

Acquisition Models in Intraoperative Positron Surface Imaging

Frédéric Monge · Dzhoshkun I. Shakir,
Ph.D · Florence Le Jeune, Ph.D, M.D. ·
Xavier Morandi, M.D. · Nassir Navab,
Ph.D · Pierre Jannin, Ph.D

Received: date / Accepted: date

Abstract Purpose: Intraoperative imaging aims at identifying residual tumor during surgery. Positron Surface Imaging (PSI) is one of the solutions to help surgeons in a better detection of resection margins of brain tumor, leading to an improved patient outcome. This system relies on a tracked freehand beta probe, using ^{18}F based radiotracer. Some acquisition models have been proposed in the literature in order to enhance image quality, but no comparative validation study has been performed for PSI.

Methods: In this study, we investigated the performance of different acquisition models by considering validation criteria and normalized metrics. We proposed a reference-based validation framework to perform the comparative study between acquisition models and a basic method. We estimated the performance of several acquisition models in the light of four validation criteria: efficiency, computational speed, spatial accuracy, and tumor contrast.

Results: Selected acquisition models outperformed the basic method, albeit with the real-time aspect compromised. One acquisition model yielded the best performance among all according to the validation criteria: efficiency (1-Spe: 0.1, Se: 0.94), spatial accuracy (max Dice: 0.77) and tumor contrast (max T/B: 5.2). We also found out that above a minimum threshold value of the sampling rate, the reconstruction quality does not vary significantly.

Frédéric Monge · Pierre Jannin · Xavier Morandi
LTSI INSERM, UMR 1099, Campus de Villejean, Université de Rennes 1
2, Avenue du Pr. Léon Bernard, Cedex 35043, Rennes, France
Tel.: +33(0)2-2323-4920
E-mail: frederic.monge@univ-rennes1.fr

Dzhoshkun I. Shakir
Translational Imaging Group / CMIC, University College London, United Kingdom

Florence Le Jeune
Centre Eugène Marquis, Rennes, F-35000, France

Xavier Morandi
CHU Rennes, Service de Neurochirurgie, Rennes, F-35000, France

Nassir Navab
CAMP, Technische Universität München, Munich, Germany

Conclusion: Our method allowed the comparison of different acquisition models and highlighted one of them according to our validation criteria. This novel approach can be extended to 3D datasets, for validation of future acquisition models dedicated to intraoperative guidance of brain surgery.

Keywords freehand beta probe · functional imaging · intraoperative imaging · neuronavigation · radiation detection physics · reconstruction · validation assessment

1 Introduction

According to the World Health Organization (WHO), gliomas are the most frequent primitive brain tumor, affecting 5 persons out of a 100,000 every year. The main treatment option is surgical resection with a careful consideration of the brain eloquent areas while achieving maximal resection, resulting in better outcome for the patient [8]. Survival prognosis from low grade glioma (LGG) is better when the maximum tumoral tissue has been removed, slowing down evolution into high grade glioma (HGG), which is much more aggressive. The extent of resection [14], correlated with the success of surgery, is assessed over postoperative MRI of the patient, to eventually detect recurrence.

In the last decade, several technologies have been developed to improve the completion of resection by allowing detection of residual tumor during surgery mainly for HGG. Intra-operative MRI yields good results towards maximum resection [7], better than conventional neuronavigation-guided glioblastoma (HGG) surgery [11]. Indeed, many limitations exist by using magnetic field devices in the operating room (OR) such as cost issues. 5-aminolevulinic acid (5-ALA) fluorescence image guided surgery is a relevant alternative to improve surgical resection outcome by giving a real-time visual histology and tumor boundaries [1]. Good results have been achieved in HGG [20]. Recently, intraoperative ultrasound (iUS) systems, performing LGG delineation, have been used to help the surgeon [21, 26] during surgery. Poor quality of images makes the interpretation subjective without training, especially for small or superficial residuals [4].

In oncology, Positron Emission Tomography (PET) is a gold standard for tumor detection and grading. Considering ^{18}F -FET (fluoro-ethyl-tyrosine) radiotracer, uptakes in brain tissues is better for detecting LGG by comparison with usual ^{18}F -FDG (fluoro-desoxy-glucose) [19]. The definition of tumor margins is more accurate with FET than FDG, according to the enhanced tumor contrast. Despite higher sensitivity and specificity in glioma diagnosis compared to MRI or CT scan [5], ^{18}F -PET scan suffers from a low spatial resolution (between 4 and 7 mm) and is rarely used in the intraoperative context. Moreover, since the gamma radiation (photons) has higher penetration ability than beta particles (positrons), if a gamma sensor is used, radiations emitted from surface or deeper tissues cannot be distinguished. In our case, beta particles is preferable because of its lower penetrating ability. Therefore, all the detected positrons are certainly emitted from the surface tissue (not deeper than 2 mm below). On the contrary, freehand nuclear beta probes, which are pen-sized devices that can detect positrons emitted by radiolabeled tissues, are more suitable in the OR for the detection of residual tumor over the resection cavity. In addition, a surgery performed with such device



Fig. 1: (a) Beta probe and its tracking target.

will be proceeded according to a normalized protocol considering radioprotection equipment for the surgical team, in agreement with the Nuclear Safety Agency. In this context, such guided surgery will be performed for LGG only, representing less than 10 tumors per year.

