Initial Information Content Analysis of Shadow Imaging of GEO Satellites

David G. Sheppard[a*], Bobby R. Hunt[a], and Peter N. McMahon-Crabtree[b]

[a]KBR, 533 Lipoa Parkway, Kihei, HI, 96753

[b]Air Force Research Laboratory, 3550 Aberdeen Ave. SE, Kirtland AFB, NM 87117

Abstract: Space Domain Awareness (SDA) is important for understanding the space environment to ensure safe operation of space missions. SDA activities include the detection, identification, tracking and characterizing of artificial satellites. SDA objectives rely greatly on the information that can be gained from ground-based sensors, such as optical telescopes. A space object may be detected by a telescope when it passes in front of a star. There have been studies of how the shadow cast to Earth from a star can be interpreted for important data, referred to as Shadow Imaging. Herein we discuss the usage of information theoretic methods to understand the limitations of stellar Shadow Imaging. These methods measure the information content in the irradiance pattern, as seen by a terrestrial observer, from the shadow cast by a geosynchronous space object passing in front of a stellar source.

Acknowledgements: The Authors thank the US Air Force Office of Scientific Research for sponsoring the efforts reported herein, under LRIR project number 22RVCOR005.

1 Introduction: Shadow Imaging

The capabilities to launch and operate satellites in orbit around the Earth by both nation-states and the private sector have significantly expanded in the past decade. The number of space systems in operation, and the pace of expanding space assets, continue to accelerate. The field of Space Domain Awareness (SDA), formerly referred to as Space Situational Awareness (SSA), is focused on the problem of understanding the space environment and potential issues such as debris or potential conjunctions. SDA encompasses the effective identification, characterization and understanding of any factor associated with the space domain that could affect space operations and thereby impact the security, safety, economy, or environment of the United States of America [1].

Traditionally, SDA has been enabled by globally distributed optical and radar systems. For example, telescopes have been deployed around the world to detect and observe space object behavior and characteristics. However, many objects are in orbits which are too high for any but the largest telescopes to produce a resolved or partially resolved image. For example, a telescope with a primary mirror of 2.5 m diameter, imaging in the center of the visible band gives diffraction limited angular resolution of details at approximately 22 cm for an object at orbital altitude of 1000 km. Atmospheric effects (such as atmospheric refraction, dispersion, line-of-sight turbulence, water vapor, ...) further impact the resolvable details. Some of these effects can be overcome by post-processing techniques, but low signal-to-noise ratio (SNR) when imaging dim objects presents a difficult challenge.

For observation of dim objects, or where traditional imaging is not feasible, the possibility has emerged of observing the shadow produced by obstruction of light from a distant source by a space object. This phenomenon is known as occultation. The observation of a shadow produced by occultation using a single collection aperture was first proposed in 1952 by Taylor [2]. A shadow of the asteroid Pallas was first successfully observed in 1978 [3]. Occultation studies have since begun to appear more frequently in astronomical journals. The search for objects and/or characterization of the Kuiper Belt, i.e., objects beyond Neptune, by occultation of stars is a notable area of application [4] [5].

The developments in astronomy are complemented by the identification of shadow imaging as a potential new tool for SDA, first proposed in 2005 [6]. Shadow imaging seeks to exploit the shadow cast by a space object when it obstructs the line of sight between an observer and a stellar source (or perhaps an object in our solar system). The shadow irradiance pattern produced at the observer contains information about the object's silhouette which is degraded mainly by free space scalar diffraction effects and detector noise. The main proposition of shadow imaging is that this pattern can be measured well enough to support solving the inverse problem of recovering the silhouette image. In this paper, we discuss the information content of the irradiance patterns and how this relates to

*corresponding author: david.sheppard@us.kbr.com

understanding the limits of performance of a shadow imaging system. This information content is approached in a general way as measured in the irradiance pattern, treated as an image, as well as how it supports a specific exploitation task, i.e., the localization of a knife edge.

1.1 Occultation by Orbiting Satellites and Shadow Imaging for SDA

The nature of the occulted light source is of central importance to understanding the measured data. Stellar sources differ in such details as size and spectral content. Many stellar sources can be treated as unresolved point sources, but there are stellar objects that are large enough or close enough to have significant angular extent. In the former case, the electromagnetic waves can be treated as plane waves. For the latter, the light can be modeled as a radially distributed set of plane waves. The orbital regime of the satellite is also of importance. Low Earth Orbit (LEO), Medium Earth Orbit, (MEO), and Geosynchronous Earth Orbit (GEO), all differ in orbital velocity and the specific track geometry across the face of the Earth for any shadow cast by occultation.

In [6], the stellar point source illumination was assumed to be an incoherent plane wave arriving at a GEO object, with the shadow occurring past the object formed by diffraction. The inverse problem of recovering the object silhouette was dealt with using a version of the Gerchberg-Saxton algorithm developed for phase retrieval [7]. For observed diffraction phenomena consistent with the Fresnel region the intensity ringing around the shadow's extent can be observed and exploited in the reconstruction process. The size and range from the Earth of GEO satellites results in shadows within the Fresnel region for terrestrial based observations using visible wavelengths.

Shadow imaging of GEO satellites was examined in further detail in 2008 [8], introducing the concept of spectrally resolved shadow imaging which sought to increase the amount of attainable spatial resolution assuming sufficiently high SNR. This was achieved by splitting the collected light into spectral bins and performing the image reconstruction separately on each bin. The final image results from stacking the individual reconstructed silhouette images produced from each spectral bin.

Several system-implementation issues have been considered in assessing the viability of collecting and using shadows of orbiting satellites cast from point sources. These initial analyses of the various challenges of shadow imaging indicated that shadow imaging from a star may be a viable way to create a representation of the silhouette of a space object [6] [9] [10].

Work by Douglas [9] examined the image resolution limits of shadow imaging using a radiometrically based wave optics simulation approach by varying environmental, observational, and light collection parameters. Recent work has concentrated on examining various reconstruction algorithms based on the Fresnel integral approach as well as predicting when and where GEO shadow events will occur on the Earth [10].

It is important emphasize that successful inversion of the diffraction pattern from a stellar occultation yields an estimate of the object silhouette, not a gray-scale image of the satellite. A silhouette lacks the interior detail that would be present in a gray-scale image, and is only the outline of an object, but can reveal details of interest for SDA. Object size, major components and their configurations, articulations of such components over time, and the object orientation with respect to the observer are all examples of relevant details potentially derived from a silhouette image of sufficient spatial resolution.

A conceptual Earth-based shadow imaging system like that proposed in [6] is depicted in Figure 1, in which a linear array of apertures arranged in a north-south direction on the Earth is used to collect light as a two-dimensional shadow moves across. For a space object at GEO distance, the sidereal motion of the Earth causes its shadow to move from west to east at approximately 2.6 km/s near the equator.

