Task-Specific Information: An Imaging System Analysis Tool

Amit Ashok[a], Pawan K. Baheti[a] and Mark A. Neifeld[a, b]

[a]Department of Electrical and Computer Engineering, 1230 East Speedway Blvd, University of Arizona, Tucson, AZ 85721, USA;
[b]College of Optical Sciences, 1630 East University Boulevard, University of Arizona, Tucson, AZ 85721, USA

ABSTRACT

We present a novel method for computing the information content of an image. We introduce the notion of task-specific information (TSI) in order to quantify imaging system performance for a given task. This new approach employs a recently-discovered relationship between the Shannon mutual-information and minimum estimation error. We demonstrate the utility of the TSI formulation by applying it to several familiar imaging systems including (a) geometric imagers, (b) diffraction-limiter imagers, and (c) projective/compressive imagers. Imaging system TSI performance is analyzed for two tasks: (a) detection, and (b) classification.

Keywords: Entropy, Mutual Information, Minimum-mean square error, Diffraction-limited imager

1. INTRODUCTION

Information content of an image is an important measure that finds applications ranging from evaluating compression algorithm performance to designing imaging systems.[1–5] However, the computation of image information content remains a challenging problem. The problem is made difficult by (a) the high dimensionality of useful images, (b) the complex/unknown correlation structure among image pixels, and (c) the lack of relevant probabilistic models. It is possible to approximate the information content of an image by using some simplifying assumptions. For example, Gaussian and Markovian models have both been used to estimate image information.[4–6] Transform-domain techniques have also been studied (e.g., wavelet prior models).[7,8]

Imagery is often used in support of a computational task (e.g., automated target recognition). It is important to note that not all information contained in the imagery is relevant to the task. For example, in a target detection task the final outcome is a binary variable, representing the two target states "target present" or "target absent". Therefore, we could say that the image may contain no more than 1 bit of *relevant* information for the target detection task. We will refer to this relevant information as task-specific information (TSI) and the remainder of this paper represents an effort to describe/quantify TSI as an analysis tool for several tasks and imaging systems of interest. What we describe here is a formal approach to the computation of TSI. Such a formalism is important primarily because it enables imager design and/or adaptation that strives to maximize the TSI content of measurements. This has two implications: (a) imager resources can be optimally allocated so that irrelevant information is not measured and thus task-specific performance is maximized and/or (b) imager resources can be minimized subject to a TSI constraint thus reducing imager complexity, cost, size, weight, etc.

The remainder of this paper is organized as follows. Section 2 introduces a formal framework for the definition of TSI. We consider two example tasks: target detection and target classification. In Section 3

Further author information: (Send correspondence to Amit Ashok)
Amit Ashok: E-mail: ashoka@ece.arizona.edu, Telephone: 1 520 621 8980
Pawan K. Baheti: E-mail: baheti@ece.arizona.edu, Telephone: 1 520 626 6443
Mark A. Neifeld: E-mail: neifeld@ece.arizona.edu, Telephone: 1 520 621 6102

Visual Information Processing XVI, edited by Zia-ur Rahman, Stephen E. Reichenbach, Mark Allen Neifeld,
Proc. of SPIE Vol. 6575, 65750G, (2007) · 0277-786X/07/$18 · doi: 10.1117/12.720878

(a) we apply the TSI framework to an ideal geometric imager and a diffraction-limited imager for each of the two tasks, and (b) extend the framework to analyze principal component projective imager for detection task. Section 4 summarizes the TSI framework and offers concluding remarks.