Freehand Positron Surface Imaging systems [16] have been introduced to generate functional images using a beta probe combined with a tracking system, as a neuronavigation workstation. Such a system is used to estimate the activity distribution of the radiotracer on 3D iso-surface images (mesh) to check resection margins. However, to perform good quality images from highly undersampled data, radiation detection physics must be modeled. In previous work, acquisition models have been proposed to address this problem, generating reconstruction images with improved spatial localization and tumor contrast. One of the previous studies provided a comparison between different acquisition models [6], pointing to the dramatic variations of reconstructions in SPECT (Single Photon Emission Computed Tomography). In Positron Surface Imaging (PSI), this variation between acquisition (or detection) models can lead to bias in the detection of residual tumor during surgery.

In this study, we propose a validation framework for comparing the performance of acquisition models. Section 2 presents the methodology of freehand PSI including activity reconstruction using detection models. We also introduce our validation framework and the acquisition models studied in this paper. Section 3 presents the results. Finally, section 4 discusses the implications of our results and concludes with an outlook on future work.

2 Materials and methods

2.1 Freehand PSI System

The proposed system consists of a freehand tracked beta probe [2]. After a scan of a radiolabeled surface of interest (SOI) with our system (see figure 2), each accumulated probe reading is synchronized with its mean corresponding pose in the reference coordinate system, giving one measurement during a time window (6D + 1D data). A sampling of the SOI is performed by generating a surface mesh with isotropic suxels¹, based on the probe positions and orientations. Thus,

¹ A suxel is defined as the smallest element of the SOI.

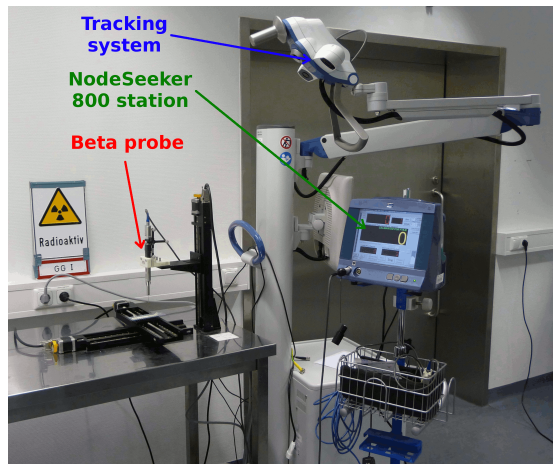


Fig. 2: A Positron Surface Imaging system: an optical tracking system (blue arrow), a beta probe (red arrow) and its workstation (green arrow).

measurements are used to determine an estimate of the surface activity distribution of the SOI (i.e. retrieve the activity value in each suxel of the surface: Positron Surface Imaging).

2.2 Surface Activity Reconstruction

Radiation detection models (or acquisition models) that allow for iterative reconstruction have been introduced and shown to substantially improve image quality in terms of spatial resolution and tumor-to-background (T/B) signal contrast [22, 23]. By taking conventions introduced in [6], the sampled SOI suxels $\mathbf{x} = (x_j)$ and the measurement vector $\mathbf{m} = (m_i)$ are considered here. Each stored measurement m_i can be written as a linear combination of the (unknown) activity values of the suxels x_j within its intersection with the FOV of the probe as follows:

$$m_i = \sum_{j=1}^N a_{ij} \cdot x_j \quad (1)$$

where coefficients a_{ij} are computed through an acquisition model, corresponding to the probability that the sensor at pose i has detected positron emissions from suxel x_j . The unknown reconstruction vector \mathbf{x} (i.e. the surface activity reconstructed, composed of N suxels) may be found by solving the system $\mathbf{m} = \mathbf{A} \cdot \mathbf{x}$, with $\mathbf{A} = (a_{ij})$ the system matrix (size $M * N$ with M the length of the measurement vector \mathbf{m}). Solutions of this system may be approximated using the Maximum Likelihood Expectation Maximization (MLEM) algorithm introduced by [25], to compute the pseudo inverse of the system matrix A .