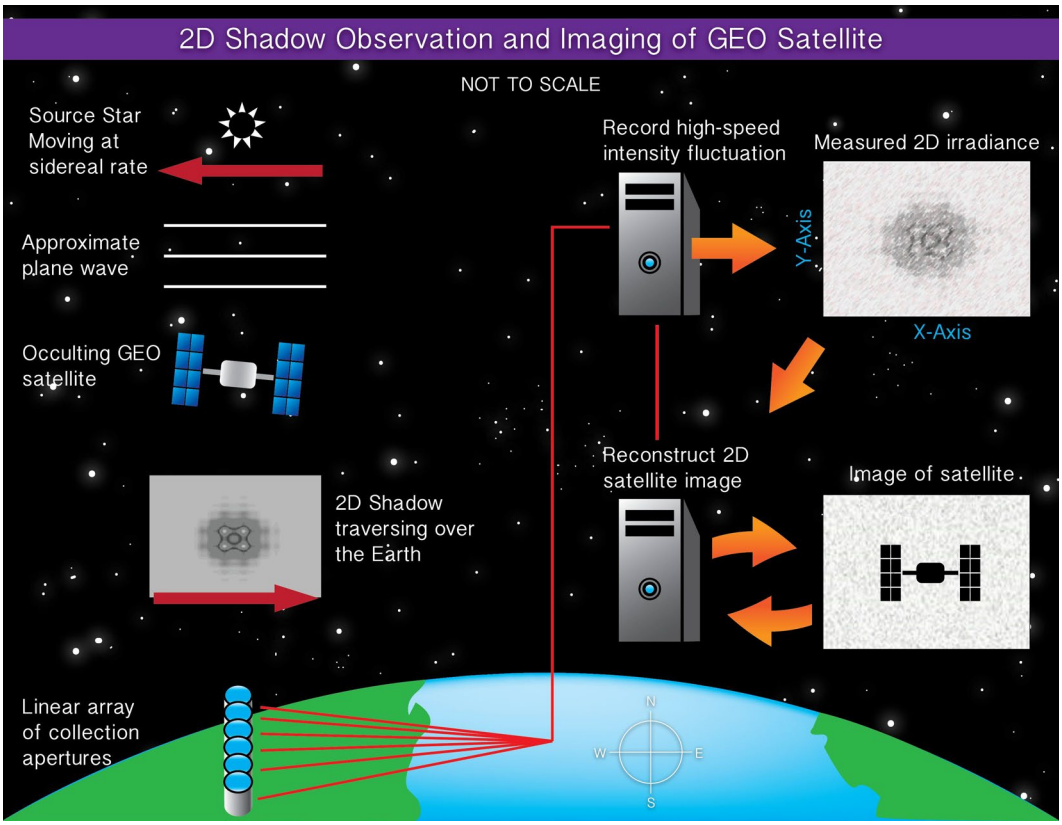

Figure 1. Two-dimensional shadow observation (imaging) of geosynchronous satellite. Note that the depiction is not to scale.

1.2 Motivation

Ultimately, we hope to show that information theory is a helpful concept in understanding the performance limits of shadow imaging systems, as well as how different system parameters come into play. Wavelength dependence can be added with obvious modifications to the above discussion. This is approached in two ways: (1) by computing information in the ground irradiance patterns as a whole; and (2) by computing information related to a particular task, i.e., knife edge localization.

2 RESOLUTION LIMITS

The fundamental resolution limits for shadow imaging were analytically derived as a function of spectral bin width and collection aperture size D_{ap} in [8] [11], and studied statistically in [12]. This section summarizes the main findings from these sources for convenience.

2.1 Analytic Studies of Shadow Imaging Resolution

In prior work, the resolution of shadow imaging was studied using analytical approaches [8] [11]. The results of these efforts, which used an equivalent Fresnel number approach, are summarized here for convenience.

- Fine features of the collected polychromatic diffraction pattern are blurred resulting in a loss of resolution in the final reconstructed image as the spectral bin width bins increases.
- The resolution limit in terms of resolvable object half-width size is given by

$$d_{lim} = 2 * \left(\sqrt{\lambda_2 z} - \sqrt{\lambda_1 z}\right) \tag{1}$$

where λ_1 and λ_2 are the wavelengths of the beginning and end of the spectral bin, respectively, and z is the propagation distance [11].

- The resolution limit as a function of spectral binning and collection aperture size is given by (augmenting the previous equation by including the factor derived in [8])

$$d_{lim} = 2 * \left(\sqrt{\lambda_2 z} - \sqrt{\lambda_1 z}\right) + 0.61 D_{ap} \quad (2)$$

- The propagation distance also impacts the resolution limit based on spectral bin width, but a considerable difference in propagation distance from the nominal GEO range of 36,000 km is required to influence the resolution limits.
- The classical resolution limit d_{lim}^{c} given by the Rayleigh criteria is given by

$$d_{lim}^{c} = 1.22 \frac{\lambda z}{D_{ap}} \quad (3)$$

By comparing these equations, the differences in how aperture size and wavelength affect the resolution limits of classical imaging and shadow imaging becomes clear. These differences can be summarized as the following [11]:

- The resolution limit becomes finer as D_{ap} increases in classical imaging, while in shadow imaging smaller apertures yield better resolution. This assumes there is a sufficient signal-to-noise ratio (SNR).
- Better resolution is attained when using shorter wavelengths in classical imaging. In shadow imaging, the spectral bin width determines the resolution limit, and the resolution is less sensitive to the starting wavelength of a given spectral bin.
- The classical resolution limit is linearly dependent on the object distance. In shadow imaging the resolution limit is proportional to the square root of the distance to the object.

2.2 Empirical Studies of Shadow Imaging Resolution

Empirical studies also exist in which simulation was used to understand shadow imaging system performance. Notably, the dissertation by Douglas [9] examines many system parameters to reveal their influences. In our work, we focus on some of the empirical studies in which visual estimates were made and relate those to information content.

In [12], the knife edge localization problem was used as a surrogate for understanding the resolution of a shadow imaging system. The root-mean-squared-error (RMSE) was computed for monochromatic and polychromatic cases and related to statistical bounds, including the Cramér-Rao and Hammersly-Chapman-Robbins bounds. Large-scale Monte Carlo methods were used to show the bounds for a set of assumed system parameters. In our work, we attempt to parallel their work in terms of task specific information, described below.

3 Shadow Imaging and Shannon Information

Information theory as applied to image processing has a long history as a means of understanding both the information content of single images, motion imagery sequences, and the optical channel itself. The motivation here is to use information theoretic techniques to understand the limits of a shadow imaging system performance, i.e., the ability to quantify it in terms of the information content in the ground irradiance patterns. Information content in the irradiance pattern is of interest, as is the information content as it relates to a specific task to be performed with the measured pattern.

3.1 Information and the Optical Channel

One of the earliest contributions to information theory as it relates to optical systems is that found in the classic paper by Fellget and Linfoot [13]. The main contribution in that paper was to quantify information transmission

through an optical channel in a very general way, and to express it in terms of power spectra of the signal and noise. This is useful in quantifying information content in the shadow ground irradiance pattern, and to understand the losses due to scalar diffraction (wave optics propagation) incurred between the aperture and the ground plane.

3.1.1 Overview

Shannon channel capacity is reviewed here as applied to the optical channel and the resulting shadow irradiance pattern. This will be used later to re-examine some of the findings from Douglas [9] and McNicholl [12]. This will be applied later to empirical studies relying on simulated shadow imaging.

3.1.2 Information Content Model for Image Channel

The shadow imaging system consists of a stellar source, a satellite situated in the aperture plane, and the ground (sensor) plane located where the shadow falls on the Earth. The satellite itself is assumed to be non-transmissive. From [13], the monochromatic mean information content observed in the ground plane irradiance pattern can be expressed as

$$\log N = \frac{1}{2}\iint \log_2\left(1 + \frac{\overline{|G(u,v)|^2}}{|W(u,v)|^2}\right) dudv \text{ (bits)} \tag{4}$$

where

- f is the satellite aperture function,
- g is the pristine, noise free, measured ground irradiance pattern,
- w is the noise in the measured diffraction pattern,
- $\overline{|G(u,v)|^2}$ is the power spectrum (periodogram) of g,
- $|W(u,v)|^2$ is the power spectrum of w.