2. TASK-SPECIFIC INFORMATION

We begin by considering the various components of an imaging system as shown in Fig. 1. In this model the scene Y provides the input to the imaging channel, represented by the operator $\mathcal{H}$, to yield $Z = \mathcal{H}(Y)$. The measurement Z is then corrupted by the noise operator $\mathcal{N}$ to yield the final measurement $R = \mathcal{N}(Z)$. The model in Fig. 1 is made task-specific via the incorporation of the virtual source and encoding blocks. The encoding operator $\mathcal{C}$ uses X to generate the scene according to $Y = \mathcal{C}(X)$. Here we consider $\mathcal{C}$ to be stochastic. The virtual source variable X represents the parameter of interest for a specific task and serves as a mechanism through which we define TSI. Other blocks in the imaging chain may add entropy to the image measurement R; however, only the entropy of the virtual source X is relevant to the task. We may therefore define TSI as the Shannon mutual-information $I(X;R)$ between the virtual source X and the image measurement R as: $TSI \equiv I(X;R) = J(X) - J(X|R)$, where $J(X) = -\mathbb{E}\{\log(pr(X))\}$ denotes the entropy of X, $J(X|R) = -\mathbb{E}\{\log(pr(X|R)\}$ denotes the conditional entropy of X given R, $\mathbb{E}\{\cdot\}$ denotes statistical expectation, $pr(\cdot)$ denotes the probability density function, and all the logarithms are taken to be base 2. Note that from this definition of TSI we have $I(X;R) \leq J(X)$ indicating that an image cannot contain more TSI than the entropy of the variable representing the task. For most realistic imaging problems computing TSI from its definition directly is intractable owing to the dimensionality and non-Gaussianity of R. Numerical approaches may also prove to be computationally prohibitive, even when using methods such as importance-sampling and/or Markov Chain Monte Carlo(MCMC).

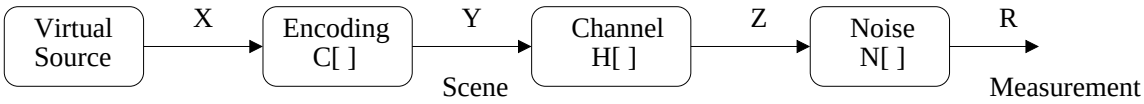


Figure 1. Block diagram of an imaging chain.

Recently, Guo et. al[9,10] have demonstrated a direct relationship between the minimum mean square error ($mmse$) in estimating $\vec{X}$ from $\vec{R}$, and the mutual-information $I(\vec{X};\vec{R})$ for a linear additive Gaussian channel, expressed as: $\vec{R} = \sqrt{s}\mathbf{H}\mathbf{C}(\vec{X}) + \vec{N}$, where $\mathcal{H}(\cdot) = \mathbf{H}(\cdot)$, and $\mathbf{H}$ denotes the matrix channel operator. Note that $\mathbf{C}(\vec{X})$ is a random function of $\vec{X}$. $\vec{N}$ denotes a zero mean addtive Gaussian noise vector with covariance $\mathbf{\Sigma}_{\vec{N}}$. The relationship stated by Guo is given as

$$\frac{d}{ds}I(\vec{X};\vec{R}) = \frac{1}{2}mmse(s) = \frac{1}{2}\mathrm{Tr}(\mathbf{H}^{\dagger}\mathbf{\Sigma}_{\vec{N}}^{-1}\mathbf{H}(\mathbf{E}_{\vec{Y}} - \mathbf{E}_{\vec{Y}|\vec{X}})), \tag{1}$$
$$\mathbf{E}_{\vec{Y}} = \mathbb{E}[||(\vec{Y} - \mathbb{E}(\vec{Y}|\vec{R}))||^2], \mathbf{E}_{\vec{Y}|\vec{X}} = \mathbb{E}[||(\vec{Y} - \mathbb{E}(\vec{Y}|\vec{R},\vec{X}))||^2].$$

Next, we apply this result to analyze imaging systems for detection, classification and localization tasks. Note that the imaging channel operator is assumed to be linear and deterministic in this work. The encoding operator $\mathcal{C}$ is assumed to be linear and stochastic.

2.1. Detection task

Let us begin by considering a target detection task, where a known target is to be detected in the presence of noise and clutter. The target position is unknown and hence for the target detection task, target position assumes the role of a nuisance parameter. Here, we have considered only one nuisance parameter, however extension to additional nuisance parameters is straightforward. The imaging model for this task is constructed as:

$$\vec{R} = \mathbf{H}\mathbf{C}_{\mathbf{det}}(X) + \vec{N}, \tag{2}$$

where the stochastic encoding operator $\mathbf{C_{det}}$ is defined as: $\mathbf{C_{det}}(X) = \sqrt{s}\mathbf{T}\vec{\rho}X + \sqrt{c}\mathbf{V_c}\vec{\beta}$. Here $\mathbf{T}$ is the target profile matrix, in which each column is a target profile (lexicographically ordered into a one-dimensional vector) at a specific position in the scene. For a scene of size $M \times M$ pixels and P different target positions, $\mathbf{T}$ is a $M^2 \times P$ matrix. The column vector $\vec{\rho}$ is a random indicator vector and selects the target position for a given scene realization. Therefore, $\vec{\rho} \in \{\vec{c}_1, \vec{c}_2 ... \vec{c}_P\}$ where $\vec{c}_i$ is a P-dimensional unit column vector with a 1 in the i^{th} position and 0 in all remaining positions. All positions are assumed to be equally probable, therefore $\Pr(\vec{\rho} = \vec{c}_i) = \frac{1}{P}$ for $\forall i$. The virtual source variable X takes the value 1 or 0 (i.e. "target present" or "target absent") with probabilities p and $1-p$ respectively. $\mathbf{V_c}$ is the clutter profile matrix whose columns represent various clutter components such as tree, shrub, grass etc. The dimension of $\mathbf{V_c}$ is $M^2 \times K$ where K represents the number of clutter components. $\vec{\beta}$ is the K-dimensional clutter mixing column vector, which determines the strength of various components that comprise clutter. $\vec{\beta}$ follows a multivariate Gaussian distribution with mean $\vec{\mu}_{\vec{\beta}}$ and covariance $\mathbf{\Sigma}_{\vec{\beta}}$. The coefficient c denotes the clutter-to-noise ratio. Note that clutter and detector noise combine to form a multivariate Gaussian random vector $\vec{N}_c = \sqrt{c}\mathbf{H}\mathbf{V_c}\vec{\beta} + \vec{N}$ with mean $\vec{\mu}_{\vec{N}_c} = \vec{\mu}_{\vec{\beta}}$ and covariance $\mathbf{\Sigma}_{\vec{N}_c} = \mathbf{H}\mathbf{V_c}\mathbf{\Sigma}_{\vec{\beta}}\mathbf{V_c}^T\mathbf{H}^T \cdot c + \mathbf{\Sigma}_{\vec{N}}$. Now, we can rewrite the imaging model as: $\vec{R} = \sqrt{s}\mathbf{H}\vec{Y} + \vec{N}_c$ where $\vec{Y} = \mathbf{T}\vec{\rho}X$. The task-specific information for the target detection task is therefore the mutual-information between the image measurement $\vec{R}$ and the virtual source X and is expressed as

$$TSI = I(X;\vec{R}) = \frac{1}{2}\int_0^s mmse_H(s')ds', \tag{3}$$

$$\text{where } mmse_H(s) = \mathrm{Tr}(\mathbf{H}^T\mathbf{\Sigma}_{\vec{N}_c}^{-1}\mathbf{H}(\mathbf{E}_{\vec{Y}} - \mathbf{E}_{\vec{Y}|X})), \tag{4}$$

$$\text{and } \vec{Y} = \mathbf{T}\vec{\rho}X. \tag{5}$$

Because X is a binary random variable with probability distribution: $\Pr(X=1) = p$ and $\Pr(X=0) = 1-p$, we can assert that

$$TSI \leq J(X) \leq 1\text{bit}, \tag{6}$$

where the entropy of X is $J(X) = -p\log(p) - (1-p)\log(1-p)$. Note that for this simple detection task the received signal $\vec{R}$ contains at most 1 bit of task-specific information.

