Special issue of the 3rd International Conference on Computational and Experimental Science and Engineering (ICCESEN 2016)

Recovering Images from PET Camera

D. Üstündağ*
The University of Marmara, the Department of Mathematics, Istanbul, Turkey

We study here one of the imaging techniques, used in nuclear medicine, called positron emission tomographic imaging, that provides information about many biological processes that are essential to the functioning of the organ, being visualized. Our emphasis is given to the application of the maximum entropy image reconstruction method called "Cambridge MaxEnt Package" for recovering images of the human brain from data obtained by positron emission tomographic camera.

DOI: 10.12693/APhysPolA.132.963
PACS/topics: 02.30.ZZ, 02.50.–R, 29.30.KV, 87.57.Q–, 87.57.U

1. Introduction

In the last few decades of the century there have been significant advances in nuclear medicine. In particularly, when Allen Cormack [1] and Godfrey Hounsfield [2] introduced X-ray computer tomography (CT) independently in the early 1970s, based on the mathematical foundation of Radon [3], for reconstruction images of an object from its projections, the field of nuclear medicine was revolutionized.

The breakthrough development of X-ray CT was made possible by continuing advances in instrumentation and computer technologies, which also accelerated the development of the multi-dimensional imaging modalities, that carry a great potential for providing, in addition to morphologic, the dynamic and functional information on biochemical and pathophysiological processes or organs of the human body.

Recent advances in basic molecular and cell biology have also made us to change in our understanding of diseases, which can be defined as alternations in cellular behaviour, that reflect functional changes, instead of defining it as structural changes. Actually, it is evident that by the time when pathological conditions are easily visible either in a conventional CT or the latest scanners, the underlying biochemical abnormality is usually well advanced. This is because X-ray CT imaging is mainly concerned with anatomical structures of body, distinguishing between different components that have different absorbing power of X-rays.

Nevertheless, the latest scanners can provide radically improved qualitative diagnostic information, relevant to a wide range of human diseases but, give limited information about the functional or physiological state of the internal organs of human body. In medicine, it is known that human diseases often occur with no specific anatomical changes and those seen may be the later effects of early biochemical process, which have remained undetected until the advanced symptoms appear in the patient.

Therefore, the goal of the radiology specialty of nuclear medicine is to provide information on the distribution of a chosen molecule in space inside the human body, so that an image of its distribution within the body, or a specific organ, provides information on the functioning of the body or organ, that is valuable for medical diagnosis. For this purpose, PET imaging techniques [4–7], for measuring concentrations of positron-emitting radioisotope within the tissue of living subjects, have been developed by researchers from diverse disciplines for the last decade. Since first introduction of PET to medical imaging applications in the late 1960's, it has already grown into a well-researched, highly evolved field [7–9].

Most of current approaches, to tackle the problem of reconstruction of isotope concentration distribution, can be classified into two general categories, namely the analytical methods [6, 10], which rely on the inversion of Radon transform, and the iterative approaches [11], which are based on a statistical description of the physical problem.

Because of the random nature of the radioactive disintegration, the tomographic data are noisy and therefore it is straightforward to regard PET reconstruction as a statistical estimation problem [12–14]. Such approaches, when reconstructing PET images, require to introduce modelling of the data statistics and to make use of some prior information about the PET imaging system. In this paper, we therefore focus on PET image reconstruction method, based on Bayesian approach [13, 18], with maximum entropy prior [16] and use here a well-known powerful algorithm of "Cambridge MaxEnt Package" (CMEP) [15] for reconstructing maximum entropy PET images. It allows incorporation of system model and measurement of statistical uncertainties, thus providing a more robust and accurate solution.

2. PET measurement modelling

The physical basis for PET imaging, shown in Fig. 1a, lies in the fact that a positron produced by a radioactive nucleus annihilates with an electron to form a pair of

*corresponding author; e-mail: dustundag@marmara.edu.tr

high energy photons, after traveling very short distance. The pair of photons travel back to back along a straight line path. These photon pairs can be detected externally within a timing window τ by two opposite detectors.