Noise originating from the object is assumed to be small compared to the measurement noise, and the log function is base 2 so that the information is in units of "bits." The polychromatic mean information content is given by

$$\log N = \frac{1}{2|B|}\iiint \log\left(1 + \frac{\overline{|G(u,v,\lambda)|^2}}{|W(u,v,\lambda)|^2}\right) dudvd\lambda \text{ (bits} - \text{nm}^{-1}) \tag{5}$$

where $|B|$ is the spectral bin width (in nanometers). Lastly, this result can be extended to the discrete-to-discrete imaging case (offered without proof):

$$\log N = \frac{1}{2|B|}\sum_{u,v,\lambda} \log\left(1 + \frac{\overline{|G[u,v,\lambda]|^2}}{|W[u,v,\lambda]|^2}\right) \text{ (bits} - \text{nm}^{-1}) \tag{6}$$

where periodograms are used to represent the power spectra. This version is the one used in the empirical studies described in subsequent sections.

3.2 Task-Based Information

An understanding of the information content in the optical channel is of central importance. However, in many cases it is desirable to understand the information content with respect to a particular image processing task, such as the detection of a tumor in medical imagery. There is a significant body of work devoted to the discovery of image quality metrics that might indicate how successful such tasks may be given a particular imaging system. For example, the relationship between Shannon information and receiver operating characteristic (ROC) curves has been published for detection problems [14]. The concept of task specific information (TSI) relates the canonical Shannon information to the minimum mean-squared-error (MSE) estimator for a given task [15]. This work focuses on the application of TSI to understanding shadow imaging system resolution through the examination of the knife edge localization task considered in [12].

3.2.1 TSI Problem Statement

The following is a brief review of the statement of the TSI problem.

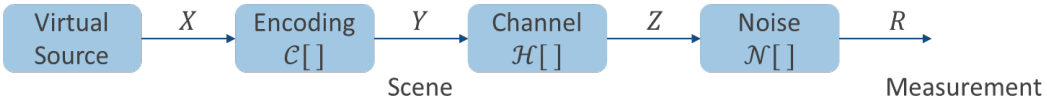

Figure 2. A model of the forward imaging channel.

The components of an imaging system are show above in Figure 2,where blocks in the diagram represent the following:

- The scene Y is processed by the optical channel to yield $Z = \mathcal{H}(Y)$
- Z is then corrupted to yield the noisy measurement $R = \mathcal{N}(Z)$
- The encoder $\mathcal{C}$ (deterministic or stochastic) produces a scene based on input from the virtual source

Task Specific Information: information in an image or signal that is *relevant* to a specific task. TSI is the Shannon mutual information $I(X;R)$ between the virtual source and the measurements:

$$TSI = I(X;R) = J(X) - J(X|R) \tag{7}$$

where $J(X) = -\mathbb{E}\{\log(pr(X))\}$ is the entropy of X, and $J(X|R) = -\mathbb{E}\{\log(pr(X|R))\}$ is the entropy of X conditioned on the measurement R, $\mathbb{E}\{\cdot\}$ denotes statistical expectation, and $pr(\cdot)$ denotes the probability density function, and logarithms are taken to the base 2. The entropy of X defines the maximum TSI content of any image measurement. Computing TSI directly from this equation is generally intractable [15]. One may consider the application of TSI to the task of knife edge localization:

TSI takes advantage of a new information theory result which provides a direct relationship between the mutual information $I(X;R)$ and the minimum mean-square error ($mmse$) in estimating X from R. Given the additive Gaussian channel $R = \sqrt{s}X + N$, where N is additive Gaussian noise with variance $\sigma^2 = 1$, and s is the signal-to-noise ratio, the relationship is given by:

$$\frac{d}{ds}I(X;R) = \frac{1}{2}mmse = \frac{1}{2}\mathbb{E}[|X - \mathbb{E}(X|R)|^2] \tag{8}$$

where $\mathbb{E}(X|R)$ is the conditional mean estimator [16]. Results for Poisson channels have also been published [17]. Extension to the linear vector Gaussian channels where $\mathcal{H}(\vec{X}) = \mathbf{H}\vec{X}$, where $\mathbf{H}$ denotes the matrix channel operator and $\vec{X}$ is the vector channel input are also available $\vec{R} = \sqrt{s}\mathbf{H}\vec{X} + \vec{N}$k, where $\vec{N}$ follows a multi-variate Gaussian distribution with covariance $\Sigma_{\vec{N}}$. In this case, the relationship is given by:

$$\frac{d}{ds}I(\vec{X};\vec{R}) = \frac{1}{2}\mathbb{E}\left[\left\|\mathbf{H}\vec{X} - \mathbb{E}(\mathbf{H}\vec{X}|\vec{R})\right\|^2\right] = \frac{1}{2}\mathrm{Tr}\left(\mathbf{H}^{\dagger}\Sigma_{\vec{N}}^{-1}\mathbf{H}\mathbf{E}\right) \tag{9}$$

where $\mathbf{H}^{\dagger}$ is the Hermitian conjugate of $\mathbf{H}$, and

$$\mathbf{E} = \mathbb{E}\left[\left(\vec{X} - \mathbb{E}(\vec{X}|\vec{R})\right)\left(\vec{X} - \mathbb{E}(\vec{X}|\vec{R})\right)^T\right] \tag{10}$$

Results are also available for a stochastic encoding of $\vec{X}$ [16]:

$$\vec{R} = \sqrt{s}\mathbf{H}\mathbf{C}(\vec{X}) + \vec{N} \tag{11}$$

The TSI estimation problem for the shadow imaging modality with a knife edge satellite aperture function can be formulated using the structure for the general localization task in [15], Section 2. This can be summarized as:

- The virtual source X localizes the knife edge to one of Q regions in the satellite aperture plane
- The encoding operator $\mathbf{C}_{ke}$ is linear, and can be stochastic

- The channel operator **H** is linear (discrete-to-discrete) and represents the wave propagation from the aperture plane to the ground plane
- The noise model $\vec{N}$ is zero-mean AWGN with covariance $\Sigma_{\vec{N}}$

The upper bound on TSI for this problem is the entropy of X

$$J(X) = -\sum_{q=1}^{Q} \Pr(X = q) \log \Pr(X = q) \leq [\log Q] \text{ bits} \tag{12}$$

From the diagram of Figure 2, the imaging model is given by

$$\vec{R} = \sqrt{s}\mathbf{H}\mathbf{C}_{ke}(X) + \vec{N} \tag{13}$$

The stochastic encoding operator is given by

$$\mathbf{C}_{ke}(X) = \mathbf{K}\boldsymbol{\Lambda}(X)\vec{\rho} \tag{14}$$

where:

- **K** is a matrix whose columns contain all possible $M \times M$ knife edge images, lexicographically stacked
- $\boldsymbol{\Lambda}(X)$ acts as a localization matrix and selects all knife edge images in the region specified by the source X
- $\vec{\rho}$ denotes the position of the knife edge within a region, and is a nuisance parameter
- If used, for $X = i$, $\vec{\rho}$ randomly selects one of the P_i available knife edge images, making $\mathbf{C}_{ke}$ a stochastic encoder
- Precise definitions of these variables generally follow Section 2.D of Ref. [15]

Substituting into Eq. (8) yields a means to compute the TSI by numerically integrating the $mmse_H$ function.

