IEEE Transactions on Nuclear Science, Vol. 36, No. 1, February 1989

A NOVEL METHOD FOR ASSESSING POSITION-SENSITIVE DETECTOR PERFORMANCE

N.H. Clinthorne, W.L. Rogers, L. Shao, A.O. Hero III, and K.F. Koral
Division of Nuclear Medicine, Bioengineering Program, and
Department of Electrical Engineering and Computer Science,
University of Michigan, Ann Arbor, MI 48109-0552

Abstract

A marked point process model of a position-sensitive detector is developed which includes the effects of detector efficiency, spatial response, energy response, and source statistics. The average mutual information between the incident distribution of γ-rays and the detector response is derived and used as a performance index for detector optimization. A brief example is presented which uses this figure-of-merit for optimization of light guide dimensions for a modular scintillation camera.

Introduction

The full-width at half-maximum (fwhm) of the point response function (prf) of a position-sensitive detector or camera has been a hallmark of sorts in performance assessment. However, to be a useful measure of resolution, it requires spatial linearity and a point response function shape invariant of location. Furthermore, the fwhm can be of marginal value in assessing requisite design tradeoffs or examining performance with respect to an optimal estimation rule or "resolution recovery" scheme.

Perhaps the best method to assess performance would be to define a measure of distortion or fidelity that has a meaning of accuracy in light of a specific task. With the complete transfer characteristics of the detector in hand, we could analyze the detector performance relative to this measure of accuracy. This is generally a difficult approach. Cameras are usually required to perform multiple tasks, many of which have poorly specified distortion criteria. Additionally, performance evaluation of the optimal estimation rule with respect to a general distortion measure may itself be an onerous task. To avoid these problems, more diffuse performance criteria are often used—mean-squared error, for example. The premise is that optimization with respect to one of these more tractable measures may optimize performance with respect to a task of interest.

The path taken here is somewhat more direct. We make use of a fundamental quantity representing the average information conveyed from a photon or particle distribution incident on a camera to our observations. This quantity is known as the average mutual information and provides a lower bound on the achievable performance with respect to any distortion measure.

This brief note will develop a marked-point process model of a camera, associated average mutual information and conclude with a practical example in which we employ these methods to assess the performance of a scintillation detector for tomographic imaging with respect to variations in light guide dimensions.

Background

The average of the mutual information between two random variables, X and Y, is defined as:

$$I(X;Y) = \int_{\mathcal{X}\times\mathcal{Y}} dP_{XY}(x,y) \ln \frac{dP_{XY}(x,y)}{dP_X(x)dP_Y(y)} \qquad (1)$$

where $P_X(x)$ denotes the probability distribution of the random variable X. The differential $dP_X(x)$ reduces to a probability density function (pdf), if one exists, or a probability mass function if X assumes discrete values. In classical Shannon terminology, $dP_X(x)$ represents the source of information and $dP_{XY}(x,y)/dP_Y(y)$ is the channel through which the information is relayed.

A particularly revealing decomposition of $I(X;Y)$ is obtained by rearranging (1) and applying Bayes' rule,

$$I(X;Y) = H(X) - H(X|Y). \qquad (2)$$

Here $H(X)$ is the average *a priori* uncertainty of the random variable X, or the source entropy,

$$H(X) = -\int_{\mathcal{X}} dP_X(x) \ln dP_X(x). \qquad (3)$$

The conditional entropy, $H(X|Y)$, is often called the *equivocation* and is the average remaining uncertainty in our knowledge of X once Y has been specified. The equivocation is defined as,

$$H(X|Y) = -\int_{\mathcal{X}\times\mathcal{Y}} dP_{XY}(x,y) \ln dP_{X|Y}(x|y). \qquad (4)$$

If Y and X are independent, the value of Y tells us nothing about X and $H(Y|X) = H(X)$—the average mutual information reaches its minimum of zero. On the other hand if knowledge of Y is entirely equivalent to knowing X then $H(X|Y) = 0$ and the mutual information attains its maximum, $H(X)$.

