Optimal Calibration of PET Crystal Position Maps Using Gaussian Mixture Models

Kelly A. Stonger and Michael T. Johnson, *Senior Member, IEEE*

Abstract—A method is developed for estimating optimal PET gamma-ray detector crystal position maps, for arbitrary crystal configurations, based on a binomial distribution model for scintillation photon arrival. The approach is based on maximum likelihood estimation of Gaussian mixture model parameters using crystal position histogram data, with determination of the position map taken from the posterior probability boundaries between mixtures. This leads to minimum probability of error crystal identification under the assumed model.

Index Terms—Crystal position map, Gaussian mixture model, maximum liklihood estimation, positron emission tomography (PET).

I. INTRODUCTION

A common detector configuration for a positron emission tomography (PET) scanner consists of a 2-D array of scintillation crystals coupled to four photomultipliers (PMTs) [1]–[4], as shown in Fig. 1. When a scintillation crystal undergoes gamma interaction, a scintillation photon is generated that is detected by all four PMTs. Each PMT produces a voltage that is used by the detector interface hardware to produce an x and y position, which represents the position of the gamma interaction on the detector crystal face. The hardware then maps the (x, y) coordinate to a specific detector crystal, typically using a lookup table [3], [5]. The position thresholds implemented in this table are typically referred to as a crystal position map (CPM). The goal in determining the CPM is to minimize the probability of error in crystal identification. We undertake to achieve this goal directly using statistical pattern recognition tools for unsupervised clustering to model the underlying distribution of scintillation photon collection in the PMTs.

Under the simple assumption that the probability of scintillation photons reaching a PMT from a particular crystal is constant, the number of photons received can be described by the binomial distribution [2], which approaches a normal distribution as the number of photons becomes large [3]. The number of photons collected by the PMTs along a particular dimension is then given by

$$p(n) = \mathcal{N}(\mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\frac{(n-\mu)^2}{\sigma^2}} \qquad (1)$$

where

$$\mu = Np$$

$$\sigma = \sqrt{Np(1-p)}$$

$$p = \frac{\text{Number of photons collected along one side}}{\text{Total number of photons collected}}.$$

Based on this, a maximum likelihood estimate of the position coordinate can be calculated using the sample mean of the PMT voltages. With a typical PMT arrangement, this leads directly to the simple coordinate equations [5]

$$x = \frac{A+B}{A+B+C+D}$$
$$y = \frac{A+C}{A+B+C+D} \qquad (2)$$

where A, B, C, and D represent the signals from the four individual PMTs.

Since the differential between two parallel axes is also Gaussian, an alternative coordinate expression [3] is

$$x = \frac{(A+B)-(C+D)}{A+B+C+D}$$
$$y = \frac{(A+C)-(B+D)}{A+B+C+D}. \qquad (3)$$

The difference between (2) and (3) is one of dynamic range—in the former case x and y are values between 0 and 1, while in the latter the values are between -1 and 1. Bayesian methods of estimating the position coordinates [6]–[8] may also be used, such as via calibrated lookup tables [6] requiring mapping each (x, y) position to corresponding measured PMT voltage levels.

With ideal hardware, the crystal position map would simply divide the dynamic range of the (x, y) locations into equal areas, one for each crystal in the detector. In reality, scintillation photon scattering effects, manufacturing differences in detectors, and nonlinearities in the system require that the region for each crystal must be uniquely discovered for each detector block.

Typically, a PET system can enter a calibration mode, where a flood phantom or source pin is used to illuminate the detectors while the hardware accumulates counts of gamma-ray events for all (x, y) coordinate pairs on a detector. This data set, as shown in Fig. 2, is called a position histogram, and can be thought of as an image of gamma-ray counts on the detector crystal face.