$$TSI(s) = I\left(X; \vec{R}\right) = \frac{1}{2}\int_0^s mmse_H(s')\, ds' \tag{15}$$

where

$$mmse_H(s) = \frac{1}{2}\mathrm{Tr}\left(\mathbf{H}^{\dagger}\Sigma_{\vec{N}}^{-1}\mathbf{H}\left(\mathbf{E}_{\vec{Y}} - \mathbf{E}_{\vec{Y}|X}\right)\right), \tag{16}$$

$$\mathbf{E}_{\vec{Y}} = \mathbb{E}\left[\left(\vec{Y} - \mathbb{E}\left(\vec{Y}|\vec{R}\right)\right)\left(\vec{Y} - \mathbb{E}\left(\vec{Y}|\vec{R}\right)\right)^T\right] \tag{17}$$

$$\mathbf{E}_{\vec{Y}|X} = \mathbb{E}\left[\left(\vec{Y} - \mathbb{E}\left(\vec{Y}|\vec{R}, X\right)\right)\left(\vec{Y} - \mathbb{E}\left(\vec{Y}|\vec{R}, X\right)\right)^T\right] \tag{18}$$

Closed form solutions for $\mathbb{E}\left(\vec{Y}|\vec{R}\right)$ and $\mathbb{E}\left(\vec{Y}|\vec{R}, X\right)$ are given in Ref. [15], Appendix A. In the following section the implementation of the $mmse_H$ estimator for this problem is discussed.

3.2.2 Knife Edge Location Estimator

In [12], a maximum likelihood estimator was derived for fixed knife edge location. The parameter which locates the knife edge in the present study is random and provides a source of entropy. The Bayes estimation using the mean-square error cost function and maximum a posteriori (MAP) estimation can be stated fairly directly. Following the development in [12], a series of photon arrivals at the observation plane may be specified by their spatial coordinates, $\{(x_j, y_j)\}$, the MAP estimate is the location where the following equation for $l(\theta)$ is maximized

$$\begin{aligned} &\ln p_{\{(X_j,Y_j)\}|\Theta}\left(\{(x_j, y_j)\}|\theta\right) = \\ &\ln p_{\Theta|\{(X_j,Y_j)\}}\left(\theta|\{(x_j, y_j)\}\right) + \ln p_{\Theta}(\theta) - \ln p_{\{(X_j,Y_j)\}}\left(\{(x_j, y_j)\}\right) \end{aligned} \tag{19}$$

$$l(\theta) \triangleq \ln p_{\Theta|\{(X_j,Y_j)\}}\left(\theta|\{(x_j, y_j)\}\right) + \ln p_{\Theta}(\theta) \tag{20}$$

and neglecting the right-most term which is not a function of the unknown parameter. It is well known that the minimum MSE estimate is given by the conditional mean:

$$\hat{\theta}_{ms} = \int_{-\infty}^{\infty} \theta \, p_{\Theta|\{(X_j,Y_j)\}}(\theta|\{(x_j,y_j)\})d\theta \tag{21}$$

See Section 2.4.1 of Ref. [18]. The PDF for the observed photon sequence (monochromatic case) is given by

$$p_{\{(X_j,Y_j)\}|\theta}(\{(x_j,y_j)\}|\theta) = \frac{e^{-\bar{N}}}{N!}\prod_{j=1}^{N}\Lambda\left((x_j,y_j);\theta\right) \tag{22}$$

where $\Lambda\left((x_j,y_j);\theta\right)$ is the intensity function at the observation plane in photons/meter^2, and N is the total number of photons received at the observation plane, given by a Poisson distribution of mean $\bar{N} = \int \Lambda\left((x_j,y_j);\theta\right)$.

If the parameter were fixed, then the likelihood function would be given by

$$l_{ml}(\theta) = \bar{N}(\theta) + \sum_{j=1}^{N} \ln\left[\Lambda\left((x_j,y_j);\theta\right)\right] \tag{23}$$

and the maximum-likelihood estimate is the value of θ which maximizes this likelihood given the data $\{(x_j,y_j)\}$

It is necessary to derive the functional form of

$$p_{\Theta|\{(X_j,Y_j)\}}(\theta|\{(x_j,y_j)\}) \tag{24}$$

used in the MAP equation. For the knife edge case, estimating the parameter θ is equivalent to estimating the intensity function since the function is completely specified by it. See Chapter 3 of Ref. [19].

$$\Lambda_{KE}\left((x_j,y_j);\theta\right) = I_{KE0}g_{KE0}\left((x_j-\theta,y_j)\right) \tag{25}$$

For a vertically oriented knife edge, we can sum over y_j

$$\Lambda_{KE}(x;\theta) = I_{KE0}g_{KE0}(x-\theta) \tag{26}$$

where I_{KE0} is now $I_0 W_y$ (where I_0 is the photoelectron rate/unit area, and W_y is the meridional length of the shadow imaging telescope array). The summed distribution can be well approximated as Gaussian by appeal to the Central Limit Theorem. The probability mass function of the meridionally summed counts in bin X_i with zonal axis bin center at x_i, is a lattice random variable, and can be approximated as:

$$P\{X_i = k|\theta\} = f_{X_i|\theta}(k) = \frac{1}{\sigma_i\sqrt{2\pi}}e^{-(k-\eta_i)^2/2\sigma_i^2} \tag{27}$$

where $k = 0,1,2,$ ..., and

$$\eta_i = \eta_{i,1} + \cdots + \eta_{i,n_y} \tag{28}$$

$$\sigma_i^2 = \sigma_{i,1}^2 + \cdots + \sigma_{i,n_y}^2 \tag{29}$$

given $n_y = W_y/\Delta y$ bins in the meridional direction. Thus,

$$\eta_i = \Lambda_{KE}(x_i;\theta) = I_{KE0}g_{KE0}(x_i-\theta) \tag{30}$$

$$\sigma_i^2 = \eta_i \tag{31}$$

The probability mass function of an observed knife edge profile bin counts $\boldsymbol{k} = \{k_1, k_2, \ldots, k_{n_x}\}$ is given by:

$$f_{\boldsymbol{X}|\Theta}(\boldsymbol{k}) = \prod_{i=1}^{n_x} f_{X_i|\Theta}(k_i) \tag{32}$$

where $n_x = W_x/\Delta x$ is the number of bins in the zonal direction. See Section 8-4, A. of Ref. [20]. The minimum MSE knife edge estimate is given by the mean of the a posteriori density:

$$\hat{\theta}_{ms} = \int_{-\infty}^{\infty} \theta \, p_{\Theta|\boldsymbol{X}}(\theta|\boldsymbol{x}) d\theta \tag{33}$$

To find $p_{\Theta|\boldsymbol{X}}(\theta|\boldsymbol{x})$, use

$$p_{\Theta|\boldsymbol{X}}(\theta|\boldsymbol{x}) = \frac{p_{\boldsymbol{X}|\Theta}(\boldsymbol{x}|\theta) p_{\Theta}(\theta)}{p_{\boldsymbol{X}}(\boldsymbol{x})} \tag{34}$$

taking logarithms and neglecting $p_{\boldsymbol{X}}(\boldsymbol{x})$,

$$l(\theta) \triangleq \ln p_{\boldsymbol{X}|\Theta}(\boldsymbol{x}|\theta) + \ln p_{\Theta}(\theta) \tag{35}$$