2.2. Classification task

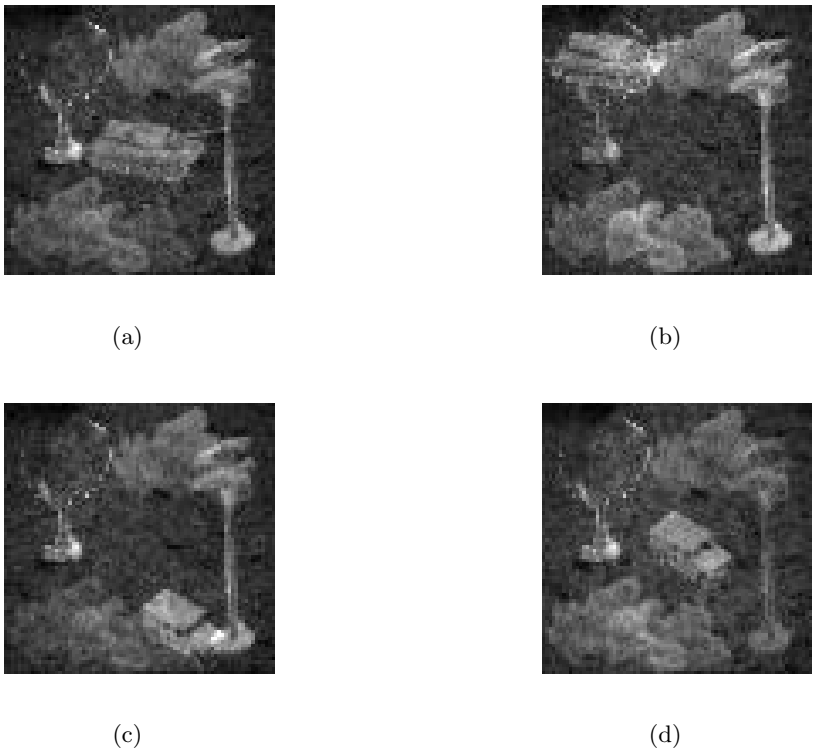
Next we consider a simple two-class classification problem for which we label the two possible states of nature (i.e., targets) as H_1 and H_2. The extension to more than two classes will be straightforward. The overall imaging model remains the same as in Eq. (2). The number of positions that each target can take remains unchanged. However, now $\mathbf{T}$ has dimensions $M^2 \times 2P$ and is given by $\mathbf{T} = [\mathbf{T_{H_1}}\mathbf{T_{H_2}}]$ where $\mathbf{T_{H_i}}$ is the target profile matrix for class i. The virtual source variable is denoted by the vector $\vec{X}$ and takes the values $[1,0]^T$ or $[0,1]^T$ to represent H_1 or H_2 respectively. The prior probabilities for H_1 and H_2 are p and $1-p$ respectively. The vector $\vec{\rho}$ from the detection problem becomes a matrix $\boldsymbol{\rho}$ of dimension $2P \times 2$ and is defined as

$$\boldsymbol{\rho} = \begin{bmatrix} \vec{\rho}_H & \mathbf{0} \\ \mathbf{0} & \vec{\rho}_H \end{bmatrix}, \tag{7}$$

where $\vec{\rho}_H \in \{\vec{c}_1, \vec{c}_2 .... \vec{c}_P\}$ and $\mathbf{0}$ is an all zero P-dimensional column vector. Once again we assume all positions to be equally probable, therefore $\Pr(\vec{\rho}_H = \vec{c}_i) = \frac{1}{P}$ for $i = \{1, 2, .., P\}$.

Consider an example that illustrates how the term $\mathbf{T}\boldsymbol{\rho}\vec{X}$ enables selection of a target from either H_1 or H_2 at one of P positions. In order to generate a target from H_1 at the m^{th} position in the scene, $\vec{\rho}_H = \vec{c}_m$ and $\vec{X} = [1,0]^T$. The product of $\mathbf{T}\boldsymbol{\rho}$ will produce a $M^2 \times 2$ matrix whose first column is equal to the H_1 profile at position m and whose second column is equal to the H_2 profile at the same position. This resulting matrix, when multiplied by $\vec{X} = [1\ 0]^T$, will select the H_1 profile. Similarly, in order to choose a target from H_2 at the m^{th} position, $\vec{\rho}_H = \vec{c}_m$ and $\vec{X} = [0\ 1]^T$.

The imaging model presented for the detection problem in Eq. (2) and the corresponding TSI defined in Eq. (3) require minor modification to remain valid for the classification problem. Specifically, we require



(a) (b)

(c) (d)

Figure 2. Example scenes: (a) Tank in the middle of the scene, (b) Tank in the top of the scene, (c) Jeep at the bottom of the scene, and (d) Jeep in the middle of the scene.

the virtual source variable to become a vector quantity $\vec{X}$, and the dimensions of $\mathbf{T}$ and $\boldsymbol{\rho}$ to be adjusted accordingly as noted above. Note that despite the increase in dimensionality, the binary source vector $\vec{X}$ results in the upper bound $TSI \leq 1$ bit for the two-class classification problem.

3. IMAGING SYSTEM ANALYSIS

The TSI framework described in the previous section allows us to evaluate the task-specific performance of an imaging system for a task defined by a specific encoding operator and virtual source variable. Two encoding operators corresponding to two different tasks: (a) detection, and (b) classification have been defined. Now we apply the TSI framework to evaluate the performance of a geometric imager and a diffraction-limited imager on these three tasks.

