Towards Optimal Model-based Partial Volume Effect Correction in Oncological PET Imaging

Frédéric Schoenahl and Habib Zaidi

***Abstract*—Quantification of tumor uptake in nuclear medicine requires the assessment of the tumor size. We propose a method based on simple ellipsoidal models which only require the reconstructed image data and scanner physical parameters. The method uses simple image and system models but is theoretically not sensitive to the reconstructed image sampling, provided that the local resolution parameters of the scanner are known. Interpolation is thus avoided and the degradation model is computed analytically. The proposed method uses the property that the blurring effect of the scanner spreads the activity around the lesion location and in the neighbor-voxels – even for sub-voxel sized lesions - and attempts to use this limited information to accurately restore lesion parameters. We propose solutions to avoid bias and reduce computation time. The uptake of simple simulated tumor objects could be restored and the influence of lesion's shape and neighborhood was investigated on real patient data for which convergence of the algorithm was demonstrated. This approach is a contribution towards in-depth investigation of reconstruction/post-reconstruction-based models for tumor activity restoration based on more complex image and system models.**

I. Introduction

Absolute quantification is a key index for clinical diagnosis and assessment of tumor response to therapy. In particular, the bias and blurring effects induced by physical degrading factors and reconstruction algorithms on oncological images give the three dimensional tumor volumes indefinite borders and fuzzy estimates for their quantification. A method has been proposed based on modeling the system response using a non-linear resolution model (*NSVOD*) [1] taking into account the scanner physical properties and position-dependent resolution within the field-of-view. These parameters are used to restore the activity of small spheroid lesions [2]. However, for very small tumors where the dimensions are at the order of the system FWHM, a model-to-image matching is not practically feasible due to the small dimensions of the object and the discretization effects induced by the low image sampling. Among the possible solutions to perform this efficiently and independently of the sampling, one can consider e.g. a three-dimensional Voronoi diagram and different volume and model sampling schemes, which corresponds to a simple nearest-neighbor interpolation approach, or more advanced interpolation techniques where computation time increases with the accuracy and the complexity of the interpolation model. We propose an alternative iterative method which leads to optimal matching between the tumor model and the reconstructed image independently of the sampling properties of the considered image.

This work was supported by the Swiss National Science Foundation under grant SNSF 3152A0-102143 and Geneva University Hospital R&D funds under grant PRD-04-1-08.

F. Schoenahl is with the Division of Nuclear Medicine, Geneva University Hospital, CH-1211 Geneva, Switzerland (e-mail: schoenahl@ieee.org)

H. Zaidi is with the Division of Nuclear Medicine, Geneva University Hospital, CH-1211 Geneva, Switzerland (e-mail: habib.zaidi@hcuge.ch)

II. Theory

As small size tumors are considered, some approximations can be made regarding the degradation model. The local resolution of the system is considered constant over the investigated object and its close neighborhood. The image object is selected as a volume of interest (VOI). The VOI should contain the points which are visually supposed to contain the actual tumor object, added to a margin of a few voxels (the "quality" of the cropping is of limited importance for the method presented here). The degradation model of the imaging system we use is simply the convolution of an ellipsoidal object with a three-dimensional anisotropic Gaussian function shifted by $\boldsymbol{X} = [X_1, X_2, X_3]$ with standard deviations $\boldsymbol{\sigma} = [\sigma_1, \sigma_2, \sigma_3]$ in the three directions of the Cartesian space. The degradation model used here is simply

$$V = G \otimes \hat{O} + n \tag{1}$$

where V denotes the selected discrete image function (VOI), $\hat{O}$ the corresponding ellipsoidal model and n is an additive noise component. For the sake of simplicity, noise was neglected in our model. The Gaussian kernel is defined as

$$G(\mathbf{X}) = \frac{1}{\sigma_1\sigma_2\sigma_3 2\pi} \cdot e^{-\frac{1}{2}\left(\frac{(x_1-X_1)^2}{\sigma^2_1}+\frac{(x_2-X_2)^2}{\sigma^2_2}+\frac{(x_3-X_3)^2}{\sigma^2_3}\right)} \tag{2}$$

$$\text{with } \sigma_d = \frac{FWHM_d}{2}$$

The lesion model O is described by the following model derived from the canonical equation of an ellipsoid:

$$O(\mathbf{x}) = Q(\mathbf{x}) \cdot \mathrm{H}_0\left(1-(x_1/s_1)^2-(x_2/s_2)^2-(x_3/s_3)^2\right) \tag{3}$$

In (3), Q is the tumor uptake distribution sampled at the location of the ellipsoid of radii $\boldsymbol{s} = [s_1\ s_2\ s_3]$ and H_0 is a modified Heaviside step function defined as $H_0(x) = 1$ if $x \geq 0$ and $H_0(x) = 0$ otherwise.