Average mutual information is a measure of how well we can reconcile the value of one random variable given the observation of another, hopefully related, random variable. Specifically, $I(X;Y)$ can be related to a lower bound on the accuracy to which we can estimate X with respect to a wide class of distortion criteria. For example, if the distortion measure is mean-squared error, we have the bound,

$$\overline{D^2} \geq \frac{1}{2\pi e} e^{2H(X)} e^{-2\max I(X;Y)}, \qquad (5)$$

where the maximization is performed over all source distributions, $P_X(x)$. Similar bounds can be derived for other distortion measures.[1]

0018-9499/89/0200-1127$01.00 © 1989 IEEE

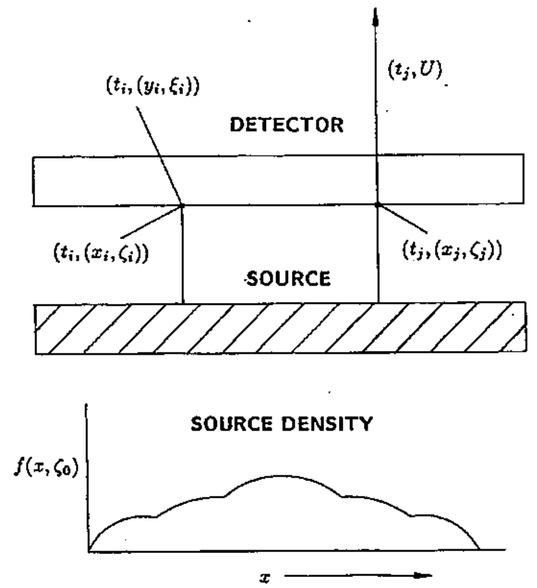

Figure 1: Marked point process camera model. Quantum incident at time t_i is detected while incident quantum at time t_j is not.

Camera Model

In what follows we will develop a marked point-process model for the observation of a photon or particle distribution incident on a camera. This approach is similar in spirit to the model developed for SPECT aperture design.[1]

Consider the attributes we wish to estimate from a sequence of events incident on a detector — usually the location and energy of each event (Figure 1). We assume that the arrival times of these quanta during the observation interval are governed by an inhomogeneous Poisson counting process, N, with instantaneous rate, $\lambda(t)$. An n-point realization of this process on the observation interval $[0,T]$ is given as the set of arrival times, $\{t_1,\ldots,t_n\}$. With each arrival time we associate the attributes of the incident radiation and form a marked point process, $(N,(X,Z))$, with corresponding realization, $\{(t_1,(x_1,\zeta_1)),\ldots,(t_n,(x_n,\zeta_n))\}$, where x_i is the location of the ith quantum and ζ_i its energy. We assume that x_i and ζ_i are drawn from the mark space, $\mathcal{X}\times\mathcal{Z}$ according to a known source pdf, $f(x,\zeta)$, which is independent of time during our observation.

This incident point process excites an observation process, (N,W), defined as follows. If the incident photon or particle is detected we associate with its detection time a collection of attributes from which we can estimate its position and energy. If the quantum fails to be detected, we simply annotate the corresponding arrival time with this fact. Accordingly, a realization of our observation process takes the form, $\{(t_1,w_1),\ldots,(t_n,w_n)\}$ where the mark w_i is,

$$w_i=\begin{cases}(y_i,\xi_i) & i\text{th quantum detected}\\ U & \text{otherwise}\end{cases}$$

where y_i and ξ_i are the estimates of position and energy respectively, and U is the mark denoting failure of detection. Conditioned on N the w_i are assumed to be independent, identically-distributed (i.i.d.) random variables in the observation interval. Knowledge of the quanta that fail to be detected is information that is not usually available from our observations; however, it allows the development of an upper bound on the true average mutual information which takes a relatively simple form.

The average mutual information between the observation process, (N,W), and the attributes of location and energy in the incident process is

$$I(((X,Z),n);(N,W))=\mathrm{E}\left[\ln\frac{dP(W|(X,Z),N)}{dP(W|N)dP(n)}\right], \quad (6)$$

where the expectation is taken with respect to the joint statistics, $dP(W,(X,Z),N)$.

Due to the fact that the w_i are conditionally independent given N we write:

$$dP(W|(X,Z),N)=\prod_{i=1}^{n}dP(w_i|(x_i,\zeta_i)). \quad (7)$$

$$dP(w_i|(x_i,\zeta_i))=\begin{cases}f(y_i,\xi_i|x_i,\zeta_i)(1-P(U|x_i,\zeta_i)) & \text{detected}\\ P(U|x_i,\zeta_i) & \text{otherwise}\end{cases} \quad (8)$$

In (8), $f(y_i,\xi_i|x_i,\zeta_i)$ is nothing more than the response of the camera to a photon detected at location x_i with energy, ζ_i, and normalized to unit area. As such it represents the normalized point-energy impulse response of the camera. $P(U|(x_i,\zeta_i))$ is the probability that an incident quantum fails to be detected.