We begin by describing the source, object and clutter components of the scene mode. The source variable X in the detection task represents "*tank present*" or "*tank absent*" conditions with equal probability i.e. $p = \frac{1}{2}$. In the classification task, the source variable $\vec{X}$ represents "*tank present*" or "*jeep present*" states with equal probability. The scene $\vec{Y}$ is of dimension 80×80 pixels ($M = 80$) and the object can be present at one of the 64 positions ($P = 64$). The number of clutter components is $K = 6$ in our model. In the simulation study, the weight vector $\vec{\beta}$ has mean equal to $\vec{\mu}_{\vec{\beta}} = [160\ 80\ 40\ 40\ 64\ 40]$ and covariance equal to $\mathbf{\Sigma}_{\vec{\beta}} = \vec{\mu}_{\beta}^{T}\mathbf{I}/5$. The clutter to noise ratio c is set to 1. The noise $\vec{N}$ is zero mean with unity covariance matrix $\mathbf{\Sigma}_{\vec{N}} = \mathbf{I}$. We use $160,000$ Monte-Carlo simulations with importance sampling to estimate the $mmse$ for the relevant task. The $mmse$ estimates are numerically integrated over a range of s to obtain the TSI.

(a) (b)

Figure 3. Example scenes with optical blur and noise: (a) Tank in the top of the scene, and (b) Jeep in the middle of the scene.

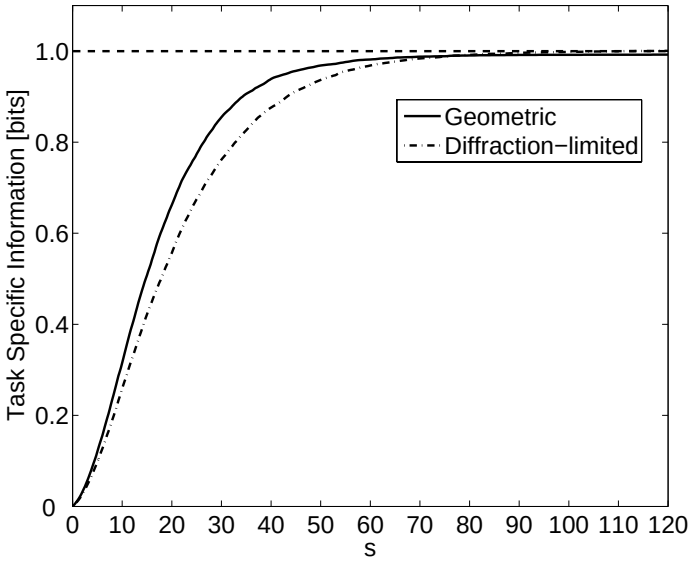
3.1. Geometric Imager and Diffraction-limited Imager

The geometric imager represents an ideal imaging system with no blur, and therefore we set $\mathbf{H}=\mathbf{I}$. Fig. 2 shows some example scenes resulting from object realizations measured in the presence of noise. Note that the object in the scene is either a *tank* or a *jeep* at one of the 64 positions. We model the diffraction-limited imager as a linear shift-invariant system with a point spread function(PSF) that is expressed as: $h_{i,j} = \int_{-\Delta/2}^{\Delta/2}\int_{-\Delta/2}^{\Delta/2} \text{sinc}^2\left(\frac{(x-i\Delta)}{W}\right)\text{sinc}^2\left(\frac{(y-j\Delta)}{W}\right)dxdy$,where Δ is the detector pitch and W quantifies the degree of optical blur associated with the imager. Lexicographic ordering of this two-dimensional PSF yields one row of $\mathbf{H}$ and all other rows are obtained by lexicographically ordering shifted versions of this PSF. The optical blur is set to $W = 2$ and the detector pitch is set to $\Delta = 1$ so that the optical PSF is sampled at the Nyquist rate. Fig. 3 shows examples of images that demonstrate the effects of both optical blur and noise.