Several algorithms exist that utilize a position histogram to produce a CPM [3], [9]. Most make only limited use of the statistics provided by the position histogram data. There has also
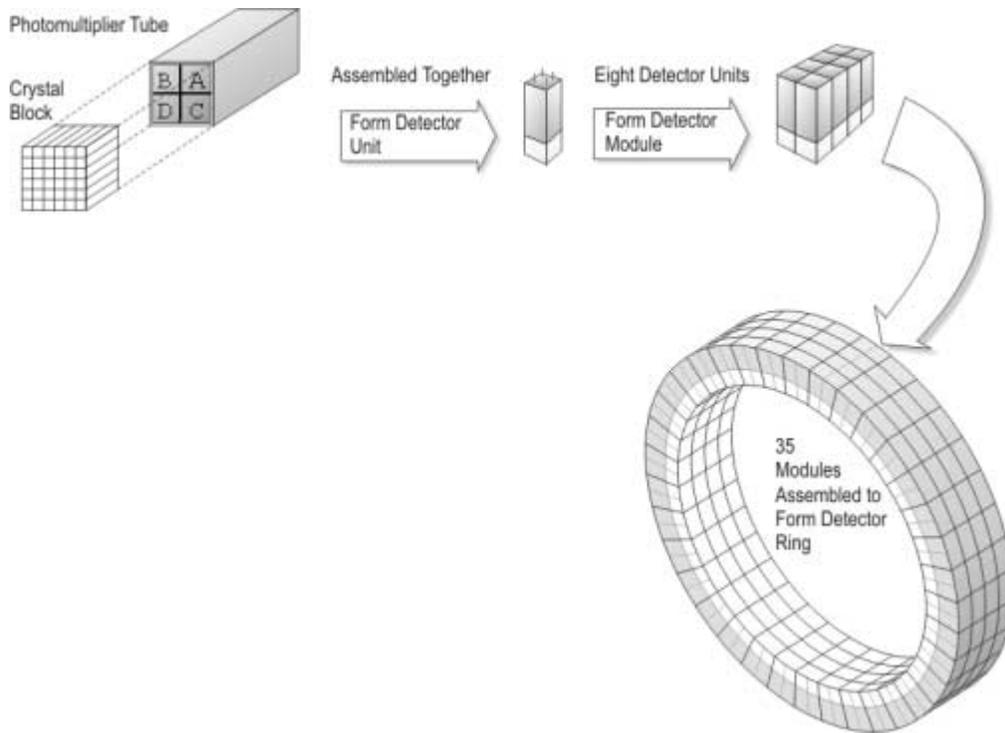
Fig. 1.   PET scanner diagram.
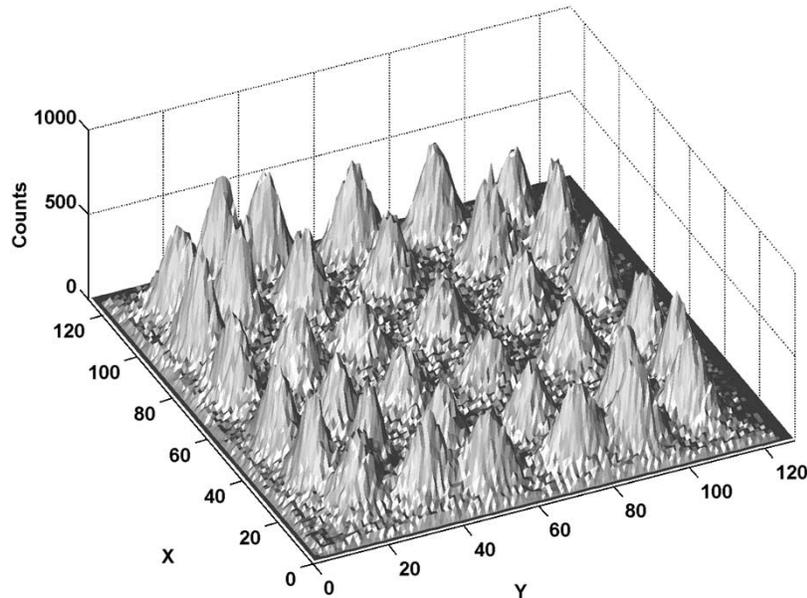
Fig. 2.   An example position histogram with a resolution of 128×128.

been work in designing photodiode implementations to improve crystal identification [7]. As mentioned previously, much of the existing work in this area [1], [2], [6]–[8] has centered on the detector itself and on calibrating the outputs of the four PMTs to produce good spatial resolution and location estimates, rather than focusing on the crystal identification task directly.