The MAP estimate $\hat{\theta}(\boldsymbol{X})$ is found by maximizing the above equation. For the case in which θ is uniformly distributed, the value of θ which maximalizes it is that which minimizes the inner term in the multivariate Gaussian distribution:

$$\tilde{\theta} = \arg\min_{\theta} \left(\boldsymbol{g}_{meas} - \boldsymbol{g}_{ref}(x-\theta) \right)^{T} \Sigma^{-1} \left(\boldsymbol{g}_{meas} - \boldsymbol{g}_{ref}(x-\theta) \right) \tag{36}$$

where $\boldsymbol{g}_{ref}$ is the reference Fresnel profile, shifted by θ. Both functions are normalized to one in the illuminated region for better stability in the computations. An alternative would be to scale $\boldsymbol{g}_{ref}$ so that it matches the value of $\boldsymbol{g}_{meas}$ in the illuminated region. The choice of Σ is important as the Gaussian approximation to the Poisson noise discussed earlier is signal-dependent. Ideally, it should be chosen with diagonal values proportional to $\boldsymbol{g}_{ref}$. For simplicity and to avoid ill-conditioning in the dimly lit region, Σ was chosen to be the identity matrix. The alternative was examined, but the dark region to the left produced very large values in the inverse and destabilized the solution in the photon scarce regime. Additionally, a stray background photon can be unduly amplified in importance and disrupt the solution. The minimization approach taken in the work was to perform a global optimization (a brute force grid search) across a range of values of θ to locate the maximum of Eq. (36).

4 Shadow Imaging Simulation Methodology

The studies discussed herein rely on Monte Carlo methods utilizing high quality forward imaging simulation as a basis for information computations. In simulation, one has the advantage of being able to separate the signal and noise, which is helpful for computations based on the Fellget and Linfoot equation. For TSI studies, one can know a priori the entropy of the virtual source used to generate scenes, and therefore possess an upper bound on the information in the output of the system.

The shadow simulation is performed using a radiometrically based numerical wave optics propagation which generates the truth shadows as described in [9]. The light collection and measurement processes are then modeled in the simulation to yield the measurement inferred shadows. The measurement inferred shadows are then used in a reconstruction process. The overall process is described in the following paragraphs.

4.1 Shadow Generation

Monochromatic irradiance patterns are generated over a wavelength range $\lambda_a \rightarrow \lambda_b$. Each pattern results from a multi-step wave optics propagation from the satellite plane to the observation plane. The following sequential steps are used to create each pattern:

1. Generate exoatmospheric complex electric field emerging from satellite plane based on source star brightness m_v and angular extent α, and the satellite's transmission function.
2. Propagate the electric field from satellite to top of atmosphere.
3. If incorporating atmospheric turbulence effects, generate altitude dependent random phase screens based on C_n^2 profile and apply to the electric field as it propagates through the atmosphere to the ground.
4. Convert the electric field on ground to irradiance.
5. Multiply the irradiance by atmospheric transmission factor.

6. Shift the irradiance pattern based on refraction per look angle relative to zenith.

4.2 Spectral Binning

The monochromatic shadow patterns are summed within each spectral bin to produce a set of polychromatic shadow patterns.

4.3 Sampling of the Patterns

The propagation through the vacuum and atmospheric regions are each performed using multiple steps by maintaining that $\Delta < \frac{\lambda z}{L}$, where Δ is the grid spacing in the propagation plane, z is the propagation distance, and L is the support length extent. The propagation plane parameters were chosen to be $\Delta = 0.1$ m, $L = 100$ m, and $M = 1000$, where M is the number of grid points, to avoid sampling issues over a 400 to 900 nm wavelength range.

4.4 Measurement Inferred Shadow

The measurement inferred shadows are generated from the pristine shadow patterns, along with the light collection system and detector parameters. Each aperture in the linear collection array has a diameter D. Thus, the light collection process requires the pristine pattern ($M \times M$ in size) to be downsampled to $M_{ap} \times M_{ap}$, where $M_{ap} = \frac{M}{D}$, while maintaining radiometric accuracy. The following steps describe the simulation process:

1. Calculate the signal and sky background photon rate patterns based on D, the optical throughput, the field of view, and the sky background brightness.
2. The signal and background photon rates are then fed into a detector model which yields a stochastic realization of the measured polychromatic photon rate patterns which includes sensor noise.
3. The $M_{ap} \times M_{ap}$ irradiance patterns are calculated then interpolated back to the final $M \times M$ measurement inferred patterns.

4.5 Image Reconstruction

In [9], a Fresnel integral version of the iterative Gerchberg-Saxton phase retrieval algorithm was proposed for image reconstruction on each simulation case to produce a silhouette reconstruction [7]. This is required under most circumstances to support the analysis of satellite visible features by an image analyst. However, such post-processing does not add information in the Shannon sense. Thus, the irradiance patterns are used directly in this work rather than the reconstructions. It is worth noting that in prior work visual estimates of the optical resolution were made using the reconstructed silhouettes.

4.6 Monte Carlo Method

Figure 3 shows the overall method used in the TSI Monte Carlo study. For a given source visual magnitude, the inner simulation loop was executed 100 times to provide statistics for the accuracy of the knife edge position estimates.

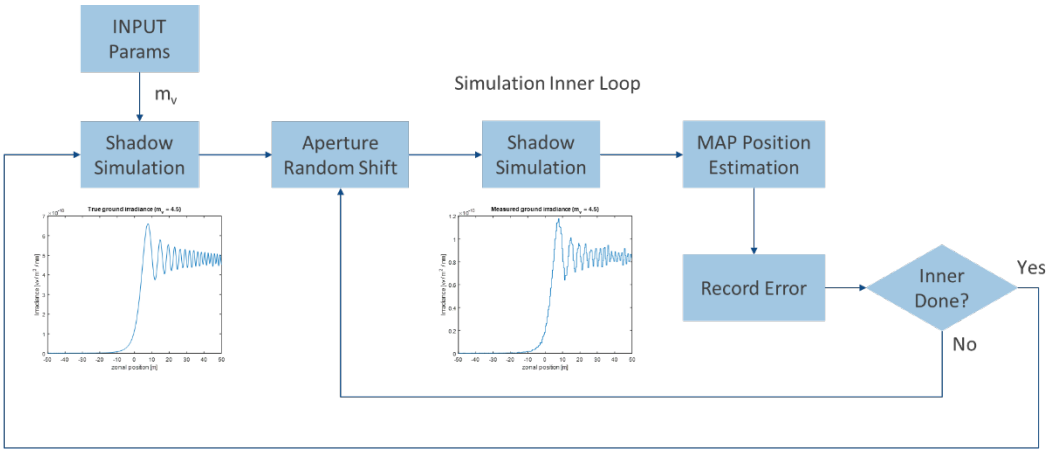

Figure 3. An overview of the Monte Carlo process for the TSI studies of knife edge localization.

5 Results and Discussion

A set of simulations were run and served as the basis for information experiments. A variety of cases are examined for the power-spectrum-based information calculations, including revisiting some earlier visual estimates of resolution using a test pattern with features of varying size. For the TSI study, a knife edge aperture function comprising a vertically oriented step function, with values of zero at left, and ones at right, was used as the transmission function. This plane was assumed to extend to infinity in all directions. Parameters for these studies are listed in the applicable table below.

5.1 Scene-Based Information Content

In this section, the Fellget and Linfoot power spectrum-based metric is used to explore the relationship between visually estimated resolution and information content. We revisit some of the earlier published results which were based on visual estimates of the smallest resolvable detail in silhouette images produced by the Gerchberg-Saxton algorithm mentioned above [9]. This is an empirical study using the shadow imaging wave propagation code.

Revisiting prior work using knife edge test patterns and estimation theory [12], results confirm the multispectral gain in terms of information. The results also indicate the presence of an information upper bound.

