Task Specific Information

Amit Ashok[1], Pawan K. Baheti[1] and Mark A. Neifeld[1,2]

Department of Electrical and Computer Engineering[1] College of Optical Sciences[2]
1230 E. Speedway Blvd., University of Arizona, Tucson, AZ 85721
Email: {ashoka,baheti,neifeld}@ece.arizona.edu

Abstract: We introduce the notion of task-specific information (TSI) to quantify the performance of imaging systems. We demonstrate the utility of TSI for evaluating the performance of conventional and projective imagers for a detection task.

OCIS codes: *110.3000,110.4280,200.3050.*

1. Introduction

The information content of an image plays an important role in a wide array of applications ranging from video compression to imaging system design[1, 2]. However, the computation of image information content remains a challenging problem. For computational tasks such as 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). In this paper, we describe a formal approach to the computation of TSI and use it to describe/quantify the task-specific performance of conventional and projective imaging systems for a target detection task.

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 $F = \mathcal{H}(Y)$. The image F is projected by the projection operator $\mathcal{P}$ to yield the measurement $Z = \mathcal{P}(F)$. 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 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, even when using methods such as importance-sampling and/or Markov Chain Monte Carlo(MCMC) prove to be computationally prohibitive.

Recently, Guo et. al [3,4] have demonstrated a direct relationship between the minimum mean square error ($mmse$) in estimating X from R, and the mutual-information $I(X;R)$ for a linear additive Gaussian channel, expressed as: $\vec{R} = \sqrt{s}\mathbf{PHC}(\vec{X}) + \vec{N}$, where $\mathcal{H}(\cdot) = \mathbf{H}(\cdot)$, $\mathcal{P}(\cdot) = \mathbf{P}(\cdot)$ and $\mathbf{H}$, $\mathbf{P}$ denote the matrix channel operator and the matrix projection operator respectively. Note that $\mathcal{C}(X)$ is a random function of X. N denotes zero mean addtive Gaussian noise with covariance $\mathbf{\Sigma}_{\vec{N}}$. The mutual-information $I(\vec{X};\vec{R})$ can be expressed as [3, 4]

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

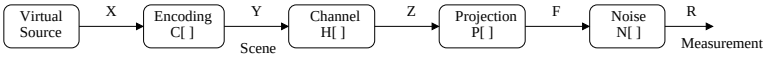
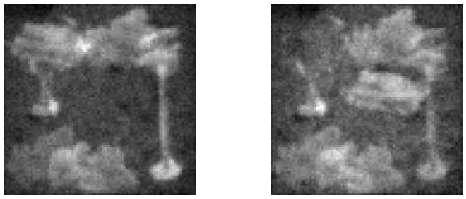


Fig. 1. Block diagram of an imaging chain.

Now let us consider 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: $\vec{R} = \mathbf{PHC_{det}}(X) + \vec{N}$, 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



(a) (b)

Fig. 2. Example scenes from conventional imager.

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 $\vec{N}_c = \sqrt{c}\mathbf{PHV_c}\vec{\beta} + \vec{N}$ with mean $\vec{\mu}_{\vec{N}_c} = \vec{\mu}_{\vec{\beta}}$ and covariance $\mathbf{\Sigma}_{\vec{N}_c} = \mathbf{PHV_c\Sigma}_{\vec{\beta}}\mathbf{V_c}^T\mathbf{H}^T\mathbf{P}^T \cdot c + \mathbf{\Sigma}_{\vec{N}}$. Now, we can rewrite the imaging model as: $\vec{R} = \sqrt{s}\mathbf{PH}\vec{Y} + \vec{N}_c$ 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

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

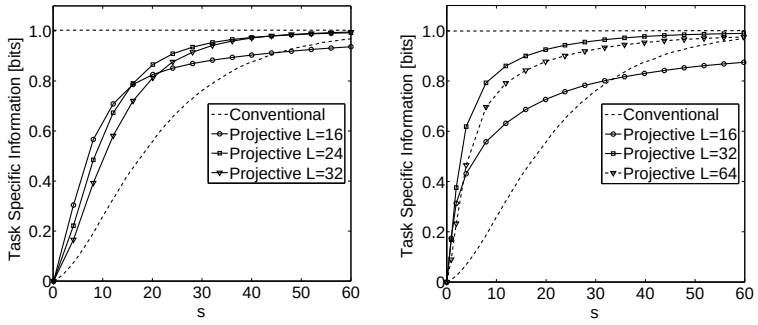
3. Conventional and Projective Imagers

We begin by describing the source, object and clutter components of the scene mode. The binary source variable X can take values 0 or 1 with equal probability. The scene $\vec{Y}$ is of dimension 80×80 pixels ($M = 80$) and the object can be present at one of 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 $\vec{\mu}_{\vec{\beta}} = [160\ 80\ 40\ 40\ 64\ 40]$ and covariance to $\mathbf{\Sigma}_{\vec{\beta}} = \vec{\mu}_{\vec{\beta}}^T\mathbf{I}/5$. 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 detection task. The *mmse* estimates are numerically integrated to obtain TSI from Eq. 2 over a range of s.