The approach introduced here uses a statistical representation to determine the optimal CPM with respect the Gaussian position model outlined above. The method, based on likelihood boundaries of Gaussian mixture models (GMMs), is robust in

that it is simple to implement and can be easily adapted for arbitrary crystal configurations. The resulting CPM has the minimum probability of error possible, given the Gaussian photon distribution model.

The basic GMM will be introduced in Section II, including parameter estimation formulas and maximum *a posteriori* map identification. Section III will illustrate the method on data from an example 6×6 crystal configuration, and quantify the difference between the new method and other algorithms. Conclusions and possibilities for future work are given in Section IV.

II. Gaussian Mixture Models

The binomial distribution of photon collection discussed in Section I leads directly to a two-dimensional Gaussian model for the (x, y) position associated with a particular crystal. For a $j \times k$ block of crystals, the overall distribution of position is then given by a GMM [10], [11]

$$
\begin{aligned}
p(x, y) &\triangleq p(\mathbf{z}) \\
&= \sum_{m=1}^{M} w_m p_m(\mathbf{z}) \\
&= \sum_{m=1}^{M} w_m \mathcal{N}(\boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m)
\end{aligned} \tag{4}
$$

where

$$
\begin{aligned}
\boldsymbol{\mu}_m &= [\mu_{x_m} \quad \mu_{y_m}]^T \\
&= \text{Centroid of crystal } m \text{ distribution} \\
\boldsymbol{\Sigma}_m &= \begin{bmatrix} \sigma_{x_m}^2 & \sigma_{x_m y_m} \\ \sigma_{y_m x_m} & \sigma_{y_m}^2 \end{bmatrix} \\
&= \text{Co} - \text{variance matrix of crystal } m \text{ distribution} \\
w_m &= \text{Weight Prior likelihood of crystal } m \\
M &= jk = \text{Number of crystals.}
\end{aligned} \tag{5}
$$

The GMM is simply a weighted sum of Gaussian distributions. For the crystal mapping problem, the prior likelihoods w_m would be uniform and the two dimensions x and y would be independent, so that the covariance matrix $\boldsymbol{\Sigma}_m$ would be diagonal. However, in reality there are some interactions between the dimensions, particularly near the corners of the map, so that a full covariance matrix is still preferable for this problem.

Parameter estimation for a GMM is commonly done using the expectation maximization algorithm [12], an iterative algorithm leading to locally maximum likelihood estimates

$$
\begin{aligned}
\hat{\mathbf{u}}_m &= \frac{\sum_{i=1}^{N} p(m|\mathbf{z}_i)\mathbf{z}_i}{\sum_{i=1}^{N} p(m|\mathbf{z}_i)} \\
\hat{\boldsymbol{\Sigma}}_m &= \frac{\sum_{k=1}^{N} p(m|\mathbf{z}_i)(\mathbf{z}_i - \hat{\mathbf{u}}_m)(\mathbf{z}_i - \hat{\mathbf{u}}_m)^T}{\sum_{i=1}^{n} p(m|\mathbf{z}_i)} \\
\hat{w}_m &= \frac{1}{N} \sum_{i=1}^{N} p(m|\mathbf{z}_i)
\end{aligned} \tag{6}
$$

where $\mathbf{z_i}$ is the ith data point and $p(m|\mathbf{z_i})$ is the mixture occupancy likelihood

$$
p(m|\mathbf{z}_i) = \frac{p_m(\mathbf{z}_i)\hat{w}_m}{\sum_{k=1}^{M} p_k(\mathbf{z}_i)\hat{w}_k}. \tag{7}
$$

Sums on the right-hand side of the re-estimation formulas are computed using the current estimates of the parameters.