5.1.1 Simulation parameters

The simulation parameters used in this study are shown in Table 1 and are as similar as possible to those used in [9] for the visual estimation of resolvable detail in the silhouette images. The test pattern used is shown in Figure 4.

Table 1. Simulation parameters for scene-based results.

Propagation Parameters	
Sampling	**Value**
support length L	100 m
grid spacing Δ	0.1m
grid samples M	1000
Object	**Value**
satellite model	barSat2
Source Star	**Value**
magnitude m_v	6 - 12
angular extent α	0 nrad
Environmental	**Value**
turbulence	Off
sky brightness m_B	20 $m_v/arc\ sec^2$
Observational	**Value**
off-zenith θ_v	0 deg
latitude ϕ_{lat}	0 deg
longitude ϕ_{long}	0 deg
altitude h	0 m

Measurement Parameters	
Light Collection	**Value**
aperture size D_{ap}	0.2, 0.4, 0.8 m
field of view θ_a	1 arc sec
Spectral	**Value**
beginning λ_a	400
ending λ_b	900
width of Λ_j	5, 10, 25, 50, 100 nm
number of bins J	100, 50, 20, 10, 5
Detector	**Value**
exposure time t	77, 154, 308 μs
dark count rate N_D	0 Hz
detection efficiency P_d	1
after-pulsing P_{ap}	0
gate time t_g	0.1 μs
duty cycle c_d	1

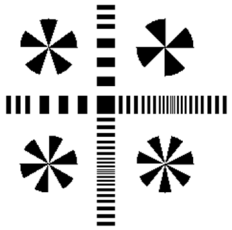

Figure 4. The barSat2 transmission function. Dark areas are opaque and represent features in the 1 – 5 m, and 1 m range.

5.1.2 Spatial Resolution (Aperture Size vs. Bin Width)

This study looks at the influence of aperture diameter and spectral bin width on system resolution. Aperture sizes of 0.2, 0.4, and 0.8 meters were used. Spectral bin widths of 5, 10, 25, 50, 100 nm were included. All combinations of these two parameters were examined.

Table 2 is a depiction of one of the main results presented in [9], where the smallest resolvable detail in the barSat2 test pattern was visually estimated. The estimates were performed using the reconstructed silhouettes produced by the Fresnel-modified Gerchberg-Saxton algorithm. Note that the test pattern does not contain features of size 0.6 – 0.9 meters, so estimates in this range were not possible.

Table 3 shows the results of computing the power spectrum-based information content in the measurement inferred ground irradiance patterns. The processing to produce a silhouette image from the patterns would not add information in the theoretical sense, so the silhouette images were not used here.

Table 2. Visual estimates of smallest resolvable detail in the silhouette images as a function of aperture size and spectral bin width. Star brightness is m_V=6.

Limits to Resolution Deduced from Reconstructed Images					
	$\Delta\lambda = 5\ nm$	$\Delta\lambda = 10\ nm$	$\Delta\lambda = 25\ nm$	$\Delta\lambda = 50\ nm$	$\Delta\lambda = 100\ nm$
$D_a = 0.2\ m$	0.2 m	0.3 m	0.4 m	0.5 m	1.0 m
$D_a = 0.4\ m$	0.3 m	0.3 m	0.4 m	0.5 m	1.0 m
$D_a = 0.8\ m$	0.5 m	0.5 m	1.0 m	1.0 m	1.0 m

Table 3. Mean spectral bin information (bits) as a function of aperture size and spectral bin width. Star brightness is m_V=6.

Mean Spectral Bin Information Capacity for Varying Aperture Size and Spectral Width					
	$\Delta\lambda = 5\ nm$	$\Delta\lambda = 10\ nm$	$\Delta\lambda = 25\ nm$	$\Delta\lambda = 50\ nm$	$\Delta\lambda = 100\ nm$
$D_a = 0.2\ m$	66512.765	43172.391	22444.090	13149.343	7353.983
$D_a = 0.4\ m$	30880.315	18076.962	8341.942	4387.670	2097.559
$D_a = 0.8\ m$	11786.258	6258.788	2407.772	930.429	172.265

The loss of information capacity is quite apparent when moving away from the upper left corner. In Figure 5, the information content is plotted against the visually estimated resolution. The visual estimates appear well correlated with information capacity plotted on a log scale. Visual assessments of 0.6 – 0.9 m features were not made as they were not present in the original test pattern. As a result, the 1 m resolution bin captures the resolutions that are not 0.5 m or less. This is likely the cause of the spread in the estimates at 1 m resolution.

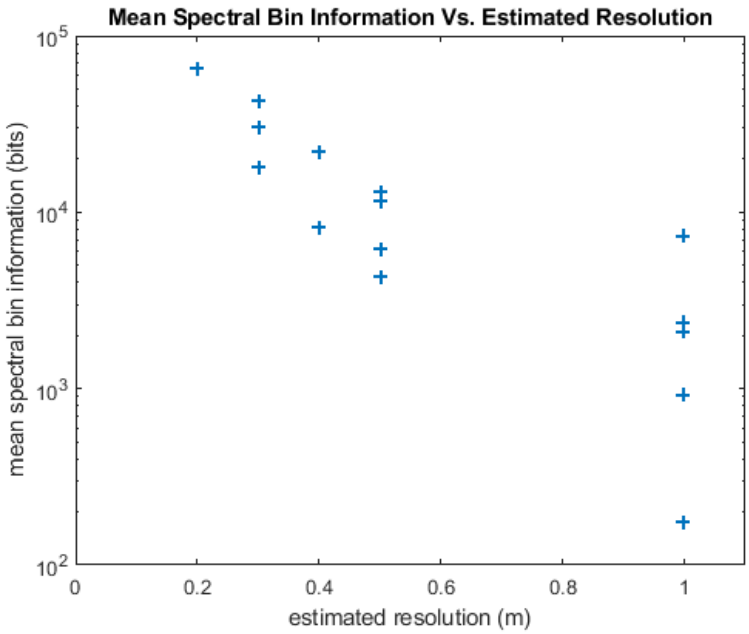

Figure 5. Information vs. Visually Estimated Resolution.

5.1.3 Varying Source Visual Magnitude

Figure 6 shows the results for fixing the aperture size and varying the stellar source magnitude. The internal consistency of the information capacity calculation is evident as is the multispectral gain in information density from using smaller spectral bins. It should be noted that this result is for $D_{ap} = 0.4$ m, with other simulation parameters as shown in Table 1.

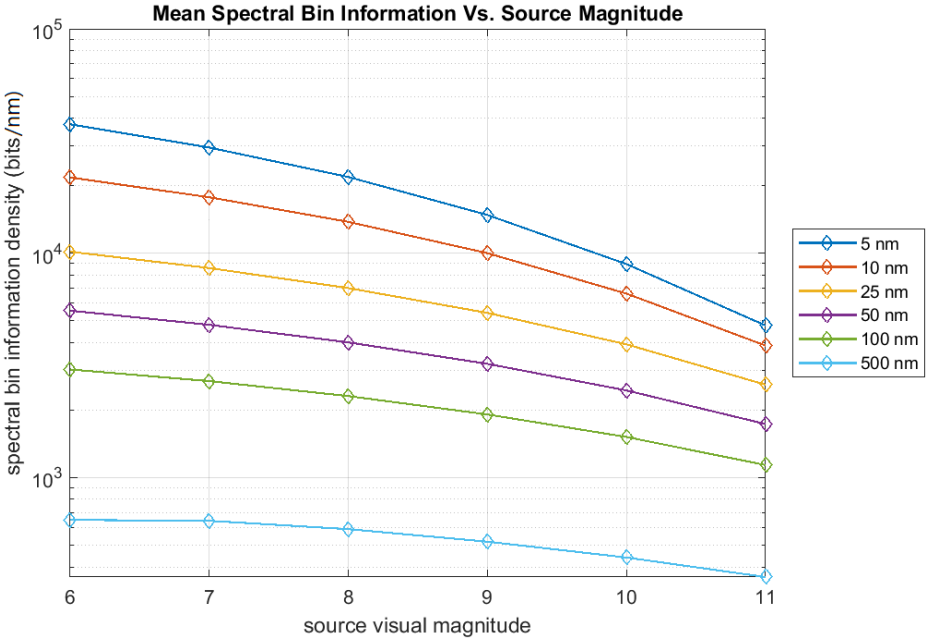

Figure 6. Information Capacity Versus Source Magnitude for the barSat2 aperture.