The hypothesis drawn on the true object parameters takes the form of a transformation matrix $\boldsymbol{T}$ which contains supposed rotation (Euler angles) and translation parameters necessary to map O onto the image-object V, giving $\hat{O}$. If $\boldsymbol{P}$ is a vector containing rotation ($\boldsymbol{r}$) and translation ($\boldsymbol{t}$) parameters, $\boldsymbol{T}$ is created as [3]:

$$\boldsymbol{P} = [\boldsymbol{r}\ \ \boldsymbol{t}\,]$$

$$\rightarrow \mathbf{T} = \begin{bmatrix} c_2c_3 & -\beta\cdot c_2s_3 & -\beta\cdot s_2 & t_1 \\ -s_1s_2c_3+\beta\cdot c_1s_3 & \beta\cdot s_1s_2s_3+c_1c_3 & -\beta\cdot s_1c_2 & t_2 \\ \beta\cdot c_1s_2c_3+s_1s_3 & -c_1s_2s_3+\beta\cdot s_1c_3 & c_1c_2 & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{4}$$

Here $c_d = \cos(r_d)$ and $s_d = \sin(r_d)$ for a rotation around axis d, β is set to +1 or -1 for a direct or an indirect coordinate system, respectively. Note that both the activity q and the size parameter $\boldsymbol{\sigma}$ are directly incorporated in the model. The object equation can thus be transformed according to

$$\hat{O}(\mathbf{x}_e) = O(\mathbf{x}_e\cdot\mathbf{T})\ ;\ \mathbf{x}_e = [\mathbf{x}\quad V(\mathbf{x})] \tag{5}$$

If the parameter values are close to the true shape parameters, the distance (*L1* or *L2*-norm) between each element of a transformed and convolved object model using G and the image-object should be minimal. The simplest and fastest approach is to draw the transformed ellipsoid $\hat{O}$ on the image grid, and then filter the grid data using an n-tap filter with identical sampling. For the smallest objects, a resampling should be done in order to increase the number of points, at the cost of using an interpolation method. Applied to small tumors (close to the scanner resolution limit), the subsequent interpolations for rotations of the model or the resampling of the image become of drastic importance.

We therefore propose to reverse the procedure. As the model is analytic, it is ideally sampled at any point in the continuous model space $\Re^3$ and likewise, can be convolved using a continuous Gaussian function in the model space. The algorithm is depicted in Fig. 2.

The VOI is rotated and shifted as a point list (thus keeping exact coordinates after transformations) without changing points intensities and using the inverse of the transformation matrix $\boldsymbol{T}^{-1}$:

$$\mathbf{X}_{\hat{V}} = \mathbf{X}_V\cdot\mathbf{T}^{-1} \tag{6}$$

The new coordinates are model-space coordinates, where the convolution of the model can be computed analytically:

$$\hat{V}(\mathbf{X}_{\hat{V}}) = \int_{\infty} G(\mathbf{X}_{\hat{V}}-\mathbf{x})\cdot O(\mathbf{x})\cdot d\mathbf{x} \quad \Rightarrow$$

$$\hat{V}(\mathbf{X}_{\hat{V}}) = \int G(X_3-x_3)\int G(X_2-x_2)\int G(X_1-x_1)\cdot O(x_1,x_2,x_3)\cdot dxdydz \tag{7}$$

In order to reduce essentially the computational complexity of Eq. (7), an interesting fact derived from the hypothesis that the convolution kernel has almost *constant parameters* over the volume of interest[1], and from the *separability property* of Gaussian functions, is that the convolution operation along the first coordinate of the ellipsoid model is equivalent to sampling a cumulated Gaussian function centered on the point for which the new intensity has to be calculated. Thus if C is the cumulated Gaussian function, Eq. (7) becomes:

$$\hat{V} = \int_{-c}^{+c} G(X_3-x_3)\int_{-b(x_3)}^{b(x_3)} qG(X_2-x_2)\left[C(-X_1+a)-C(-X_1-a)\right]dydz \tag{8}$$