Given a set of histogram data, the above formulas can be easily implemented to estimate the means, variances, and weights of the crystal distributions, given initial estimates for these parameters. For the crystal map task, the estimation algorithm above is fairly robust with respect to initialization, since the number of mixtures is known *a priori* and the mixture distributions are relatively well-separated. However, initialization too far from the histogram peak values can occasionally cause errant results, as discussed further in Section III. Initialization methods range from the simplicity of a uniform grid to peak selection algorithms or clustering algorithms like K-means [13]. For these experiments we use a simple low-pass filter of the histograms and windowed peak selection, described in detail in Section III. The covariance matrices can be initialized fairly arbitrarily; the method used here is to set the initial covariances at about half the distance between crystal centers.

After initialization and estimation using the histogram data as outlined above, the CPM is determined using the equi-likelihood boundaries between the classes. From a statistical pattern recognition perspective, this is equivalent to associating each location in the CPM with the most likely crystal according to the maximum *a posteriori* (MAP) decision rule [13]

$$
\text{crystal} = \underset{m=1...36}{\arg\max} \{p_m(\mathbf{x})w_m\}. \tag{8}
$$

This rule leads to the minimum probability of error decision by identifying the crystal most likely to have generated voltages corresponding to each particular position. If the crystals can be expected to be equally likely, the prior likelihood terms w_m can be dropped from this expression, giving the maximum likelihood (ML) decision rule. The scalar valued functions $(p_m(\mathbf{x})w_m)$ used for classification in (8) are called the discriminant functions for each class, and the equi-value boundaries between the discriminant functions give the classification decision boundaries. For Gaussian distributions, the decision boundaries are quadratic functions of the means and covariance matrices of the crystals.

The algorithm for setting the CPM is summarized as follows:
1) initialize the GMM parameters;
2) run the EM re-estimation procedure given in (6) until parameters converge (typically only a few iterations);
3) identify each point in the CPM with the associated MAP crystal according to (8).

III. Experimental Example

A. Description of Detector System

The system used in this study was a GE Discovery ST PET system. The detectors consist of a 6×6 array of bismuth germanate (BGO) crystals coupled to four PMTs. The position histograms of 280 detector blocks were used as training data for the algorithm.

B. Algorithm Preprocessing and GMM Initialization

The dynamic range of the (X,Y) coordinates of the position histogram produced by the hardware is 256×256. This data set was reduced to 128×128 by summing four adjacent pixels to produce one pixel in the 128×128 position histogram. This reduction has a low pass filtering effect and decreases processing time. Additional smoothing of the histogram, as illustrated in
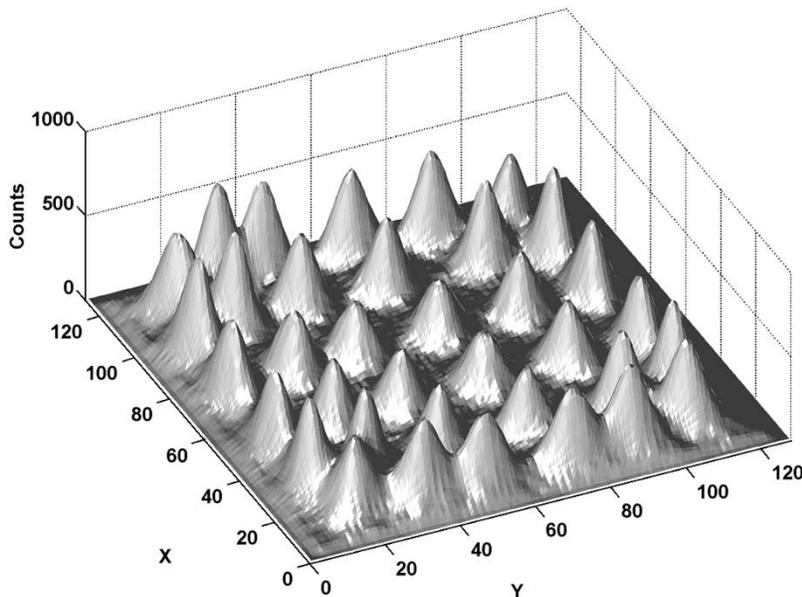
Fig. 3.   Smoothed crystal position histogram.

Fig. 3, is accomplished by a convolution with a simple low-pass filter with kernel:

$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$
$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$
$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$

The initialization procedure used to set the initial GMM means was adopted from part of the current CPM calibration procedure, and is given as follows.

