Optimal Nonlinear Line-of-Flight Estimation in Positron Emission Tomography

Alexander M. Bronstein, Michael M. Bronstein, Michael Zibulevsky, and Yehoshua Y. Zeevi

***Abstract*—The authors consider detection of high-energy photons in positron emission tomography using thick scintillation crystals. Parallax effect and multiple Compton interactions in such crystals significantly reduce the accuracy of conventional detection methods. In order to estimate the photon line of flight based on photomultiplier responses, the authors use asymptotically optimal nonlinear techniques, implemented by feed-forward and radial basis function neural networks. Incorporation of information about angles of incidence of photons significantly improves accuracy of estimation. The proposed estimators are fast enough to perform detection, using conventional computers. Monte Carlo simulation results show that their approach significantly outperforms the conventional Anger algorithm.**

***Index Terms*—Artificial neural network, emission tomography, gamma camera, scintillation detector.**

I. Introduction

DETECTION of high-energy photons emitted as the result of positron decay is one of the most important low-level stages in positron emission tomography (PET) imaging. In this study we consider a detector based on the Anger scintillation camera [1]. Incident high-energy gamma quanta, generated due to positron decay, produce scintillation effect in the crystal. As the result, a shower of low energy photons in the visible and UV spectra is emitted. These photons are collected by an array of photo-multipliers (PMTs), optically coupled to the scintillation crystal, and invoke electric impulses in them. The PMT responses are utilized in estimation of the scintillation point coordinates.

A noncollimated Anger camera, based on thick crystals with high photon penetration depth such as NaI(Tl), is considered in this study. Application of such thick crystals in PET scanners is desirable, due to their low cost and very high light output; they were previously used primarily in gamma ray astronomy [2].

The majority of existing scintillation position estimation algorithms are based on centroid arithmetic, usually combined with correction maps [3]. Their application appears, however, to be problematic in the case of thick crystals due to significant parallax observed at large radiation incidence angles.

Tomitani *et al.* [4] proposed an iterative maximum likelihood algorithm for position estimation and depth encoding in thick scintillation crystals, in order to compensate for the parallax effect. However, an iterative approach necessitate extensive computations that prohibit real-time implementation.

Delorme *et al.* [5] and Clément *et al.* [6] have implemented artificial neural networks in a depth-encoding scintillation detection. Their approach is flexible and offers advantages over iterative algorithms. Still, it does not resolve the problem of multiple Compton interactions, which make the conception of "depth of interaction" ambiguous.

Our study presents a solution for these problems, incorporating side information on the photon incidence angle into the process of position estimation. We use localized, asymptotically optimal, nonlinear estimators, implemented by feed-forward and radial basis functions (RBF) neural networks. As a byproduct, we get accurate position estimation over the entire area of detector including the edges. This is difficult to obtain with centroid arithmetics algorithms. We present a comparison of algorithms on a Monte Carlo simulation and discuss the prospects for practical implementation.

II. Estimation of Scintillation Coordinates

In order to estimate the line of flight, it is usual to estimate coordinates of interaction independently in each detector. The situation becomes more complicated in presence of Compton interactions. Even in the simple case of a single Compton interaction, with consequent photoelectric absorption, one can attempt to estimate the coordinates of both interactions, and choose one with lower penetration depth. A more accurate decision would be to choose the first interaction along the possible line of flight (LOF). This observation demonstrates the importance of prior knowledge of photon direction.

Due to limited light statistics, the coordinate estimation process is not perfect, and one can use the Bayesian framework to optimize it. This would involve knowledge of the distribution of energy deposit and consequent flight directions at every Compton interaction, probability of photoelectric absorption, distribution of visible photons after each interaction, etc. Finally, we would obtain a complicated estimator, which provides the coordinates of the first interaction. The resulting model should include the initial photon direction.

Since this does not seem to be a straightforward problem to manipulate or solve, especially in real time, we propose a learning approach as an alternative. We use a learning approach in order to develop and solve the approximation of the optimal statistical model automatically, using training data, which can be available in large amounts from simulation or from physical experiment. Relatively computationally intensive training process results in fast and accurate on-line estimator.

Manuscript received March 8, 2002. This work was supported in part by the Ollendorff Minerva Center, by the Fund for Promotion of Research at the Technion, and by the Israeli Ministry of Science.

The authors are with the Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa 32000, Israel (e-mail: bron@aluf.technion.ac.il; mbron@tx.technion.ac.il; mzib@ee.technion.ac.il; zeevi@ee.technion.ac.il).