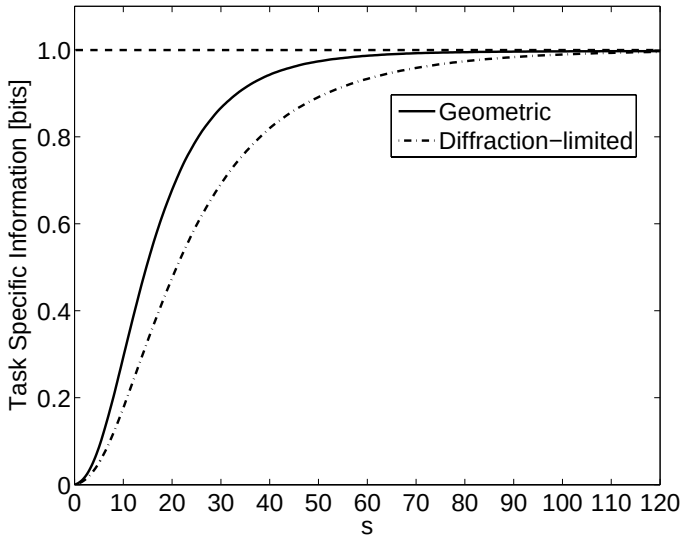
We begin by describing the TSI results for the detection and classification task. Fig. 4(a) shows the plots of TSI versus s for the target detection task. The dashed curve represents the geometric imager and the dash-dot represents the diffraction-limited imager. We observe that the TSI increases with signal to noise ratio, eventually saturating at 1 bit. This result is according to our expectations that (1) TSI increases with increasing signal to noise ratio and (2) TSI is upper bounded by $J(X)$. The TSI metric verifies that imager performance is degraded due to optical blur compared to the geometric imager. For example in the detection task, $s = 34$ yields $TSI = 0.9$ bit for the geometric imager, whereas a higher signal to noise ratio $s = 43$ is required to achieve the same TSI for the diffraction-limited imager. Fig. 4(b) compares the TSI performance of the two imagers for classification task. Again the solid curve represents the geometric imager and dash-dot represents the diffraction-limited imager. Recall that for the classification task we treat the position as the nuisance parameter and so the equi-probable assumption results in a virtual source entropy of 1 bit. As expected the TSI saturates at 1 bit and it gets degraded due to optical blur.

3.2. Projective Imager

For task-specific applications (e.g. detection) an isomorphic measurement (i.e. a pretty picture) may not represent an optimal approach for extracting TSI in the presence of detector noise and a fixed photon budget. A projective imager attempts to directly extract the scene information by measuring linear projections of the scene, while minimizing the number of detector measurements and thereby increasing the measurement signal to noise ratio.[11] The imaging chain is now modified so that the measurement is given by $R = \mathcal{N}(\mathbf{P}(Z))$, where $\mathbf{P}$ is the optical projection operator in matrix form. The imaging model in Eq. (2) and TSI in Eq. (3) can be simply modified by replacing $\mathbf{H}$ with $\mathbf{PH}$. With this motivation we apply the TSI analysis to evaluate the target-detection performance of two projective imagers based on: (a) principal component projections and (b) matched filter projections.



(a)



(b)

Figure 4. TSI versus signal to noise ratio using geometric and diffraction-limited imagers for (a) Detection task, and (b) Classification task.

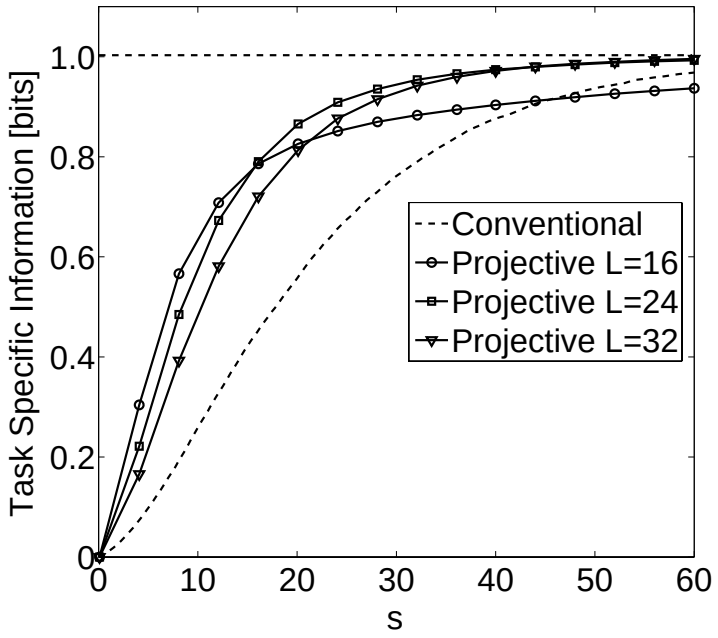


Figure 5. Detection task: TSI versus signal to noise ratio for PC projective imager.