We model the conventional imager as a linear shift-invariant system with a diffraction-limited point spread function(PSF) that is expressed as: $h_{i,j} = \int_{-\Delta/2}^{\Delta/2}\int_{-\Delta/2}^{\Delta/2}\mathrm{sinc}^2\left(\frac{(x-i\Delta)}{W}\right)\mathrm{sinc}^2\left(\frac{(y-j\Delta)}{W}\right)dxdy$,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. We set $\mathbf{P} = \mathbf{I}$ for coventional imager. Fig. 2 shows examples of images that demonstrate the effects of both optical blur and noise. Fig. 3(a) shows the plot of TSI versus s for the conventional imager in dash-dot curve for the target detection task. 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 $J(X)$. Although these general trends were known in advance of our analysis, we are encouraged by our ability to quantify these trends using a formal approach.

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. 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.

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



(a) (b)

Fig. 3. Detection task: TSI versus signal to noise ratio (a) PC projective imager and (b) MF projective imager.

$\mathbf{P}^*$ to enforce this photon constraint resulting in the projection matric $\mathbf{P} = \frac{1}{cs}\mathbf{P}^*$, where $cs = \max_j \sum_{i=1}^{F} |\mathbf{P}^*|_{ij}$.

Fig. 3(a) 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 $L = 32$ 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.

In fact, for a detection problem it is well known that the generalized matched filter (MF) approach is optimal in terms of the Neyman-Pearson criterion. Recall that in our target detection problem the target position is a nuisance parameter that must be estimated implicitly. In such a case, instead of a matched filter (e.g. correlator) we consider a set of matched projections. Each matched projection corresponds to the target at a given position. We define the projection matrix $\mathbf{P}$ as $\mathbf{P} = \bar{\mathbf{T}}\mathbf{\Sigma}_{\vec{N}_c}^{-1}$, where $\bar{\mathbf{T}}$ is the modified target profile matrix with each row corresponding to a target profile at a specific position. The number of positions chosen is L yielding a $L \times M^2$ $\bar{\mathbf{T}}$ matrix. The target positions for constructing $\bar{\mathbf{T}}$ are chosen such that they are equally spaced with some overlap between the profiles at the adjacent positions. The target profile matrix $\bar{\mathbf{T}}$ is post-multiplied by $\mathbf{\Sigma}_{\vec{N}_c}^{-1}$ to account for the effects of colored noise $\vec{N}_c$ resulting in $\mathbf{P}$ of size $L \times M^2$. Similar to PC projections, the MF projection matrix $\mathbf{P}$ is normalized to enforce the photon constraint. Note that $L = 64$ is the maximum number of projections, as the number of target positions is limited to 64. Fig. 3(b) shows the plot of TSI versus s for the MF projective imager with $L = 16, 32,$ and 64. As before, we see that TSI increases with number of projections L. However, at $L = 32$ the TSI shows the rollover effect due to the signal to noise ratio cost associated with increasing L. Ideally, we expect that the maximum TSI is obtained for $L = 64$, as it includes all possible target positions. However, there is some overlap between the target profiles at adjacent positions and so $L \leq 64$ projections are sufficient to extract the detection-task related information. As expected, the MF projection imager yields better performance compared to the PC projection imager. For example, to achieve $TSI = 0.9$ bit the MF projections with $L = 32$ requires $s = 17$ compared to $s = 23$ for the PC projections.

4. Conclusions

The TSI results obtained from the simulation study confirm our intuition about the performance of the conventional and projective imaging systems. We note that TSI can also be used to optimize imaging system design for maximizing the task-specific performance/information. We will present the results of TSI optimized imagers at the COSI presentation.

References

[1] A. Ortega and K. Ramchandran, "Rate-distortion methods for image and video compression," IEEE Sig. Proc. Magazine **15**, 23-50 (1998).

[2] A. Ashok and M. A. Neifeld, "Information-based analysis of simple incoherent imaging systems," Opt. Express **11**, 2153-2162 (2003).

[3] D. Guo, S. Shamai and S. Verdu, "Mutual information and minimum mean-square error in Gaussian channels," IEEE Trans. on Inform. Theory **51**, 1261-1282 (2005).

[4] D. P. Palomar and S. Verdu, "Gradient of mutual information in linear vector Gaussian channels," IEEE Trans. on Inform. Theory **52**, 141-154 (2006).