Digital Object Identifier 10.1109/TNS.2003.812434

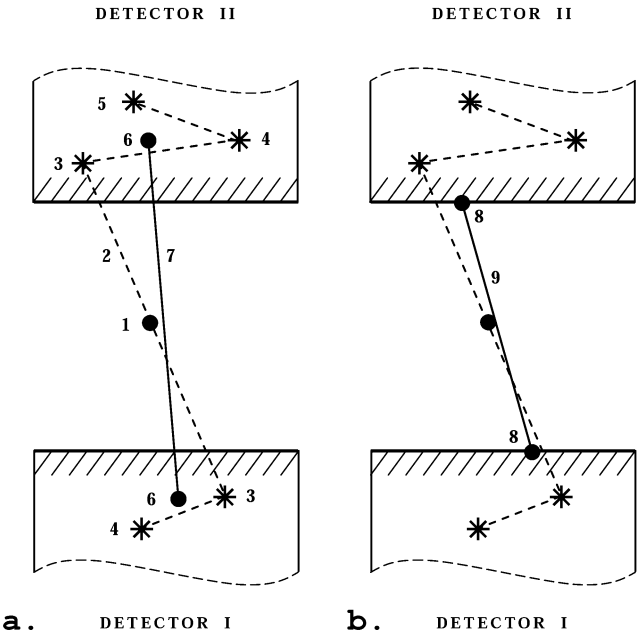

Fig. 1. Scheme of photon coordinate estimation using the Anger algorithm: (a) and ANN-based estimator (b) 1—photon pair emission source, 2—actual LOF, 3—first interaction, 4—secondary interaction, 5—tertiary interaction, 6—scintillation point estimated by Anger algorithm, 7—LOF estimated by Anger algorithm, 8—entrance point estimated by ANN, and 9—LOF estimated by ANN.

Our crucial observation is that using the knowledge of photon direction, one can get a more accurate estimate. In this case, one do not even need to estimate the three–dimensional (3-D) coordinates of each interaction. Instead, the two–dimensional (2-D) coordinate of photon entrance into the detector crystal can be estimated directly. Together with the incidence angle, this gives full description of the LOF.

Such an approach reduces the influence of the parallax effect and multiple Compton interactions, since the point of photon entrance is well defined in this case as well (Fig. 1). Unlike many existing approaches, our method does not rely on the average interaction depth.

III. Parametric Estimation Using Neural Networks

Scintillation detector can be considered to be a complicated nonlinear stochastic system that maps the photon LOF into a vector x of PMT responses. Given the incidence angle, LOF is defined by planar coordinates $\boldsymbol{y} = (y_1, y_2)$ on the surface of the crystal. For every incidence angle, we implement an optimal nonlinear estimator of $\boldsymbol{y}$ of the form, $\hat{\boldsymbol{y}} = \Phi(\boldsymbol{x}; W)$, where $\Phi(\boldsymbol{x}; W)$ is a family of functions, parameterized by the vector of parameters W.

A reasonable criterion for estimator optimality is the expectation of some error function $\mathbb{E}\{\varepsilon(\Phi(\boldsymbol{x}; W) - y)\}$, for example, the expected squared error $\mathbb{E}\left\{||\Phi(\boldsymbol{x}; W) - \boldsymbol{y}||_2^2\right\}$.

We are interested in forms of $\Phi(\boldsymbol{x}; W)$, that possess the property of a universal approximator: when the number of parameters W is large enough, any bounded function $f(\boldsymbol{x})$ can be approximated with given accuracy over a bounded domain by an appropriate choice of W.

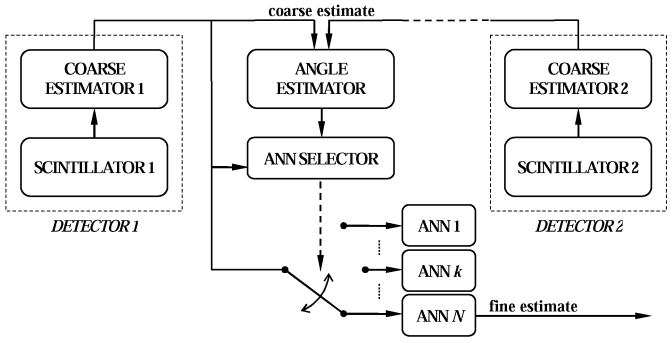

Fig. 2. Block diagram of a practical ANN-based scintillation coordinates estimation algorithm. Estimation of scintillation coordinates in detector 1 using side information from detector 2.