The shaded line of response corresponds to the volume between the detectors and is defined by cross-sectional area of the coincidence detectors 1 and 2. C_1 and C_2 are the single count rates, recorded by detector 1 and 2, respectively. However, in reality, the coincidence events may also include those in which two gamma rays originate from two unrelated position annihilations and those in which the annihilation photons lose their direction and energy before they arrive to the detector system.

Actually, the raw data recorded by the PET camera consist of 10^6 random coincidence events, which are represented by a pair of coordinate values, one from each detector. Thus, image reconstruction may then proceed in two stages. The first stage is to back-project the raw data. The second stage is to deconvolve the back-projected data with the point-spread function, shown in Fig. 1b.

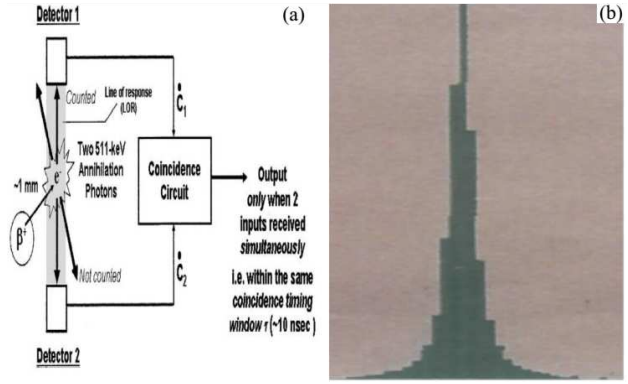

Fig. 1. (a) The physical basis for PET: annihilation of positron and electron produces a pair of back to back 511 keV photons that are detected by a pair of detectors. (b) The point-spread function, experimentally obtained by a "spread" out version of the point source, placed in the system and reconstructed into an image, which looks like a Gaussian shaped function with very long tails.

To form a back-projection, the image volume (typically a cube, $600 \times 300 \times 124$ mm^3) consists of K equally spaced planes with equal dimensions, each of which is divided into (256×128) square elements (pixels). Then, a simple back projected data is obtained by a computer program, which forms a line of response between a pair of coordinate-values, recorded as raw data and increments the count in the pixel where this line passes in each back-projection plane.

3. Mathematical problem of PET system

In PET, the back projected image $\boldsymbol{D} = \{d_i\}$ $(i = 1, ..., M)$ is composed of M pixels. R_{ij} is proportional to the probability that a photon pair, emitted from pixels j is detected by the tube i and is called the point-spread function or transfer matrix with a dimension $(M \times N)$. Then, the image formation equation can be implemented in the form:

$$d_i = \sum_{j=1}^{N} R_{ij} f_j, \ (i = 1, 2, 3, ..., M), \tag{1}$$

where the point spread function $\boldsymbol{R}$, illustrated in Fig. 1b, is assumed to be modelled correctly, e.g., sampling effects and scattering contributions are properly incorporated. Given the back-projected data $\boldsymbol{D}$ and the knowledge of $\boldsymbol{R}$, the mathematical problem is to estimate $\boldsymbol{f}$.

Mathematically, such an inverse problem is an ill posed one, making the solution non-trivial and ambiguous: very small variations in the experimental data $\boldsymbol{D}$ lead to very large variations in the solution $\boldsymbol{f}$ and generally requires some kind of regularization in order to generate physically plausible solutions.

This can be done in a variety of ways [14], but a commonly used idea to realize regularization techniques with statistical motivation is the Bayesian model, using the posterior probability density function (PDF) $p(\boldsymbol{f}\,|\boldsymbol{D})$, given according to Bayes formula;

$$p(\boldsymbol{f}\,|\boldsymbol{D}) = \frac{p(\boldsymbol{D}\,|\boldsymbol{f})p(\boldsymbol{f})}{p(\boldsymbol{D})} \propto p(\boldsymbol{D}\,|\boldsymbol{f})p(\boldsymbol{f}), \tag{2}$$

where the term $p(\boldsymbol{D})$ may be considered necessary only for normalization purposes, so that it is often written as a proportionality, leaving out the denominator. The Bayesian approach has the advantage that it allows for incorporating additional prior information on $\boldsymbol{f}$, via the a-priori PDF $p(\boldsymbol{f})$, into the reconstruction process. According to Skilling [16], the prior PDF $p(\boldsymbol{f})$ for an unnormalised positive additive image $\boldsymbol{f}$ is proportional to $\exp(\alpha S(\boldsymbol{f}, \boldsymbol{m}))$, where the entropy $S(\boldsymbol{f}, \boldsymbol{m})$ of the required map $\boldsymbol{f}$, relative to an initial map $\boldsymbol{m}$, is given by

$$S(\boldsymbol{f}, \boldsymbol{m}) = \sum_{j=1}^{N} f_i - m_j - f_j \log\left(\frac{f_j}{m_j}\right). \tag{3}$$