Again, using the independence of the w_i, we have,

$$dP(W|N)=\prod_{i=1}^{n}dP(w_i). \quad (9)$$

From the Poisson assumption on N:

$$dP(n)=\frac{\Lambda^n e^{-\Lambda}}{n!}. \quad (10)$$

Substituting (7), (9) and (10) into (6), and noting that the w_i are i.i.d., we obtain the result:

$$I(((X,Z),n);(W,N))=\Lambda I((X,Z);W)+H(n), \quad (11)$$

where Λ is the expected number of incident quanta during the observation interval, $[0,T]$ and $H(n)$ is the entropy of a Poisson random variable with rate Λ.

Since Λ and $H(n)$ are independent of the camera parameters, we focus on the rate independent average mutual information,

$$\begin{aligned}I((X,Z);W) &= \mathrm{E}\left[(1-P(U|x,\zeta))\ln\frac{f(y,\xi|x,\zeta)}{f(y,\xi)}\right]\\ &+ \mathrm{E}\left[\ln\frac{P(U|x,\zeta)}{P(U)}\right]. \quad (12)\end{aligned}$$

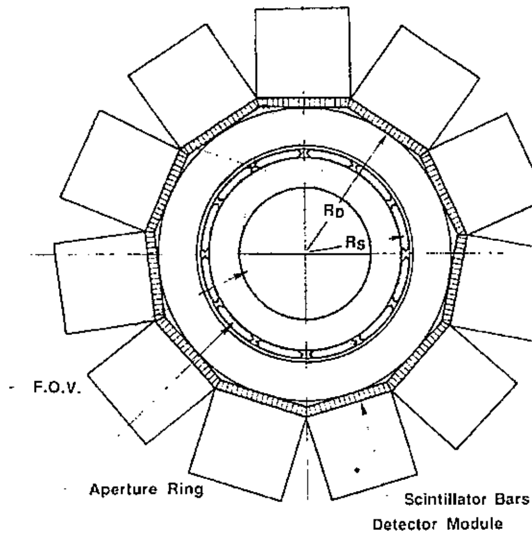

Figure 2: SPRINT II tomograph. The eleven modular detectors are arranged around the periphery.

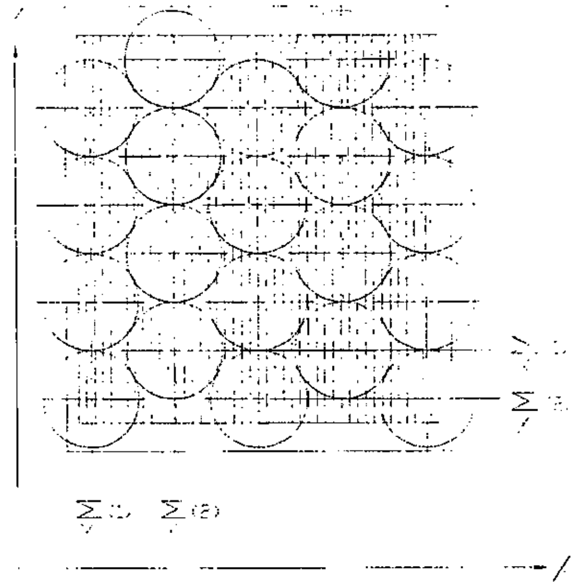

Figure 3: Pictorial diagram of the modular detector showing the individual NaI(Tl) scintillator bars and the photomultiplier arrangment. The light guide is between the detector and the PMTs.

The expectation in the first term is taken with respect to $f(y,\xi|x,\zeta)f(x,\zeta)$ and the second with respect to $P(U|x,\zeta)f(x,\zeta)$. Inspection of (12) reveals that the camera prf enters only into the first term. If changes to the detector are made which do not modify its detection efficiency, maximizing (12) corresponds to maximizing its first term with respect to the point-energy impulse response. It is also interesting to note that if the detection probability is constant over the region where the source density is non-zero, the second term vanishes.

The maximization of the average mutual information often coincides with what we intuitivley feel should be high-information. As an example, consider only the position dependence of (12) and fix energy response and detection efficiency. For gaussian, spatially invariant prfs, and a gaussian or extended source density, maximizing (12) is equivalent to minimizing the fwhm of the point response function.[1]

Applications

We applied average mutual information as a tool for optimizing the modular detector performance of SPRINT II—a second generation ring geometry tomograph for SPECT.[2] The arrangement of the eleven camera modules in this instrument is shown in Figure 2. Each modular detector in this unit is comprised of 45 discrete bars of NaI(Tl) scintillator, optically coupled to a light guide setup whose thickness must be chosen to yield optimum performance. A 4×5 hex-packed array of PMTs is affixed to the light guide and their outputs are routed to a position computer which calculates maximum-likelihood position estimates in real-time.[3] A pictorial diagram of this detector is shown in Figure 3.