Given the PMT responses to a set of known LOFs $\{\boldsymbol{y}'; \boldsymbol{x}' = f(\boldsymbol{y}') = (x'_1, \ldots, x'_n)\}_{i-1}^{N}$ (referred to as a *training set*), we find such W, that minimizes the mean-squared error (MSE) on the training set, i.e:

$$W^* = \underset{w}{\operatorname{argmin}} \sum_{i=1}^{N} (\Phi(\boldsymbol{x}^i; W) - y')^2.$$

This process is referred to as *training*. When the training set is sufficiently large, the MSE approximates the expected squared error with any desired accuracy. Under such conditions, a universal approximator $\Phi(\boldsymbol{x}; W^*)$ with sufficient parameters approaches the optimal nonlinear estimation. In this study we used two types of universal approximators implemented as artificial neural networks (see Appendix for description).

A. Two-Level Scheme Using Localized Estimators

A possible implementation can be based on a combination of coarse and fine estimators (Fig. 2). Fine estimators, implemented as artificial neural networks, are trained on scintillation events in different (possibly overlapping) regions at a range of calibrated incidence angles. Coarse estimators, based, for example, on the Anger algorithm determine the rough position and incidence angle of the photon. According to this information, the appropriate fine estimator is selected. Such a combination of estimators allows reduction in the size of each network and accelerates the training.

IV. Simulations

In order to test the proposed approach and compare it with other algorithms, we performed a Monte Carlo simulation of ray tracing and gamma quanta interaction in a scintillation detector. The simulation was performed using a slightly modified version of TRIUMF detector modeling platform introduced by Tsang *et al.* [8], [9]. The region of interest and the incidence angle involved in the estimator selection were assumed to be known.

A model of a NaI(Tl) scintillation crystal of size $210 \times 210 \times 45$ mm, separated with a 20 mm glass light guide was simulated. The detector consisted of seven circular PMTs, each of radius 30 mm, with inter-tube gaps of 10 mm. The inter-tube area was assumed to consist of an ideal light-absorbing material (Fig. 3). Fig. 4 depicts the coordinate system of the detector model.

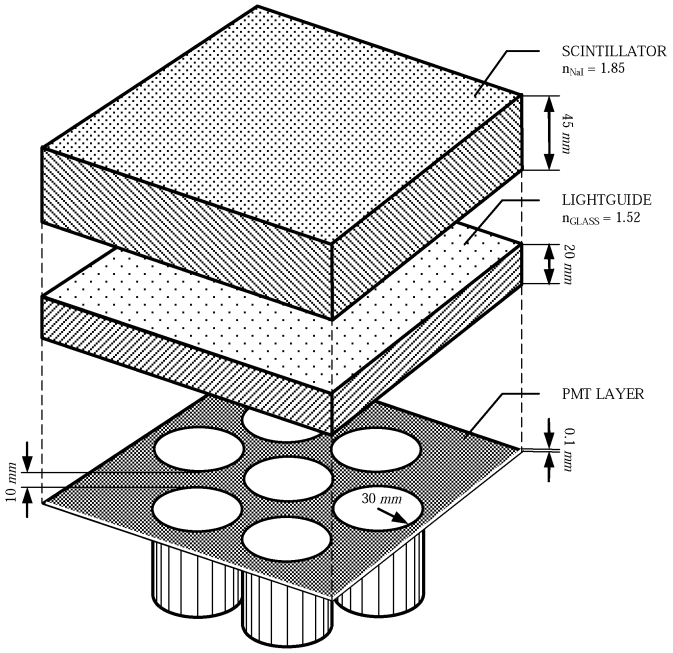

Fig. 3. Detached NaI scintillation detector used in the simulation.