1) For each coordinate in the filtered position histogram with more than a threshold (100) number of counts, determine if it is a local maximum. If it has more counts than any of its 8 immediate neighbors, then save it in a peak candidate array.
2) Find the average counts of the peak candidates and remove any that are less than 25% of the mean, until no longer possible or there are 36 peaks remaining.
3) Remove any peaks that are close (less than 20 pixels) by Euclidean distance, until no longer possible or there are 36 peaks remaining.
4) Remove any peaks that are less than 30% of the mean, until no longer possible or there are 36 peaks remaining.
5) Repeat step 4, incrementing the mean threshold by 5% each time, until there are 36 peaks remaining.

The final 36 coordinate pairs are used as the initial mean values in the GMM. The priors w_m of each mixture are initialized to 1/36, and the covariance matrices are initialized as diagonal with $\sigma_x^2 = \sigma_y^2 = 5$.

To investigate robustness of the algorithm with respect to parameter initialization, the initial mean values were modified by adding white noise over a range of variances.

C. Position Histogram Sampling

Processing time for the algorithm can be substantially reduced by down-sampling the data, since the number of counts in a typical position histogram is much larger than needed for accurate boundary identification. This process can also include an optional noise floor on the histogram.

Both full and down-sampled histograms, with and without noise floor constraints, were investigated, with the resulting maps changing only minimally.

D. Results

This algorithm was implemented in Matlab and trained on the crystal position histograms of 280 detector blocks. A typical GMM crystal map can be seen in Fig. 4, where the lines in the image represent the decision boundaries between the crystals and the GMM mean value locations are shown as points inside the regions. Fig. 5 is a crystal map produced using the current method used by GE for this crystal configuration, with histogram peaks shown as points. Fig. 6 is the pixel by pixel difference between this method and the GMM method. The total pixel difference is 5.3%, which was a typical result of the 280 detectors sampled. Fig. 7 is a crystal map produced by following the valleys between the crystal peaks [3], which are again shown as points. Fig. 8 is the pixel by pixel difference between this valley-tracking method and the GMM method. The total pixel difference in this case is 9.4%.