Here, q is assumed to be a constant term $Q(\mathbf{x}) = q$ within the object function. The minimization takes the final form:

$$\arg\min_{\mathbf{P}} f(\mathbf{P}) \qquad \text{for} \quad f(\mathbf{P}) = w\left|\hat{V}-V\right| \tag{9}$$

This general term assumes in the ideal case that the ellipsoid is drawn on a flat-null background ($w(\mathbf{x}) = 1$). In reality PET acquisitions induce a typical background level of ~10-60% of the maximum lesion intensity. The background can be very inhomogeneous at tissue interfaces. The main pitfall for the optimization is a proposed $\boldsymbol{P}$ which corresponds to a very large ellipsoid of intensity roughly equal to the noise level and leads to a better match (local minimum) with the VOI point-set. This might occur due to inhomogeneous background or if the VOI selection around the tumor is too large. In order to reduce this effect we propose to apply a penalty function which cancels the insignificant voxel intensities. The penalty function is quadratic and anisotropic having the following formula:

$$w(\mathbf{X}_{\hat{V}}) = 1/\sqrt{\left(\left|\mathbf{X}_{\hat{V}}\right|/\mathbf{s}\right)^2} \qquad \text{for} \quad \mathbf{X}_{\hat{V}} \subset \Re^{*3}, \mathbf{s} \subset \Re^{+*3} \tag{10}$$
$$\text{and}\ \ w(\mathbf{X}_{\hat{V}}) = 1 \quad \text{else}$$

Equation (10) has only element-wise divisions, and is shown in Fig. 1. This ratio will increase for distant points and has its maximum at the model centre.

[1] The approximation of a local-shift independent resolution has a minor effect as the variation of a typical scanner intrinsic resolution is locally very small (radial ΔFWHM = ~0.01 mm^{-1} for our CTI ECAT-ART PET scanner). As a consequence, the local resolution was set to the FWHM of the reconstructed image at the tumor maximum intensity point.

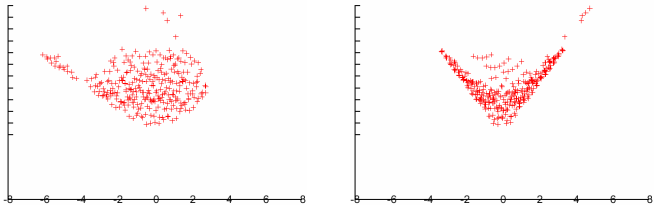

Fig. 1. Anisotropic penalty function projected on the X and Z planes (isotropic units and identical display ranges).

One of the major drawbacks of iterative methods is the high computational time required to reach convergence. Nevertheless, the proposed algorithm is only subject to fixed precision errors for basic operations and transcendental functions calculations. An additional advantage is that the orthogonality property of the 3x3 rotational sub-matrix in Eq. (4) allows a quick and exact computation of the inverse matrix for $\boldsymbol{T}$.

The recalculation of the analytic convolution is a critical issue, which necessitated testing the speed and accuracy of some available numerical methods. The best compromise was found when using a quadratic non-adaptative numerical integration [4]. By reducing the function integration to Eq. (8), a reduction of the computational time by a factor ~10 is obtained. The Nelder and Mead simplex algorithm was used for optimization as it does not need time consuming Hessian or derivatives matrix calculations. As the volume of interest was intended to be small, constraints were set on the radius, center location, and intensity. One also expect them to avoid local minima – in some case, added to the quadratic constraint of Eq. (10). Another issue concerning performance is that the algorithm can process any number of points from the set, provided that the most significant points are included in the computations. A global thresholding strategy at the background estimated mean level (e.g. from the VOI-histogram maximum in the case of a supposed random noise), or above can be performed to suppress unnecessary points. As a consequence, the speed is increased, but requires an increased accuracy (fractional index) for breaking iterations and the tumor to be identified as unique within the VOI. The method is theoretically capable of matching only the ellipsoid points which significantly belong to the shape in the selected VOI. Finally, the VOI selection in our study was automatic by taking points in a neighborhood of 2 to 6 voxels away from the first estimate of the largest lesion volume radius, thus providing rectangular planes. Finer cropping of the image object could be performed by using the first estimates and a quadratic distance function similar to Eq. (10). Besides increasing the computational speed, it would impeach the penalty function to "kill" the tumor volume if the model center reaches the borders of the selected VOI.