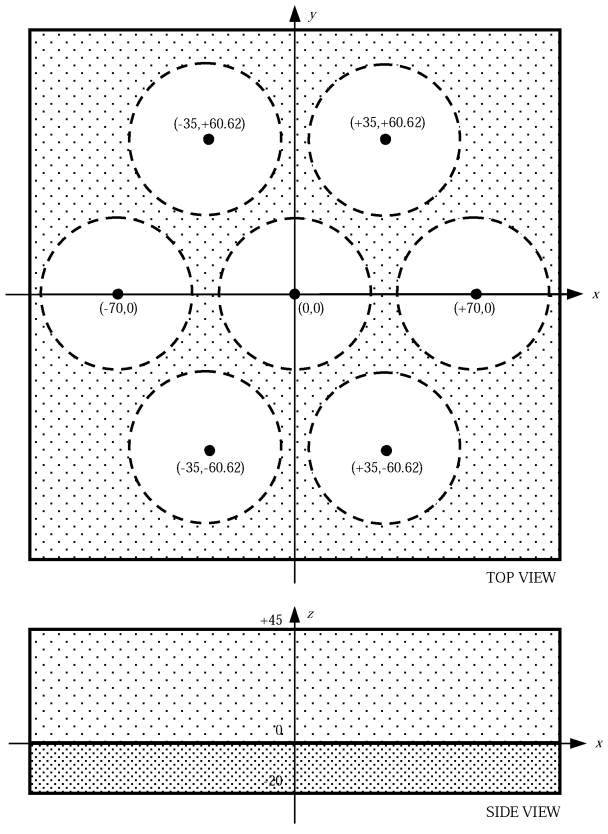

Fig. 4. Detector coordinate system. All coordinates are given in millimeters.

Using TRIUMF, we simulated narrow beams of 512 KeV gamma quanta impinging at the detector crystal surface in a given point and a given incidence angle. The point at which the beam entered the scintillator surface was recorded for each simulated event and served as reference for position estimation. Incidence angle was measured as shown in Fig. 5.

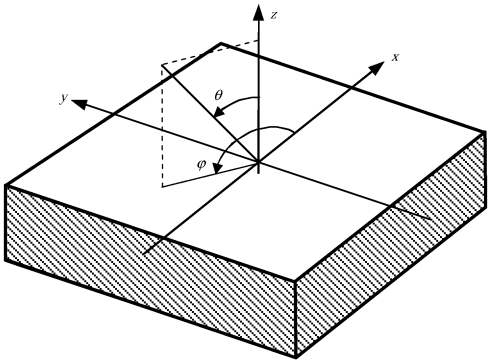

Fig. 5. Incidence angle measurement: normal angle θ and azimuthal angle φ.

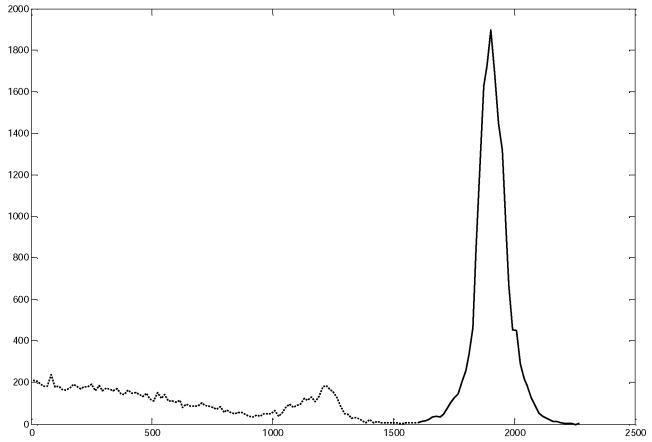

Fig. 6. Estimated pulse spectrum: number of scintillation events versus the number of photoelectrons produced in an event. Events below the photopeak (dotted) were rejected.

PMT responses to the simulated scintillation events, expressed as the number of the produced photoelectrons (assuming quantum efficiency of 25%) were recorded. Pulse spectrum (Fig. 6) was estimated in order to determine the threshold for low-energy events rejection.

Four tests were performed in order to analyze the effectiveness of different scintillation coordinates estimation algorithms. The tests were performed in small regions of the detector. In each test, two sets of data were simulated: a training set and a test set.

The training set consisted of simulated PMT responses to scintillation events resulting from gamma quanta impinging at the crystal surface on an evenly-spaced grid (31×31 points with 1 mm step, 500 gamma quanta per grid point). The test set was used for error estimation and error distribution analysis over the entire tested region and was created on an evenly-spaced 16×16 grid with 2 mm step, 5,000 gamma quanta per grid point.

The following tests were performed:

1) Region: [0 mm, 30 mm]×[0 mm, 30 mm];
   Incidence angle: $\theta = 0°$, $\varphi = 180°$
2) Region: [0 mm, 30 mm]×[0 mm, 30 mm];
   Incidence angle: $\theta = 10°$, $\varphi = 180°$
3) Region: [0 mm, 30 mm]×[0 mm, 30 mm];
   Incidence angle: $\theta = 30°$, $\varphi = 180°$
4) Region: [40 mm, −15 mm]×[70 mm, 15 mm];
   Incidence angle: $\theta = 0°$, $\varphi = 180°$.