α, denoted as a regularization parameter, is initially undetermined, so the solution is conditional on two quantities, α and $\boldsymbol{m}$. For a Poisson-distributed event counting experiment, the conditional probability of the whole data set is

$$p(\boldsymbol{D}\,|\boldsymbol{f}) = \frac{1}{\prod_{i=1}^{N}\sqrt{2\pi}\,\sigma_i} \exp\left(-\frac{\chi^2(\boldsymbol{f})}{2}\right), \tag{4}$$

where $\chi^2(\boldsymbol{f})$ is usually used to measure the misfit:

$$\chi^2(\boldsymbol{f}) = \sum_{i=1}^{M} \frac{1}{\sigma_i^2}\left(d_i - \sum_{j=1}^{N} R_{ij} f_j\right)^2, \tag{5}$$

where σ_i is the standard deviation of the noise in the ith pixel. Then the posterior PDF in Eq. (2) turns into the following form:

$$p(\boldsymbol{f}\,|\boldsymbol{D}) \sim \exp\left(\alpha S(\boldsymbol{f}, \boldsymbol{m}) - \frac{1}{2}\chi^2(\boldsymbol{f})\right). \tag{6}$$

The computationally interesting Bayesian approach, called the maximum a-posteriori probability (MAP)

estimation, consists of computing an estimate $\boldsymbol{f}$ of the unknown object by maximizing the a-posteriori PDF $p(\boldsymbol{f}\,|\boldsymbol{D})$, given in Eq. (6). Various numerical algorithms for solving this non-linear optimisation problem have been suggested by various researchers [13, 14]. However, we shall only consider here a complicated, but highly successful scheme, developed by Skilling and collaborators [15, 17], wherein a maximum is repeatedly sought not along a single search direction, but in a small dimensional subspace, spanned by vectors that are calculated at each landing point. The subspaces basis vectors are chosen in such a way, as to avoid directions leading to negative values. One of the most successful choices is the three-dimensional subspace spanned by the vectors with components given by

$$\begin{aligned} \boldsymbol{e}_1 &= \boldsymbol{f}\,(\nabla S)\,, \\ \boldsymbol{e}_2 &= \boldsymbol{f}\,(\nabla\chi^2)\,, \\ \boldsymbol{e}_3 &= |\nabla S|^{-1}\,\boldsymbol{f}\,(\nabla^2\chi^2)\,\boldsymbol{f}\,(\nabla S) \\ &\quad - \left|\nabla\chi^2\right|^{-1}\boldsymbol{f}\,(\nabla^2\chi^2)\,\boldsymbol{f}\,(\nabla\chi^2)\,. \end{aligned} \tag{7}$$

Here the entropy metric $(-\nabla\nabla S)$ is used to define the lengths

$$|\nabla S| = \left(\sum f_i \left(\frac{\partial S}{\partial f_i}\right)^2\right)^{\frac{1}{2}},$$

$$|\chi^2| = \left(\sum f_i \left(\frac{\partial \chi^2}{\partial f_i}\right)^2\right)^{\frac{1}{2}}. \tag{8}$$