III. Results

3D Images which contain ellipsoidal shapes of various intensities and known dimensions using variable system characteristics (FWHM of 2, 3, 4, 5 voxels) were analytically simulated using MEDx (*Medical Numerics, Inc*., Sterling, VA, USA). The background was set with a fixed Gaussian noise level with a mean of 2000 counts and a standard deviation of 500. Scanner resolution parameters were input to the optimizer as static parameters. Estimates were fixed randomly at the tumor size scale and compared after optimization to the actual known parameters. The parameters found for three ellipsoids are reported in Table 1.

TABLE I

Summary of parameter estimates. As more than one combination of parameters can give the same ellipsoid, additional rotations were applied ($\pm \pi/2$) for the radii to be compared. Units for distances are voxels and bold characters are the expected values. TB is the tumor maximum-to-background mean ratio.

P	Simulated images		
	ELL1 (TB=2.5)	ELL2 (TB=2)	ELL3 (TB=2)
t_1	25.50 ± 0.00 (**25.50**)	25.50 ± 0.00 (**25.50**)	27.50 ± 0.08 (**27.50**)
t_2	82.04 ± 0.08 (**82.00**)	82.05 ± 0.06 (**82.00**)	84.57 ± 0.09 (**85.00**)
t_3	9.26 ± 0.27 (**9.50**)	9.7 ± 0.25 (**9.50**)	8.72 ± 0.49 (**9.00**)
s_1	4.64 ± 0.33 (**5.00**)	4.90 ± 0.35 (**5.00**)	5.20 ± 0.56 (**2.50**)
s_2	3.26 ± 0.29 (**3.50**)	3.90 ± 0.37 (**3.50**)	2.50 ± 0.08 (**2.50**)
s_2	3.26 ± 0.90 (**3.00**)	3.55 ± 0.44 (**3.00**)	6.15 ± 0.24 (**5.50**)
Activity (cts)	5298 ± 312 (**5000**)	4741 ± 311 (**4500**)	3584 ± 540 (**4000**)
Total cts x 10^3	1500 ± 300 (**1100**)	1395 ± 212 (**989**)	949 ± 200 (**686**)

Results of Table 1 are in good agreement with expected values considering the facts that the voxelisation of the analytic phantom induced an overestimate in all measurements and that large filters ($FWHM_d$ > 4 voxels) results in an overestimation in the standard deviations. The restored values for 2 and 3 voxels FWHM images have produced excellent final estimates (50% less error in estimated activity).

For testing the effect of different shapes and background on the quality of the matching, the algorithm was run on selected lesions using whole-body clinical datasets reconstructed with OSEM statistical iterative reconstruction algorithm (Fig. 3-a), the lesions were cropped with a rectangular VOI and more accurate margins for very inhomogeneous backgrounds (Fig. 3-II and 3-III).