Because of the geometry of SPRINT II, detector performance requirements are critical in the azimuthal (x) direction. An error in position in this direction on the detector results in a 2:1 error at the center of object space. By varying the light guide thickness, scintillation light can be redistributed to the PMTs altering the both the extrema of spatial resolution and its distribution across the detector. The choice of optimal thickness was hindered by lack of a suitable performance measure. The fwhm of the point response function was a poor choice because of the discrete nature of the detector and matters were further complicated by the wide variation in prf shapes with position. We applied the average mutual information measure as an aid in determining the optimal light guide thickness.

To simplify the application, our analysis focused entirely on the azimuthal detector performance. We justified this by noting that in the axial direction, performance requirements are not stringent therefore slight alterations in light guide thickness are of little consequence. It was also assumed that varying the light guide thickness would not affect either the detection efficiency or the energy response significantly, hence the second term of (12) was ignored for intercomparisons. The first term was approximated discretely by,

$$I(X;Y) = \sum_i \sum_j P(y_j|x_i)P(x_i)\log_2 \frac{P(y_i|x_j)}{\sum_k P(y_j|x_k)P(x_k)}, \quad (13)$$

where $P(x)$ is now a discrete, one-dimensional source distribution, x being the location of incidence, and $P(y|x)$ are the normalized azimuthal prfs of the detector where y is the detector output. The base of the logarithm has been changed to give the resulting information rate in bits / detected photon.

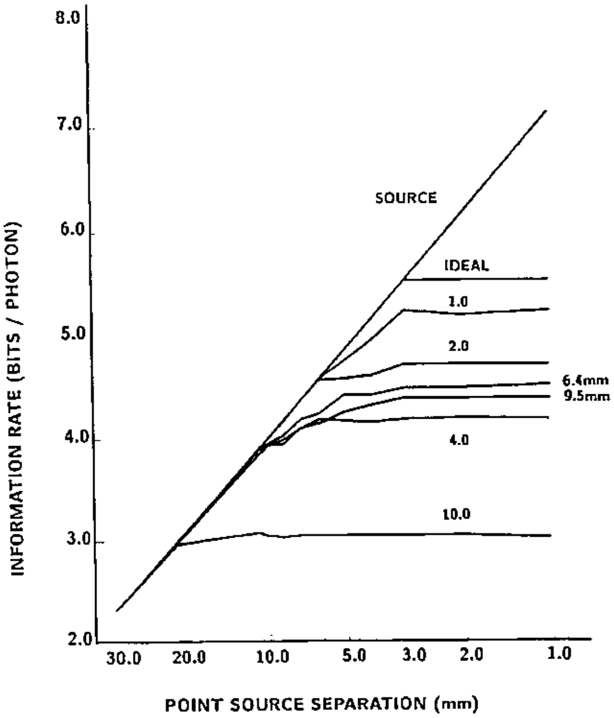

Figure 4: Results using comb function sources with simulated prfs (Ideal, 1.0mm, 2.0mm, 4.0mm, 10.0mm) and the actual detector prfs (6.4mm, 9.5mm). The information rate is in bits/detected photon.

The point response functions for each light guide were estimated by translating a highly collimated line source of ^{99m}Tc over the 140 mm detector width in 1 mm increments and histogramming the detector outputs of 20,000 detected events at each location. This sampling density was deemed sufficient because: the source had a finite width (0.3 mm), the detector does not support a resolution much below 3.0 mm fwhm, and dropping the sampling density by a factor of two did not alter the results.

The evaluation of (13) also required that we choose an appropriate set of source distributions. Since the acquired prfs completely characterized the camera response in terms of our model, the source distributions were computer generated. A particularly interesting class of sources investigated were comb functions, essentially "picket-fences" of point sources in which the spacing between pickets was gradually reduced to 1 mm. At this spacing, the source was indistinguishible from a discrete, uniform source.

To assist in analyzing the results a computer model was implemented which generated simulated point response functions. We employed the follwing model for the modular detector. Photons interacting at any location within a 3 mm bar were assumed to give the same mean position. An underlying gaussian positioning error, whose width could be varied was combined with this response and the result was again quantized in terms of the 3 mm bars.

Results

Two light guide thicknesses were evaluated, 6.4 mm (0.25 in.) and 9.5 mm (0.375 in.). Simulated point response functions were also generated using several values for the underlying gaussian position error including the ideal detector which unequivocally identified the 3 mm bar in which a photon interacted. Results using the comb function sources are shown graphically in Figure 4 where average mutual information in bits / detected photon is plotted against the point source separation of the comb functions for both the simulated and actual detector. In this analysis we assumed that detection efficiency was 100%. For comparison we also plotted the source entropy for each comb source, which upper-bounds the average mutual information.