The pixel differences given above represent an artificially high measure of error, since the boundaries are located where the relative probability is fairly low. An alternative way of estimating the relative error, both within the GMM method and differentially compared to the other methods, is to use the integral of each crystal's GMM mixture distribution outside of its decision region as an estimate of the probability of error. Using this metric, the probability of error for the GMM method is 0.7%. For the method in Fig. 5, the estimated error is 2.8%, and for the method in Fig. 7 the estimated error is 3.6%. Note that the two comparative error estimates are biased in favor of the GMM approach since the metric assumes that the Gaussian location distribution of the GMM is the ideal decision boundary. One
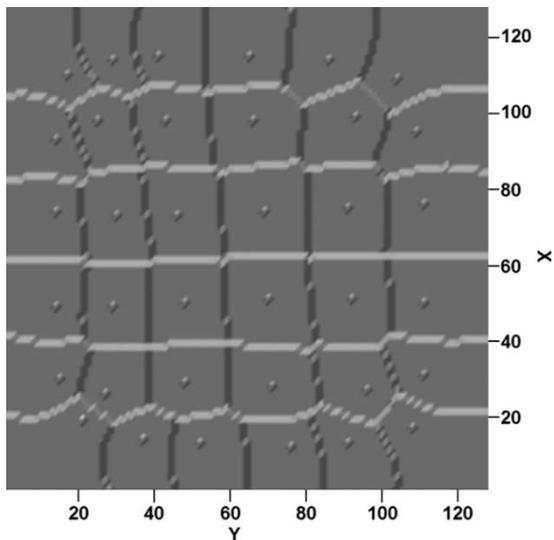
Fig. 4.   GMM CPM method.
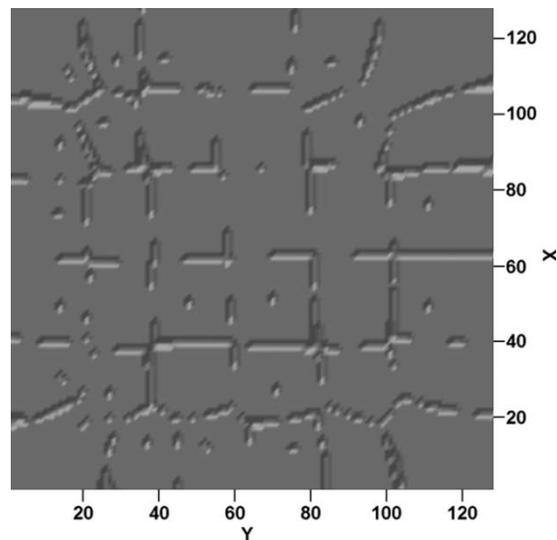
Fig. 6.   Subtraction of GMM and GE methods (5.3% difference).
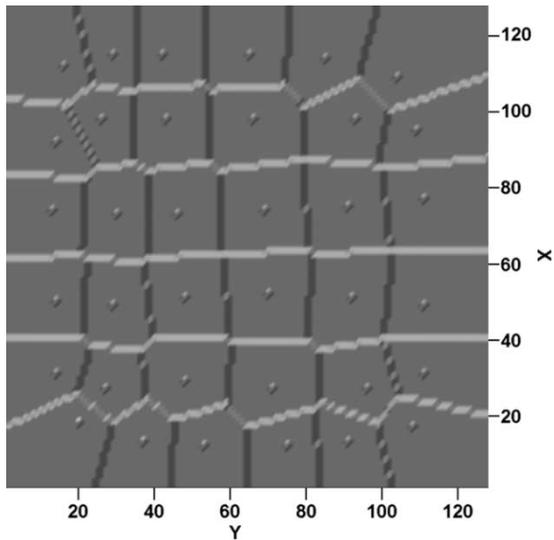
Fig. 5.   Current GE CPM method.
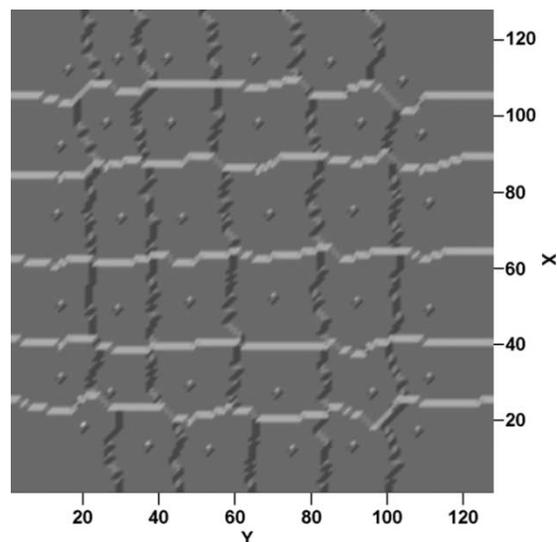
Fig. 7.   Histogram tracking CPM method.

substantial benefit of this error assessment method is that it can estimate and track the expected error of the crystal mappings as it changes over time due to physical system deterioration and drift.

By examining the error of the maps as a function of noise added to the initial mean estimates, it was found that the final mixture locations was relatively unaffected by mean initializations within about 10% of the true histogram peaks, relative to distance between crystal centers. Outside this range, the algorithm demonstrates an increasing likelihood of final map error as a function of the error in the initial values.

IV. Conclusion

An ML method based on GMMs has been presented for constructing crystal position maps of PET gamma-ray detectors. Under the assumption that the underlying crystal probability distribution is Gaussian, this method will produce the minimum probability of error crystal position map.
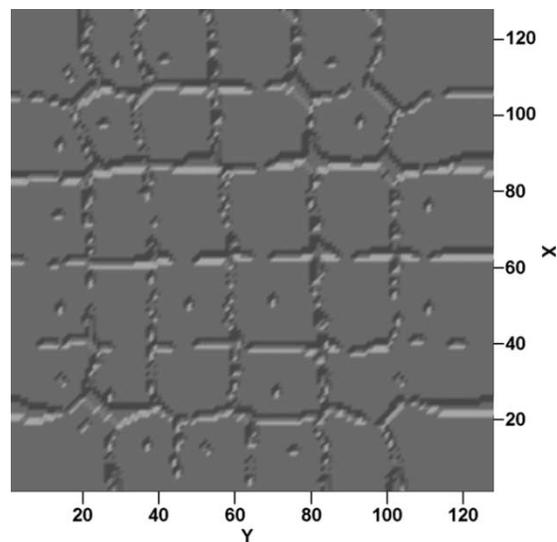
Fig. 8.   Subtraction of GMM and histogram tracking methods (9.4% difference).