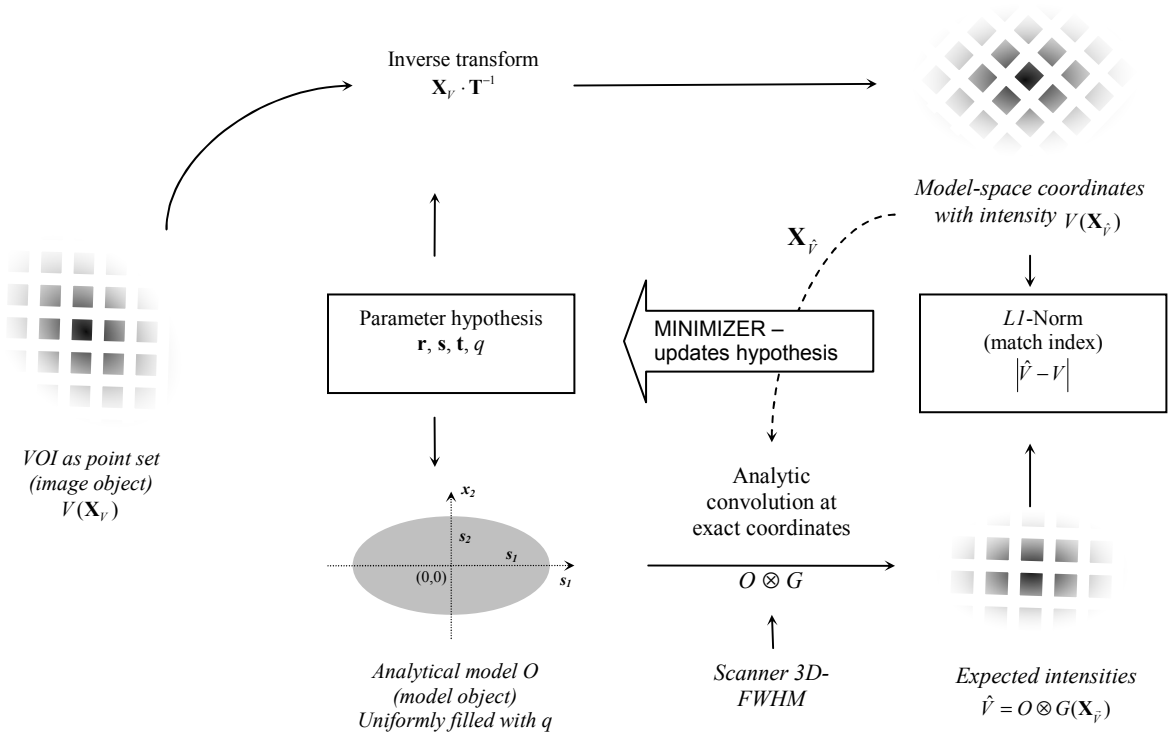

Fig. 2. Ellipsoid model parameters assessment (2D representation). Under the assumption that the parameter hypothesis is exact, the *L1*-norm should be minimal.

IV. Conclusion

The method presented in this work is intended for post-processing-based quantification restoration and converges even for inhomogeneous neighborhood of lesions using clinical reconstructed images. It is efficient because no bias is introduced in the matching process, such as grid resampling and subsequent interpolations, or approximation in the convolution calculation. We neglected in this study the initial averaging per voxel inherent to the reconstruction algorithm, especially if the voxels are large; we expect the ellipsoid model to be slightly biased and the matching index to have a residual non linear component. Taking into account the additional scatter component would lead to the same conclusions. In the examples shown, this error did not affect the convergence of the algorithm towards the expected shape. The ellipsoidal model should fit most of the situations encountered in clinical practice if we consider that the tumors are issued from a very small region. The weakest issue is the simplicity of the system and image models which could be improved further in order to support a more accurate degradation and image model to be integrated in a comprehensive maximum a-posteriori (MAP) model framework at the cost of a larger computational burden. This algorithm works for small tumors, but obviously, should provide similar results for larger objects for which the computational cost would then increase dramatically.

V. References

[1] C. H. Chen, R. F. Muzic, A. D. Nelson, and L. P. Adler, "A nonlinear spatially variant object-dependent system model for prediction of partial volume effects and scatter in PET," IEEE Trans Med Imag, vol. 17, pp. 214-27, 1998.

[2] C. H. Chen, R. F. Muzic, A. D. Nelson, and L. P. Adler, "Simultaneous recovery of size and radioactivity concentration of small spheroids with PET data," J Nucl Med, vol. 40, pp. 118-30, 1999.

[3] J. T. Ashburner, "Computational Neuroanatomy," PhD Thesis Wellcome Department of Cognitive Neurology, University College London, 2000.

[4] R. Piessens, E. de Doncker-Kapenga, C. W. Uberhuber, and D. K. Kahaner, QUADPACK A subroutine package for automatic integration: Springer Verlag, 1983.