Several interesting features are immediately discernible. At low source information rates or widely separated point sources, the information transfer through the simulated and actual detectors is essentially lossless. Evaluation of (5) yields a lower bound of zero on the mean-squared error and in the limit of an infinite number of events, these sources could be recovered exactly. (This is only true for discrete source distributions and detection efficiencies of 100%.)

At higher source entropies, corresponding to more point sources at a smaller separation in the field-of-view, some of the curves break away from the source entropy curve with the lowest resolution systems deviating first. At this point information is lost, however, if the operating point is before the curves flatten, the information could conceivably be sent without loss by properly encoding the source distribution.

Where these curves plateau is a fundamental limit known as the *channel capacity*. We verified this by using the iterative Arimoto-Blahut algorithm which estimates the capacity of a discrete channel.[4] Information can not be sent distortion free at a rate greater than the channel capacity and some error in the source estimate is inevitable. The branch of information theory relating information rates and channel capacities to the minimum achievable distortion is known as *rate-distortion theory*.

Results from the light guide evaluations demonstrated that the thinner light guide exhibited a slightly higher information transfer rate. The channel capacities were 4.53 bits / photon and 4.40 bits / photon for the 6.4 mm and 9.5 mm thicknesses respectively. This is in contrast to the capacity for the ideal SPRINT II modular detector which is $\log_2 45 = 5.49$ bits / photon.

Discussion

The marked point process model and average mutual information can be valuable tools in performance evaluation. Like all tools, however, they must be used cautiously. Particularly, our model does not include the count-rate dependent effects that plague many camera systems. Incorporating this into the model, while conceptually not difficult, could lead to intractable expressions for the average mutual information.

The amount of data required to characterize even modest detector systems can be staggering. A two-dimensional camera with 64×64 positions would require at least 4096 64×64 images just to model its spatial response at one energy level. Some parametric modeling of the detector response could potentially decrease this data requirement tremendously.

The source distributions chosen for analysis should reflect actual distributions that will be incident on the camera. The comb functions were chosen here because of their obvious analogy with the true meaning of resolution. However, our results were relatively insensitive to the choice of source distribution. The conclusions drawn above were the same as those resulting from using gaussian sources of various widths and multi-modal rect function sources—the thinner light guide always yielded a slightly higher mutual information. This insensitivity to source may not hold in another application.

The source distribution which yields the channel capacity can be obtained from the Arimoto-Blahut algorithm and represents the source which transfers information most efficiently through the detector system. Knowledge of this distribution could be useful in cases where some latitude exists in how the source is projected onto the detector (γ-ray astronomy, for example). The sensitivity of the average mutual information to small variations in the capacity-matching source is also useful in determining the robustness of a particular detector system.

To realize the potential benefit of increased information transfer, the optimal estimation rule with respect to a specific distortion criterion must be employed. This should always be the final analysis because it is the ultimate benchmark with respect to the task of interest.

Conclusion

We developed a marked point process model of a position-sensitive detector and an expression for the average mutual information between an incident source of photons or particles and the detector's observations of this process. The model includes the spatial and energy responses of the camera and its detection efficiency. We feel that this model, along with the average mutual information measure can be useful in evaluating necessary tradeoffs in camera design and we used it to assist in our choice of light guide dimensions for the SPRINT II modular detectors. While we do not advocate replacing the traditional performance measures such as fwhm, there are many design situations where the traditional measures are inadequate and this technique can be a useful adjunct.

Acknowledgement

This work was supported by the National Cancer Institute, DHHS, under PHS grant CA32856.

References

[1] L. Shao, A.O. Hero III: "Information criteria for SPECT image reconstruction and aperture optimization." Technical Report No. 261, Communications and Signal Processing Laboratory, Dept. of EECS, University of Michigan, Ann Arbor, MI. October 1988.

[2] W.L. Rogers, N.H. Clinthorne, L. Shao, P. Chiao, Y. Ding, J.A. Stamos, K.F. Koral: "SPRINT II: A second generation single photon ring tomograph.", IEEE Trans on Medical Imaging. In press.

[3] N.H. Clinthorne, W.L. Rogers, L. Shao, K.F. Koral: "A hybrid maximum likelihood position computer for scintillation cameras." IEEE Trans Nucl Sci NS-34:97–101, February 1987.

[4] H. Jürgensen: "A note on the Arimoto-Blahut alogorithm for computing the capacity of discrete memoryless channels." IEEE Trans on Information Theory IT-30:376, March 1984.