Contour plot of resolving capability based on bar polarizations. Signal levels shown are 500 (top left), 1000 (top center), 5000 (top right), 10000 (bottom left), 50000 (bottom right)

The second test was focused on estimating r_0 with no prior knowledge other than the two-channel image data. Simulated data is created using the Komologov long exposure OTF with a given r_0 value. Incrementing through a set of r_0 values, the algorithm is run repeatedly. After each iteration of the algorithm, when it reaches the stopping criteria, the likelihood was recorded. Once all r_0 values had been incremented through, the likelihood was plotted as function of r_0, Fig. 5. This Fig. is similar to all of the likelihood curves obtained from testing in that it rises very steeply below the actual r_0 value and then slowly decrease after the optimal solution has been passed. In this case an r_0 value of 6cm was used and likelihood curve was able to pick that out.

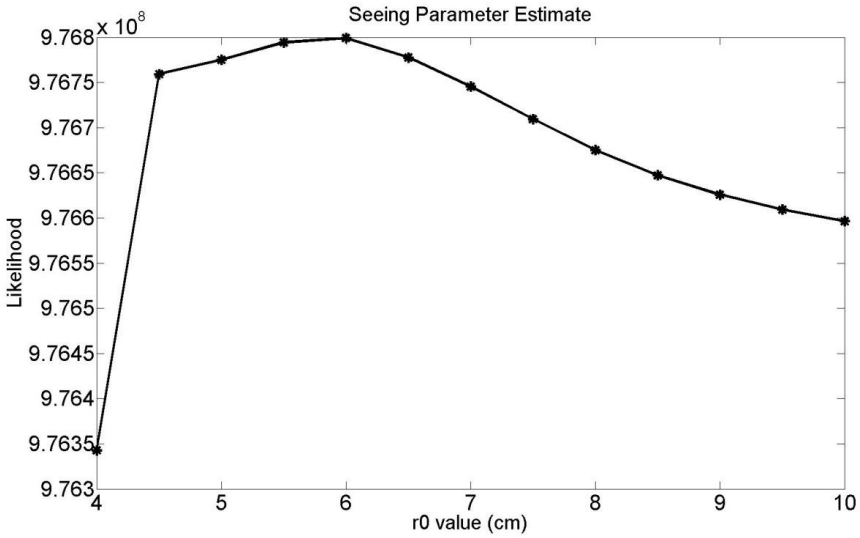

Fig. 5: After the algorithm produces an estimated image for a range of r_0 values, the likelihood is calculated and plotted. It reaches maximum at or near the true seeing parameter. In this case, the r_0 value was 6cm and was estimated exactly.

6. Conclusions

The purpose of the algorithm developed in this thesis was to produce improved object estimates beyond that of single channel models currently employed. The process is dependent on not just polarization but polarization diversity. When two points in the image have significantly different polarizations, one horizontal and one vertical for instance, the ability of the algorithm to resolve them is greatly increased. In contrast to this, when there is little to no diversity the algorithm will only resolve as well as a single channel system would.

The first unexpected consequence of using polarization diversity is that it is invariant to signal level. When there is high diversity, low signal levels can be resolved just as well as much higher ones. This continues down to SNR's below 10. However, as the signal decreases, points with low polarization diversity cannot be distinguished as well or at all. The required level of diversity increases as the signal fades. The points must be increasingly polarized to achieve the same resolution but, at the extremes points, they are always resolved to the same high level.

The other goal of the algorithm was to determine the Fried's parameter for the atmosphere without having any knowledge of the seeing conditions. The work done by MacDonald [3] required a secondary method to calculate the r_0 value from the likelihood curve. Without a separate estimation of the r_0 value, the likelihood curve rises steeply and levels off but never reaches an apex. In the case of polarization diversity, this additional calculation is not needed. The curve does reach a maximum and then begins to decrease. With the additional information provided by a two channel system, the correct seeing parameter naturally falls out of the likelihood.

7. References

1. Committee on Armed Services of the United States Senate and House of Representatives. *Report of the Commission to Asses United States National Security Space Management and Organization*. Pursuant to Public Law 106-65: January 11, 2001.
2. Stong, David M. *Polarimeter Blind Deconvolution Using Image Diversity*. Air Force Institute of Technology (AU), Wright-Patterson AFB OH, September 2007.
3. MacDonald, Adam. *Blind Deconvolution of Anisoplanatic Images Collected by a Partially Coherent Imaging System*. Institute of Technology (AU), Wright-Patterson AFB OH, June 2006.
4. Schultz, Timothy J. "Multiframe blind deconvolution of astronomical images," *J. Optical Society of America A*, Vol. 10, No. 5, 1064-1073 (May 1993).