TABLE I
ERROR DISTRIBUTION—TEST 1 ($\theta = 0°$)

	Bias (*mm*)	RMS Error (*mm*)	Std. Deviation (*mm*)	FWHM (*mm*)
Anger	—	5.8185	5.8185	8.86
LLR	0.4927	4.1869	4.1579	4.49
MLP	0.0070	3.8833	3.8832	3.63
RBF	0.0563	4.0083	4.0080	4.40

TABLE II
ERROR DISTRIBUTION—TEST 2 ($\theta = 10°$)

	Bias (*mm*)	RMS Error (*mm*)	Std. Deviation (*mm*)	FWHM (*mm*)
Anger	—	6.5805	6.5805	11.26
LLR	0.8106	4.2798	4.2025	6.17
MLP	0.1601	4.1601	4.1572	5.18
RBF	0.6935	4.0225	3.9624	5.75

TABLE III
ERROR DISTRIBUTION—TEST 3 ($\theta = 30°$)

	Bias (*mm*)	RMS Error (*mm*)	Std. Deviation (*mm*)	FWHM (*mm*)
Anger	—	10.8209	10.8209	25.30
LLR	-1.0584	5.3644	5.2590	10.75
MLP	0.2323	4.9454	4.9400	13.39
RBF	0.5754	5.0396	5.0067	8.23

TABLE IV
ERROR DISTRIBUTION—TEST 4 (PERIPHERAL REGION, $\theta = 0°$)

	Bias (*mm*)	RMS Error (*mm*)	Std. Deviation (*mm*)	FWHM (*mm*)
Anger	—	11.8644	11.8644	26.76
LLR	-0.2680	5.2147	5.2078	8.71
MLP	-0.0596	4.3069	4.3065	5.13
RBF	-0.3079	4.2259	4.2147	6.69

Test 1 served as a reference. Tests 2 and 3 were used to examine the sensitivity to parallax effect. Test 4 was used to examine the influence of edge effects. The following algorithms for estimation of scintillation coordinates were compared:

- *Anger algorithm* with threshold pulse pre-amplification, which is used in many modern scintillation detectors [3], [12], [13]. The output of the Anger algorithm was corrected by an ideal infinitely-dense bias correction map (unachievable in practice).
- *Localized linear regression* (LLR). Position estimation was carried out according to the formula

$$y = a_0 + \boldsymbol{a}^T x$$

where Y is the estimated position (X coordinate), x is the vector of PMT responses, and $(a_0, \boldsymbol{a})$ is the vector of regression coefficients found by minimizing the MSE over the training set in the tested region

$$\boldsymbol{a} = \underset{a}{\operatorname{argmin}} \left\{ \sum_{i=1}^{N} \left(a_0 + \boldsymbol{a}^T x^i - y^i \right)^2 \right\}$$

where $\{\boldsymbol{x}^i, y^i\}_{i=1}^N$ is the training set.
- *Multilayer perceptron* (MLP) with three nonlinear (tanh) layers (ten neurons each) and one linear neuron, implemented using the MathWorks MATLAB NN Toolbox [14].
- *Radial basis function network* (RBF) with 100 Gaussian neurons, implemented using the MATLAB RBF toolbox by Müller *et al.* [10].

In all of the estimations employing these algorithms, the input was a vector of 7 PMT responses and the output was the X coordinate of the estimated LOF entrance position. Root mean squared (RMS) estimation error was calculated at each node of the test set grid according to

$$\mathrm{RMS} = \left[\frac{1}{N} \sum_{i=1}^{N} (\hat{y}^i - y^i)^2 \right]^{1/2}.$$

Minimum, maximum and average RMS error over the test set grid was found. Position estimation error histogram was computed on the data of the test set and error distribution parameters such as bias (mean error), RMS error, standard deviation and FWHM were estimated.

A. The Influence of Multiple Interactions

An additional test (Test 5) was performed to study the influence of multiple interactions in the scintillation crystal on scintillation position estimation accuracy. For this purpose, a test set, consisting of 100 000 gamma quanta impinging the crystal at incidence angle $\theta = 30°$ in the center of the region used in Test 3, was generated. For each event, the total number of interactions, of which it comprised, was recorded. Then, the test set was divided into subsets containing events with fixed number of interactions (from 1 to 5). RMS error of the estimators obtained in Test 3 was calculated on these sets.

V. RESULTS

Tables I–IV summarize the estimation error distribution parameters for different algorithms obtained in Tests 1–4, respectively. Figs. 7–10 show the position estimation error histograms obtained in Tests 1–4, respectively. All the histograms are normalized by the peak value to make easier FWHM comparison.

Anger algorithm appeared the worst method in all tests. Local linear regression, which did not appear to be the best among the compared adaptive estimation algorithms, yielded in all tests better results than Anger's algorithm. This fact emphasizes the importance of locality. Nonlinear estimation, using MLP and RBF networks, showed the best results in all tests.