With these three search directions, quadratic models for $S(\boldsymbol{f},\boldsymbol{m})$ and $\chi^2(\boldsymbol{f})$ are constructed in the subspace given in Eq. (6) and then the algorithm proceeds by determining coefficients of the search vector $\delta\boldsymbol{f} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + x_3\boldsymbol{e}_3$, that gives the maximum of $S(\boldsymbol{f},\boldsymbol{m})$, subject to the constraint $\chi^2(\boldsymbol{f}) \le N$, by implicitly adjusting α, so that the constraint is satisfied. More precisely, in each iteration, $S(\boldsymbol{f},\boldsymbol{m})$ and $\chi^2(\boldsymbol{f})$ are projected onto a small dimensional subspace where $S(\boldsymbol{f},\boldsymbol{m})$ attains its maximum, while $\chi^2(\boldsymbol{f}) \le N$ is satisfied. Finally, the current $\boldsymbol{f}$ is moved to the new location by $\hat{\boldsymbol{f}} = \boldsymbol{f} + \delta\boldsymbol{f}$, while the updated $\hat{\boldsymbol{f}}$ needs to be protected against stray on negative values. The iteration is repeated until

$$\text{TEST} = \frac{1}{2}\left|\frac{\nabla S}{||\nabla S||} - \frac{\nabla\chi^2}{||\nabla\chi^2||}\right|^2 < \varepsilon, \tag{9}$$

where ε is a preassigned value.

4. Recovering PET images

The concentration of a chemical compound, labeled gallium isotope (^{68}Ga) with the half-life of 68 minutes, in the brain was measured by using the PET camera. Data was provided by the positron emission group at the School of Physics and Space Research at the University of Birmingham. What is recorded by the PET camera is a pair of coordinate values, one from each detector, which represents an annihilation event. By using these recorded pairs of coordinate values, the back-projected images of brain, shown in Figs. 2a and 3a, were constructed. Imaging or data analysis problems of PET camera is only considered here, however it is known that such studies provide quantitative functional information, which helps to quantify normal human brain physiology and highlights the changes in metabolism and function caused by cerebral disease.

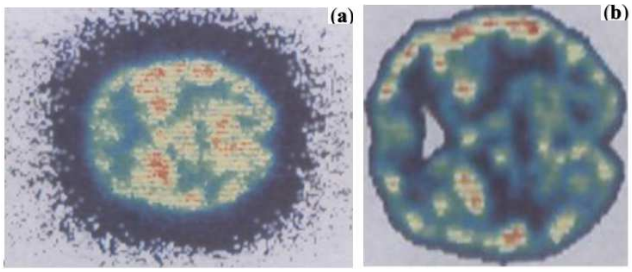

Fig. 2. (a) The back-projected data. (b) The reconstructed image formed by the CMEP.

The images, shown in Figs. 2a and 3a, are the sectional images of the brain, that are colour-coded to show differences in the level of activity from point to point. The number of random coincidence events collected by the PET camera distinguishes these two sets of data, taken from the central plane. The back projected data in Fig. 3a was therefore obtained by discarding 10% of information from the back-projected data, given in Fig. 2a. It is expected that with more counts collected, the better will be the resulting image quality.

The CMEP requires a specification of a model image and the variance of the noise in the data. The model image $\boldsymbol{m}$ represents our prior state of knowledge or ignorance about the original image. It is usually taken to be uniform, but the use of a non-uniform model image $\boldsymbol{m}$ allows us to introduce partial correlations across the image we want to infer. Therefore, at the kth iteration we estimated the model $\boldsymbol{m}^k$ in the following form:

$$\boldsymbol{m}^k = \boldsymbol{B}\boldsymbol{f}^k, \tag{10}$$

where $\boldsymbol{B}$ is a blurring matrix, which averages the densities over neighbouring pixels (3×3). The uncertainty in the data, associated with each pixel, varies from pixel to pixel according to the number of photons counted. Therefore, the variance σ_i^2 of the noise was estimated in the form:

$$\sigma_i^2 = \begin{cases} d_i & \text{if } d_i \neq 0 \\ 1 & \text{if } d_i = 0 \end{cases}, \quad i = 1, ..., M. \tag{11}$$

A fictional photon is added here to avoid the discontinuity arisen from the ith component of data d_i. The CMEM was run on an IBM3090 computer. After approximately 30 iterations the convergence criterion (TEST $\le$ 0.01, $\chi^2\,(\boldsymbol{f}) \le 16384$) was satisfied.

The colour-coded images of the reconstructed densities are shown in Figs. 2b and 3b. The red areas in the pictures indicate high levels of activity; dark blue or white areas indicate little or no activity in those brain

structures. It is also observed that these images have almost similar brain structures, even though the data shown in Fig. 3a contain less information than the one in Fig. 2a. As expected, CMEP requires large consumption of CPU time on the computer, because it is non-linear and iterative.