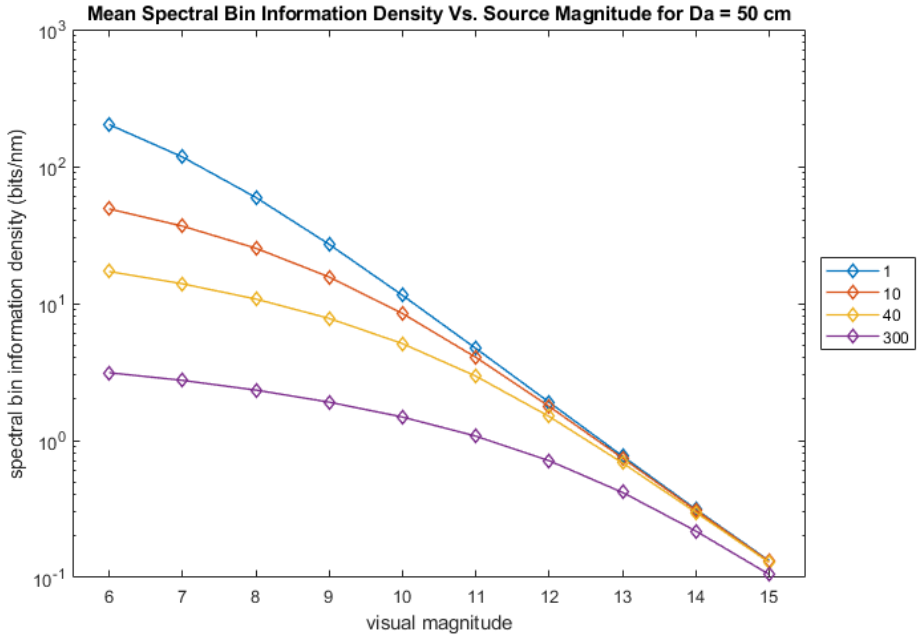

Figure 7. Multi-spectral bin information density versus source visual magnitude.

5.1.4 Knife Edge Aperture Function and Source Visual Magnitude

In this case, information capacity was calculated from the signal and noise functions in the knife edge diffraction pattern, with the multispectral gain evident again in Figure 7. The curves appear to be approaching an upper bound, likely due to photon scarcity and the absence of dark current noise, i.e., for a vanishing signal-to-noise ratio. Simulation parameters were chosen to match [12] as closely as possible and are shown in Table 4.

5.2 Task-Specific Information Content

5.2.1 Overview

The relationship between mutual information and minimum mean-squared error is explored here to understand the limits of resolution of a shadow imaging system. The ability to locate the position of a knife edge profile serves to estimate the resolution. Empirical results using the wave propagation software and Monte Carlo studies are presented in the following sections.

5.2.2 Simulation Parameters

The simulation parameters are summarized in Table 4. The transmission function is a simple step function located at the mid-point of the object plane. The resulting diffraction pattern resulting from placing the transmission function at GEO distance in a vertical orientation is shown in Figure 8. This is a horizontal slice through the pattern after normalization of the illuminated region to one. A collection aperture size of $D_{ap} = 0.5$ m was used along with quasi-monochromatic spectral bin widths of 1 nm to follow the approach taken in [12] as closely as possible. However, this study utilized circular apertures rather than square ones.

Table 4. Simulation parameters for TSI study.

Propagation Parameters	
Sampling	**Value**
support length L	100 m
grid spacing Δ	0.1m
grid samples M	1000
Object	**Value**
satellite model	vertical_500
Source Star	**Value**
magnitude m_v	6 - 15
angular extent α	0 nrad
Environmental	**Value**
turbulence	Off
sky brightness m_B	20 $m_v/arc\ sec^2$
Observational	**Value**
off-zenith θ_v	0 deg
latitude ϕ_{lat}	0 deg
longitude ϕ_{long}	0 deg
altitude h	0 m

Measurement Parameters	
Light Collection	**Value**
aperture size D_{ap}	0.5 m
field of view θ_a	1 arc sec
Spectral	**Value**
beginning λ_a	550
ending λ_b	551
width of Λ_j	1
number of bins J	1
Detector	**Value**
exposure time t	7.6828e-05 s
dark count rate N_D	0 Hz
detection efficiency P_d	0.3
after-pulsing P_{ap}	0
gate time t_g	0.1 μs
duty cycle c_d	0.95

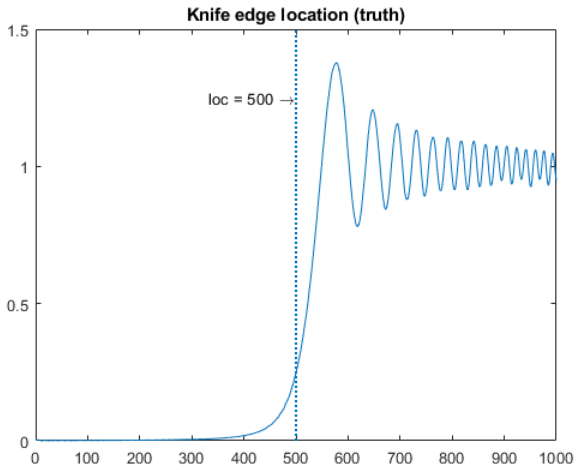

Figure 8. One-dimensional horizontal slice through the knife edge diffraction pattern profile produced by a step function aperture at GEO distance. Illuminated region normalized to a mean value of one.

Other notable experimental parameters are:

- For each stellar source visual magnitude, 100 random trials were conducted.
- Source magnitudes ranged from 6 – 12 in 0.1 magnitude increments.
- The random knife edge profile positions were drawn with replacement from the interval [-50, 50] pixels, corresponding to an interval of 10 m in 0.1 m steps.
- The TSI is limited to an upper bound of $\log_2(101) = 6.6582$ bits by design.

Note that the source visual magnitude here acts as a stand-in for the parameter s in the TSI discussion earlier. One could relate the mean signal-to-noise ratio in the measurement inferred diffraction patterns in the illuminated region to this parameter, but we have omitted this detail.

5.2.3 TSI for Knife Edge Localization Task Result

Figure 9 shows the result from the Monte Carlo study. The TSI values are the result of summing minimum mean-squared error to a point along the source visual magnitude axis as in Eq. (15). The summation is from the right to be in the direction of increasing SNR. The TSI is co-plotted with the RMSE rather than the MMSE (on a log scale) to support comparison with the results in [12]. By design, the TSI for this problem is limited to $\log_2(101) = 6.6582$ bits. The TSI curve saturates at a significantly lower value, however. This phenomenon was also seen in the results presented in [15] for the projective imaging system. There are several possible contributing factors here which could introduce information losses:

- Free space propagation and diffraction imposes a loss of information.
- The calculation might benefit from extension to even fainter sources.
- The GMAPD detector model was used which includes the effect of gating and may lower signal-to-noise.
- A sky background source was included, which although faint, may contribute to loss of signal-to-noise.
- The asymptotic behavior of Poisson noise with vanishing SNR may be different than that of AWGN.
- There is no dark current “floor” in this study as signal strength fades.