In addition to the optimality of the resulting map with respect to crystal identification, benefits of this approach include adaptability and ease of implementation for any possible crystal configuration, robustness with respect to crystal alignment and noise, and the ability to assess expected error of the crystal maps over time.

Future work in this area includes the possibility of improving the underlying model of position distribution by incorporating effects of nonlinear effects such as scattering. Such effects may cause the true distribution to deviate from Gaussian, which could be captured by using mixtures of other elliptically symmetric distributions, for example by using multiple Gaussian mixtures for each crystal, with tied means.

REFERENCES

[1] T. D. Milster, I. A. Selberg, H. H. Barrett, A. L. Landesman, and R. H. Seacat, "Digital postition estimation for the modular scintillation camera," *IEEE Trans. Nucl. Sci.*, vol. NS-32, pp. 748–752, 1985.

[2] M. Casey and R. Nutt, "A multicrystal two dimensional BGO detector system for position emission tomography," *IEEE Trans. Nucl. Sci.*, vol. NS-33, pp. 460–463, 1986.

[3] M. Dahlbom and E. J. Hoffman, "An evaluation of a two-dimensional array detector for high resolution PET," *IEEE Trans. Med. Imaging*, vol. 7, pp. 264–271, 1988.

[4] W. W. Moses, S. E. Derenzo, R. Nutt, W. M. Digby, C. W. Williams, and M. Andreaco, "Performance of a PET detector module utilizing an array of silicon photodiodes to identify the crystal of interaction," *IEEE Trans. Nucl. Sci.*, vol. 40, pp. 1036–1040, 1993.

[5] J. W. Young, J. C. Moyers, and M. Lenox, "FPGA based front-end electronics for a high resolution PET scanner," *IEEE Trans. Nucl. Sci.*, vol. 47, pp. 1676–1680, 2000.

[6] R. A. Mintzer, J. N. Aarscold, N. J. Yasillo, C. E. Ordones, and K. L. Matthews, "Maximum likelihood calibration of small gamma cameras for 511 kdV positron annihilation radiation," in *Proc. IEEE Nuclear Science Symp. Medical Imaging Conf.*, vol. 3, San Francisco, CA, Oct. 1995, pp. 1567–1570.

[7] N. H. Clinthorne, W. L. Rogers, L. Shoa, and K. F. Koral, "A hybrid maximum-likelihood position computer for modular scintillation cameras," *IEEE Trans. Nucl. Sci.*, vol. NS-34, pp. 97–101, 1987.

[8] R. M. Gray and A. Macovski, "Maximum *a posteriori* estimation of position in scintillation cameras," *IEEE Trans. Nucl. Sci.*, vol. NS-23, pp. 849–852, 1976.

[9] J. G. Rogers, R. Nutt, M. Andreaco, and C. W. Williams, "Testing 144 and 256 crystal BGO block detectors," *IEEE Trans. Nucl. Sci.*, vol. 41, pp. 1423–1429, 1994.

[10] G. McLachlan and D. Peel, *Finite Mixture Models*. New York: Wiley, 2000.

[11] M. Agha and D. S. Branker, "Algorithms AS 317: maximum likelihood estimation and goodness-of-fit tests for mixtures of distributions," *Appl. Statist.*, vol. 46, pp. 399–407, 1997.

[12] T. K. Moon, "The expectation-maximization algorithm," *IEEE Signal Processing Mag.*, pp. 47–59, Nov. 1996.

[13] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed. New York: Wiley, 2001.