The best nonlinear methods showed about 33%, 40%, 54% and 11% of improvement in the RMS error compared to the

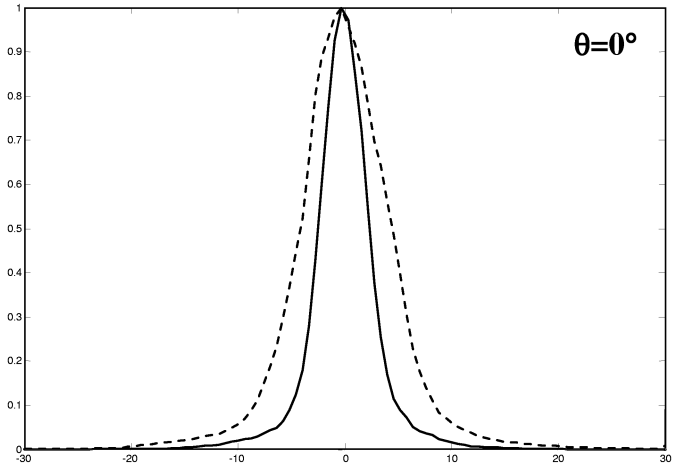

Fig. 7. Error histogram of the unbiased Anger algorithm (dashed) and the best ANN estimator (solid) in Test 1. X-axis represents error in millimeters.

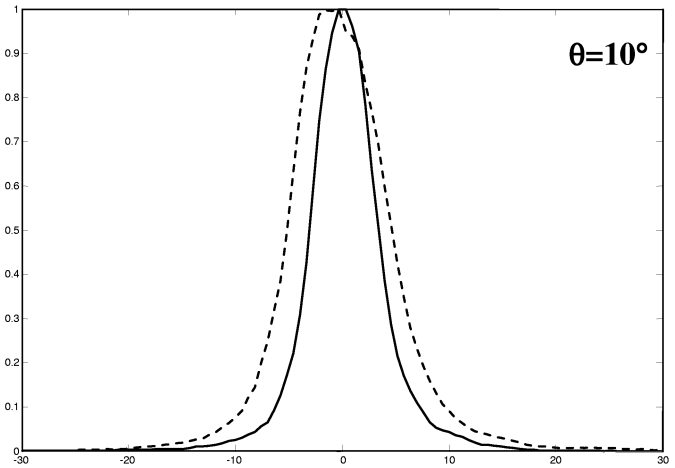

Fig. 8. Error histogram of the unbiased Anger algorithm (dashed) and the best ANN estimator (solid) in Test 2. X-axis represents error in millimeters.

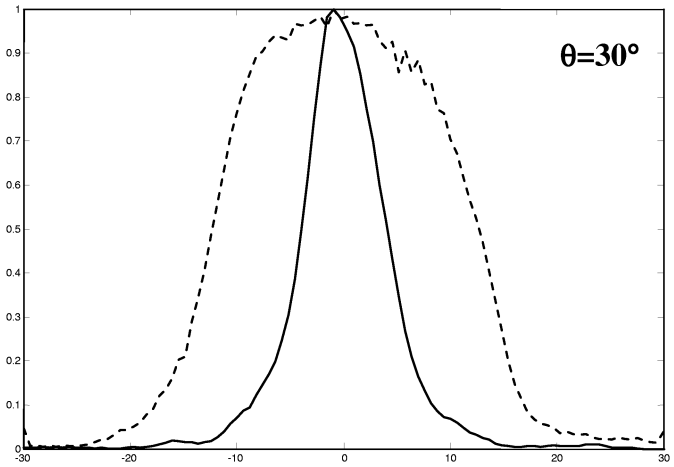

Fig. 9. Error histogram of the unbiased Anger algorithm (dashed) and the best ANN estimator (solid) in Test 3. X-axis represents RMS error in millimeters.

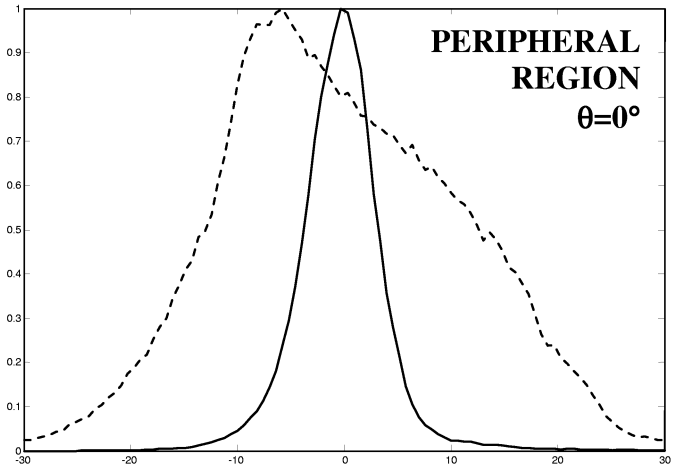

Fig. 10. Error histogram of the unbiased Anger algorithm (dashed) and the best ANN estimator (solid) in Test 4. X axis represents RMS error in millimeters.