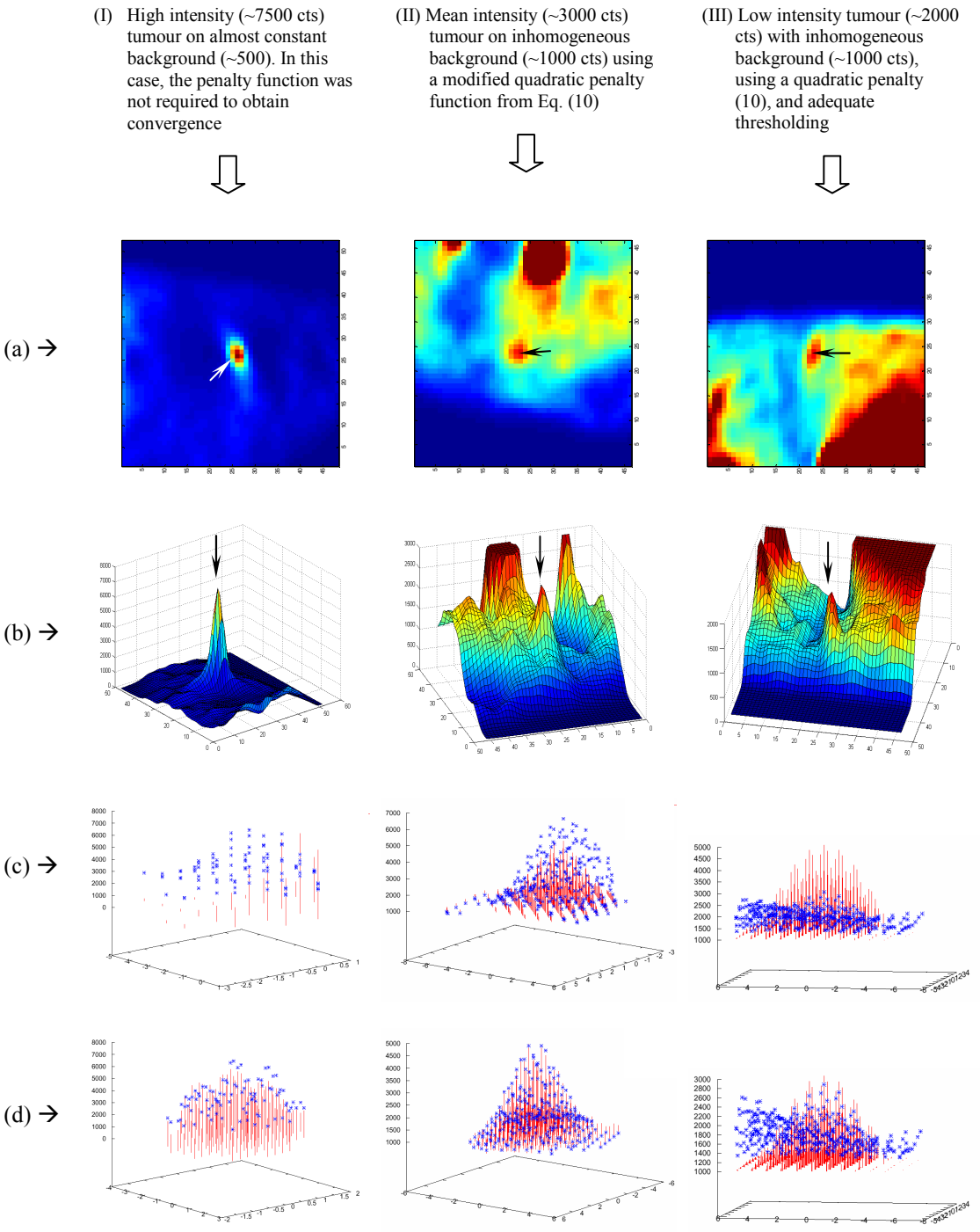

Fig. 3. Study of the algorithm convergence on reconstructed patient data (AW-OSEM reconstructed whole body PET images). (a) and (b) are representations of the tumor location and its neighborhood, (c) is the cropped and thresholded VOI-point-set drawn close to the initial estimates and (d) is the final obtained estimate. Bars depict the estimated model, and crosses are the VOI elements (thresholded). The most influent initial estimate is the center of the tumor as the model could escape the region, however, it is the easiest to determine automatically. For cases (II) and (III), lesion maximum intensity point were carefully set as start estimate, and only limited shifts around this value were allowed by setting the step sizes of the minimization algorithm.