For a set of objects O, the PC projections are defined as the eigenvectors of the object auto-correlation matrix R_{OO} given by: $R_{OO} = \mathbb{E}(oo^T)$, where $o \in O$ is a column vector, a one dimensional lexicographic representation of a two-dimensional object. In our simulation study, we use $10,000$ such object realizations to estimate R_{OO}. The projection matrix $\mathbf{P}^*$ consists of the L dominant eigenvectors of R_{OO} of length $M^2 = 6400$ arranged as rows. To ensure a fair comparison of the projective imager with the conventional imager, we constrain the total number of photons used by the former to be less than or equal to the total number photons used by the latter. We normalize $\mathbf{P}^*$ to enforce this photon constraint resulting in the projection matrix $\mathbf{P} = \frac{1}{cs}\mathbf{P}^*$, where $cs = \max_j \sum_{i=1}^{F} |\mathbf{P}^*|_{ij}$.

Fig. 5 shows the TSI for this projective imager plotted as a function of s for the detection task. Note that the TSI for this projective imager increases as the number of PC projections L is increased from 16 to 32. This can be attributed to the reduction in truncation error associated with increasing L. However, there is also an associated signal to noise ratio cost with increasing L as we distribute the fixed photon budget across more measurements while the detector noise variance remains fixed. This effect is illustrated by the case $L = 24$ where the TSI begins to deteriorate. Notwithstanding this effect, the PC projective imager provides an improved task-specific performance compared to the conventional imager, especially at low signal to noise ratio. For example, the PC projective imager with $L = 24$ achieves a $TSI = 0.9$ bit at $s = 18$; whereas, the conventional imager requires $s = 34$ to achieve the same TSI performance. Although we have shown that the PC projective imager provides larger TSI than the conventional imager we cannot claim that the PC projections are an optimal choice.

4. CONCLUSIONS

Objective metrics that measure imagery are often used to accomplish some task. In these cases task-specific performance is crucial. The task-specific information content of the image measurement can serve as a measure of imaging system performance in such a case. In this paper, we have proposed a framework for the definition of TSI in terms of the well known Shannon mutual information measure. The TSI data obtained from the simulation study confirm our intuition about the performance of the conventional and projective imaging systems. Note that TSI may serve as an upper bound on the performance of any algorithm that

attempts to extract task-specific information from the measurement data. We also note that TSI can also be used to optimize imaging system design for maximizing the task-specific performance/information: an area of ongoing research in our group.

REFERENCES

1. R. E. B. J. A. O'Sullivan and D. L. Snyder, "Information-theoretic image formation," *IEEE Trans. on Image Processing* **44**, pp. 2094–2123, 1998.
2. A. Ortega and K. Ramchandran, "Rate-distortion methods for image and video compression," *IEEE Signal Processing Magazine* **15**, pp. 23–50, 1998.
3. A. Ashok and M. A. Neifeld, "Information-based analysis of simple incoherent imaging systems," *Opt. Express* **11**, pp. 2153–2162, 2003.
4. C. L. F. F. O. Huck and Z. Rahman, "An information theory of visual communication," *Phil. Trans. R. Soc. A: Phys. Sci. and Engr.* **354**, pp. 2193–2248, 1996.
5. F. O. Huck and C. L. Fales, "Information-theoretic assessment of sampled imaging systems," *Optical Engineering* **38**, pp. 742–762, 1999.
6. N. F. S. P. Awate, T. Tasdizen and R. T. Whitaker, "Adaptive, nonparametric markov modeling for unsupervised, mri brain-tissue classification," *Med. Image. Anal. (to be published)* .
7. J. Liu and P. Moulin, "Information-theoretic analysis of interscale and intrascale dependencies between image wavelet coefficients," *IEEE Trans. on Image Processing* **10**, pp. 1647–1658, 2001.
8. L. Zhen and Karam, "Mutual information-based analysis of jpeg2000 contexts," *IEEE Trans. on Image Processing* **14**, pp. 411–422, 2005.
9. S. S. D. Guo and S. Verdu, "Mutual information and minimum mean-square error in gaussian channels," *IEEE Trans. on Inform. Theory* **51**, pp. 1261–1282, 2005.
10. D. P. Palomar and S. Verdu, "Gradient of mutual information in linear vector gaussian channels," *IEEE Trans. on Inform. Theory* **52**, pp. 141–154, 2006.
11. M. A. Neifeld and P. Shankar, "Feature-specific imaging," *Applied Optics* **42**, pp. 3379–3389, 2003.