Anger algorithm in Tests 1–4, respectively. The improvement rose with the incidence angle and was especially significant in Test 3 (incidence angle 30°). This fact demonstrates the ability of our method to treat large-incidence-angle events more accurately than the Anger algorithm.

Error distributions produced by the nonlinear estimators tend to have smaller FWHM:standard deviation ratio, compared with that characteristic of the Anger algorithm; this can potentially result in additional image resolution improvement [11].

Fig. 11 demonstrates the RMS error in Test 5 as a function of the number of interactions in the crystal. ANN estimators appear to be less sensitive to multiple interactions than the Anger algorithm.

VI. Computational Complexity

MLP network used in this study performs 280 multiplications, 31 additions and 30 computations of the nonlinear function per detected photon. Nonlinear function can be efficiently computed using look-up tables and, in general, its computation is estimated to be equivalent to 5–6 multiplication operations. Preliminary coarse scintillation position estimation and network selection can be carried out in practically negligible time. Hence, about 500 FLOPs are required per scintillation event.

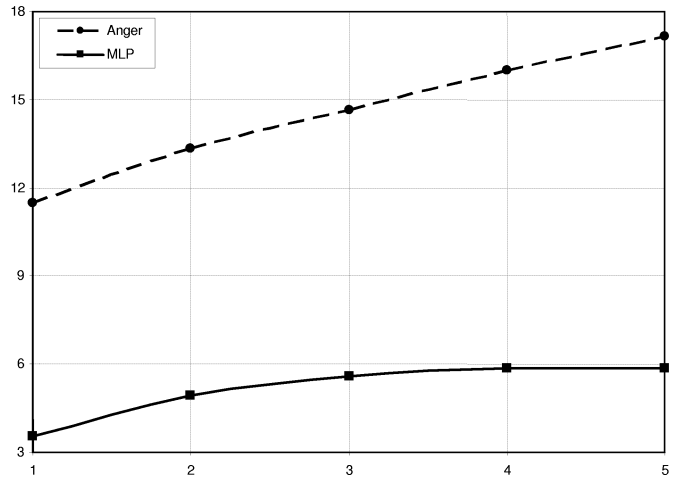

Fig. 11. Test 5: RMS error in mm versus number of interactions. Anger algorithm (dashed) and the best ANN estimator (solid).

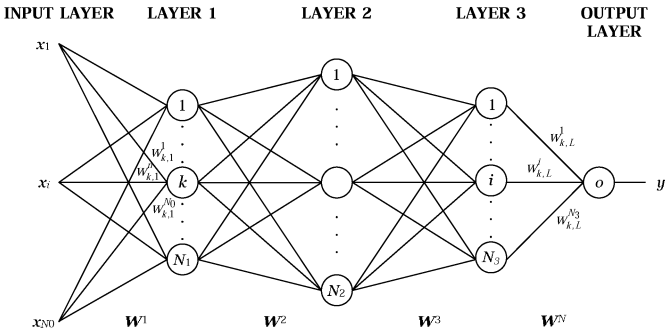

Fig. 12. Multilayer preceptron with three hidden layers and single output neuron.

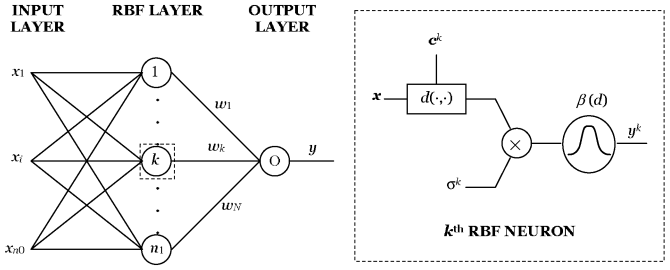

Fig. 13. RBF network with single output neuron.

The RBF networks used in this study require about 700 additions, 100 multiplications and 100 computations of the nonlinear function, which requires about 1000 FLOPs per scintillation event. A 1 GFLOPS processor would be capable of handling about 1×10^6 events per second, which is sufficient for PET applications. Since most types of artificial neural networks have a highly parallelizable pipeline-like architecture, implementation using dedicated hardware may be advantageous.