Additional investigation is needed, although this initial result appears to be promising.

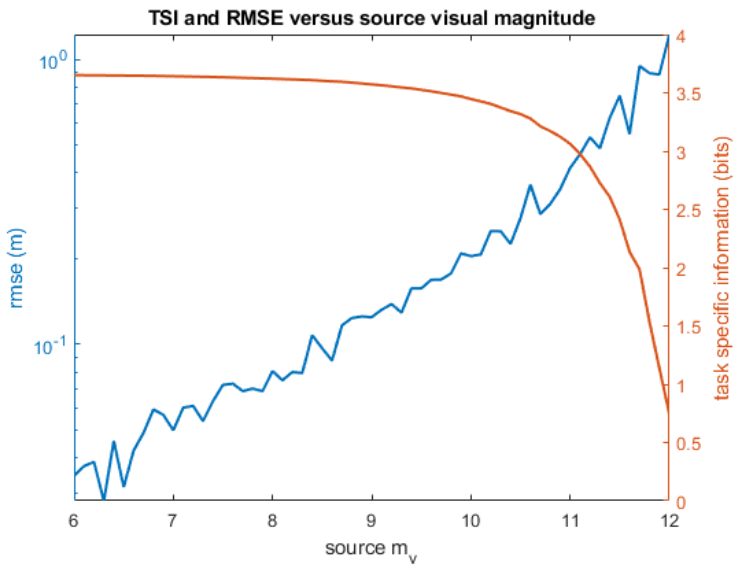

Figure 9. RMSE and TSI versus source visual magnitude.

6 Summary

In this work, both scene-based information and task specific information were investigated. This initial investigation of shadow imaging performance as a function of Shannon information has resulted in the following insights:

- The visual resolution assessments in [9] were revisited using the power spectrum-based information computational approach which demonstrated a reasonable correlation between the visual assessments and measured information content.
- The knife edge diffraction pattern information content was similarly investigated, corroborating the observed multi-spectral gain reported in [12], but in terms of information content.
- It has been shown that the knife edge localization problem can be approached using task specific information (TSI).
- Further study is needed to fully understand why the TSI saturates at a value lower than the upper bound, and to quantify the influence of the various contributing factors discussed above.

We anticipate future work in the following areas:

- Relating the observed upper bounds on information for the power spectrum-based information metric more closely with the bounds examined in [12].
- Consider Guo's Poisson channel result for applicability and tractability regarding TSI [17].
- Explore other tasks for use with TSI which are of value in SDA, e.g., closely spaced object detection.

7 References

[1] U. S. Space Force, "Space Power: Doctrine for Space Forces," 2020. [Online]. Available: https://apps.dtic.mil/sti/pdfs/AD1129735.pdf. [Accessed 16 January 2024].

[2] G.E. Taylor, "An occultation by a minor planet," *MNASSA : Monthly Notes of the Astronomical Society of South Africa*, **11**(4), pp. 33–34, January 1952. https://hdl.handle.net/10520/AJA00248266_1218.

[3] L.H. Wasserman, *et al*., "The diameter of Pallas from its occultation of SAO 85009," *The Astronomical Journal*, **84**(2), pp. 259-268, February 1979. https://doi.org/10.1086/112416.

[4] M.W. Buie, *et al*., "Size and Shape Constraints of (486958) Arrokoth from Stellar Occultations," *The Astronomical Journal*, **159**(4), p. 130, 26 February 2020. https://doi.org/10.3847/1538-3881/ab6ced.

[5] F. Roques, G. Georgevits and A. Doressoundiram, "The Kuiper Belt Explored by Serendipitous Stellar Occultations," in The Solar System Beyond Neptune, M.A. Barucci, H. Boehnhardt, D.P. Cruikshank and A. Morbidelli, Eds., Tucson, University of Arizona Press, May 2008.

[6] R.H. Burns, V. Gamiz, J.J. Dolne, J. Lambert and S. Long, "Shadow imaging of GEO satellites," *Proc. SPIE* 5896, Unconventional Imaging, 58960C (23 August 2005); https://doi.org/10.1117/12.634820.

[7] R. W. Gerchberg and W.O. Saxton, "A practical algorithm for the determination of phase from image and diffraction plane pictures," *Optik*, **35**, pp. 237-246, 1972.

[8] J. Luu, L. Jang and B. Willard, "Shadow Imaging Efforts at MIT Lincoln Laboratory," in Advanced Maui Optical and Space Surveillance (AMOS) Technologies Conference, 2008.

[9] D.M. Douglas, "Shadow imaging of geosynchronous satellites," PhD Thesis, University of Arizona, 2014.

[10] D. Douglas, B.R. Hunt and D.G. Sheppard, "Recent Developments in Shadow Imaging of Geosynchronous Satellites," in Advanced Maui Optical and Space Surveillance (AMOS) Technologies Conference, 2016.

[11] D.M. Douglas, B.R. Hunt, D.G. Sheppard, "Resolution limits for shadow imaging of geosynchronous satellites: analytic and simulated approaches," *Proc. SPIE* 10410, Unconventional and Indirect Imaging, Image Reconstruction, and Wavefront Sensing 2017, 104100O (6 September 2017); https://doi.org/10.1117/12.2274837.

[12] P.J. McNicholl and P.N. Crabtree, "Statistical bounds and maximum likelihood performance for shot noise limited knife-edge modeled stellar occultation," *Proc. SPIE* 9227, Unconventional Imaging and Wavefront Sensing 2014, 922707 (18 September 2014); https://doi.org/10.1117/12.2059317.

[13] P.B Fellgett and E.H. Linfoot, "On the assessment of optical images," *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences*, **247**, pp. 369-407, 1955. https://doi.org/10.1098/rsta.1955.0001.

[14] E. Clarkson and J.B. Cushing, "Shannon information for joint estimation/detection tasks and complex imaging systems," *J. Opt. Soc. Am. A* **33**(3), 286-292 (2016). https://doi.org/10.1364/JOSAA.33.000286.

[15] M.A. Neifeld, A. Ashok and P.K. Baheti, "Task-specific information for imaging system analysis," *J. Opt. Soc. Am. A*, **24**(12), B25-B41 (December 2007). https://doi.org/10.1364/JOSAA.24.000B25.

[16] D. Guo, S. Shamai and S. Verdu, "Mutual information and minimum mean-square error in Gaussian channels," in *IEEE Transactions on Information Theory*, 51(4), pp. 1261-1282 (April 2005). https://doi.org/10.1109/TIT.2005.844072.

[17] D. Guo, S. Shamai and S. Verdu, "Mutual Information and Conditional Mean Estimation in Poisson Channels," in *IEEE Transactions on Information Theory*, 54(5), pp. 1837-1849 (May 2008). https://doi.org/10.1109/TIT.2008.920206.

[18] H.L. Van Trees, Detection, Estimation, and Modulation Theory, Part I, New York: John Wiley & Sons, 1968.

[19] R.L. Streit, Poisson Point Processes, New York: Springer, 2010.

[20] A. Papoulis, Probability, Random Variables, and Stochastic Processes, Third Edition, New York: McGraw-Hill, 1981.