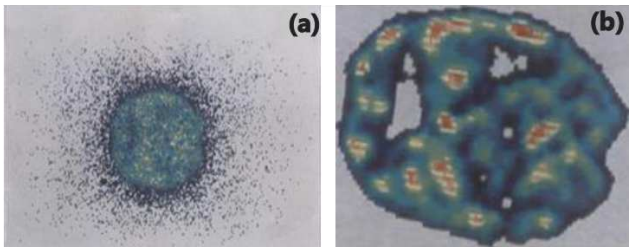

Fig. 3. (a) The back-projected data. (b) The image reconstructed by the CMEP.

3. Conclusions

In this work, CMEP is used to reconstruct PET images with Poisson noise model, including partial correlations. It is based on a subspace of several search directions, instead of a line search, and the noise level constraint over the likelihood function is controlled, so that the positivity during each subspace search is directly enforced. These features make the algorithm attractive for solving high dimensional nonlinear problems.

Overall, PET reconstructions indicate that the application of the maximum entropy method leads to considerably improved reconstruction, especially in the case where the data are very noisy and the sampling is incomplete. Along with improving the reconstruction methods for solving the basic PET reconstruction problem, there are many novel challenges, facing this field, and new advances in PET instrumentation require hybrid reconstructions algorithms to reach their potential improvements in signal to noise ratio. Clinical importance of PET imaging will continue to increase.

Acknowledgments

We would like to express our thanks to Dr. Hawkesworth and Dr. D. Paker for providing the back projected PET data and to Dr. J. Skilling for sending us the Cambridge MaxEnt program and for his comments about the results. This work is a part of project whose name is "Positron lifetimes spectrum analysis with Bayesian statistical inference" with a number FEN-A-060510-0135 supported by Marmara University, Istanbul, Turkey.

References

[1] A.M. Cormack, *J. App. Physics* **35**, 2908 (1964).

[2] G.N. Hounsfield, *Brittan J. Radiol.* **46**, 1016 (1973).

[3] J. Radon, *Math. Phys. Klass.* **69**, 262 (1917).

[4] N.M. Ter-Pogossian, M.E. Raichle, B.E. Sobel, *Scientific America* **243**, 140 (1980).

[5] J. Verhaeghe, A.J. Reader, *Phys. Med. Biol.* **58**, 393 (2013).

[6] G.L. Zeng, *Medical image reconstruction: A conceptual tutorial*, Springer Heidelberg Dordrecht London, New York 2010.

[7] H. Zaidi, M. Becker, *IEEE Signal Process. Mag.* **33**, 67 (2016).

[8] A. DelGuerra, N. Belcari, M. Bisogni, *Riv. Nuovo Cimento* **39**, 155 (2016).

[9] M.J. Ehrhardt, K. Thielemans, L. Pizarro, D. Atkinson, S. Ourselin, B.F. Hutton, S.R. Arridge, *Inverse Problems* **31**, 015001 (2015).

[10] V. Bettinardi, L. Presotto, E. Rapisarda, *Med. Phys.* **38**, 5394 (2011).

[11] J. Qi, R.M. Leahy, *Phys. Med. Biol.* **51**, 541 (2006).

[12] R. Leahy, J. Qi, *Stat. Computing* **10**, 147 (2002).

[13] K. M. Hanson, *Proc. SPIE* **1898**, 716 (1993).

[14] A.R Davies, R.S. Anderssen, *J. Austr. Math. Soc. Series B* **28**, 114 (1986).

[15] J. Skilling, R.K. Bryan, *Monthly National Radio Astronomy Soc.* **211**, 111 (1984).

[16] J. Skilling, in: *Maximum Entropy and Bayesian Methods* Ed. F. Fougčre, Kluwer Academic Publishers, Dordrecht 1990, p. 341.

[17] R. Bryan, *J. Phys. Colloques* **47**, 43 (1986).

[18] M. Cevri, D. Üstündağ, *Acta Phys. Pol. A* **130**, 45 (2016).