VII. Conclusion

The proposed method of LOF estimation in PET, based on training of and processing with artificial neural networks, incorporates information about the incidence angle in the estimation algorithm. Unlike the conventional algorithms, which estimate the scintillation coordinates, our approach estimates directly the photon LOF, given PMT responses from a pair of detectors. This allows compensation for the parallax effect, multiple Compton scattering and increases effective detection area.

In practice, a different version of the algorithm can be implemented. The neural networks can be fed with the angle as additional input and trained over small regions of the detector on a range of angles with given angular resolution. The implementation of localized estimators allows reduction in the complexity of each estimator. The proposed algorithm is sufficiently fast to be implemented in real time using standard software or special purpose hardware.

Appendix

The following neural network architectures were used in this work.

1) Multilayer perceptron (MLP), shown in Fig. 12

$$y_k^n = \sum_{i=1}^{N_{n-1}} \varphi\left(w_{k,n}^i y_i^{n-1}\right) + b_{k,n}; \quad y_k^0 = x_k$$
$$\Phi = \sum_{i=1}^{N_{L-1}} w_L^i Y_i^{L-1} + b_L$$

where L is the number of layers; N_l is the number of neurons in each layer; y_k^n and $x_{k,n}, b_{k,n}$ are the output and the parameter vector of kth neuron in nth layer, respectively; x and Φ are the network input and output, respectively; φ is some nonlinear function, usually of a sigmoidal type [7].

2) Radial basis function (RBF) network, shown in Fig. 13

$$\Phi = \sum_{k=1}^{N} w_k \beta(d(x, c_k); \sigma_k) + b$$

where N is the number of neurons in the nonlinear layer; $\{w, b\}$ are the network weights; β is a Gaussian with controllable variance, σ_k, and mean, c_k [7].

References

[1] H. O. Anger, "Scintillation camera," *Rev. Sci. Instrum.*, vol. 29, pp. 27–33, 1958.
[2] W. R. Cook, M. Finger, and T. A. Prince, "A thick anger camera for gamma-ray astronomy," *IEEE Trans. Nucl. Sci.*, vol. NS-32, pp. 129–133, 1985.
[3] H. Barrett and W. Swindell, *Radiological Imaging*. New York: Academic, 1981.
[4] T. Tomitani, Y. Futami, Y. Iseki, S. Kouda, T. Nishio, T. Murakami, A. Kitagawa, M. Kanazawa, E. Urakabe, M. Shinbo, and T. Kanai, "Depth encoding of point-of-interaction in thick scintillation cameras," in *Proc. IEEE MIC*, Seattle, WA, 1999.
[5] S. Delorme, R. Frei, C. Joseph, J.-F. Loude, and C. Morel, "Use of a neural network to exploit light division in a triangular scintillating crystal," *Nucl. Instrum. Methods Phys. Res.*, vol. A 373, pp. 111–118, 1996.
[6] D. Clément, R. Frei, J.-F. Loude, and C. Morel, "Development of a 3D position sensitive scintillation detector using neural networks," in *Proc. IEEE Medical Imaging Conf.*, Toronto, ON, Canada, Nov. 1998.
[7] S. Haykin, *Neural Networks: A Comprehensive Foundation*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1999.
[8] T. Levin and C. Moisan, *A Crash Course on Using the TRIUMF Detector Modeling Platform*: Triumf, Canada's Nat. Lab. Particle and Nuclear Phys., 1996.
[9] G. Tsang, C. Moisan, and J. G. Rogers, "A simulation to model positron encoding multicrystal PET detectors," *IEEE Trans. Nucl. Sci.*, vol. 42, pp. 2236–2242, Dec. 1995.
[10] K. R. Müller, A. Smola, G. Rätsch, B. Schölkopf, J. Kohlmorgen, and V. Vapnik, "Using support vector machines for time series prediction," in *Advances in Kernel Methods—Support Vector Learning*. Cambridge, MA: MIT Press, 1998.
[11] A. M. Bronstein, M. M. Bronstein, M. Zibulevsky, and Y. Y. Zeevi. (2002) High energy photon detection in PET using neural networks research, Rep. Technion. [Online]. Available: http://visl.technion.ac.il/bron/works
[12] G. Kulberg and G. Muehllehner, "Scintillation camera with improved resolution," U.S. Patent 3 732 419, May 8, 1973.
[13] D. Inbar, "Method and means for improving the resolution of a Gamma camera," French Patent 2 530 824, Jan. 27, 1984.
[14] *ANN Toolbox User's Guide*, MathWorks, Inc., 2000.