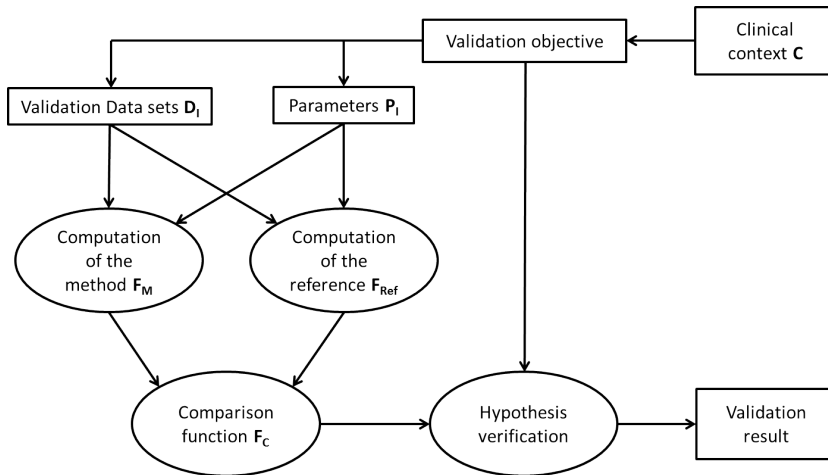


Fig. 3: Main steps of the reference-based validation protocol for medical image processing (figure reproduced from [10]): from a defined validation objective related to a clinical context C , results of the method to be assessed F_M applied on validation data sets D_I with parameters of the method P_I are compared with a reference through a comparison function F_C . Then, the final result is obtained by testing the output of F_C against the validation objective.

2.3 Reference-based assessment of reconstruction

In this section, we present our validation scheme for reconstructions generated from Positron Surface Imaging System. The need for a dedicated assessment scheme in this novel imaging approach drives us to develop a specific framework, according to the methodology proposed in [9]. Our focus is to propose a comparative study allowing the performance characterization of acquisition models and the identification of parameters which influence reconstruction outputs, for intraoperative use in neurosurgery. We focused on assessment levels 1 and 2, dedicated to technical parameters and clinical reliability setting respectively. “Biological phantom” datasets acquired in simulated OR conditions with a minimal setup are considered, where the clinical realism is close to a real surgery with very few controlled phantom parameters. Based on this methodology, assessed features concern the reconstruction efficiency, computation speed and tumoral characterization accuracy. Our validation protocol is based on [10], where a scheme for medical image processing validation procedure is described via a structured framework. It relies on the comparison between the results of a method and a reference, assumed to be close or equal to the correct solution. This scheme has been successfully used for designing and reporting medical image processing validation studies, such as in [12]. The main components of this reference-based validation scheme are presented in figure 3. The validation objective of our study was to evaluate detection performances and clinical relevance (V_C) of reconstructed images generated with acquisition model F_M , on in-vitro data sets in the context of neurosurgery, for the detection of resection margins (C). Following sections provide an instantiation of the framework, where each component is detailed in figure 4.

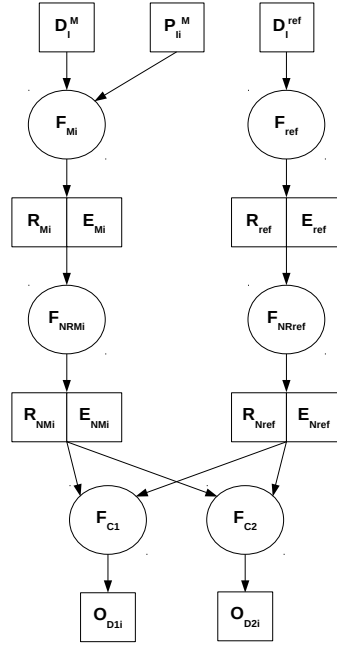


Fig. 4: Adapted reference-based validation process according to our evaluation context.

2.3.1 Validation data sets D_I

Our validation datasets consist of acquisitions with an in-vitro phantom as described in [22].

Bioluminescence images D_I^{ref} Cancer cells were supplied with luciferine (0.1 ml per dish) to perform imaging with a photographic camera of each Petri dish. We define D_I^{ref} the obtained 2D images considered as the reference to which reconstructions \hat{R}_{M_i} (see below) will be compared for each configuration (shown in figure 5, lower row). Petri dish 1 (figure 5d) : two high-contrast tumors. Petri dish 2 and 3 (figures 5e and 5f) : three tumors within regular contrast.

Raw positron surface data D_I^M Petri dish surfaces were scanned with a 3-axis step motor setup in 1 mm steps. For each position of the step motor, probe readings were accumulated for 3 seconds. Due to the acquisition frequency of the beta probe (≈ 10 Hz), each position is associated with radioactive counts in the time window (i.e. ≈ 30 probe readings during each 3 second acquisition). A long

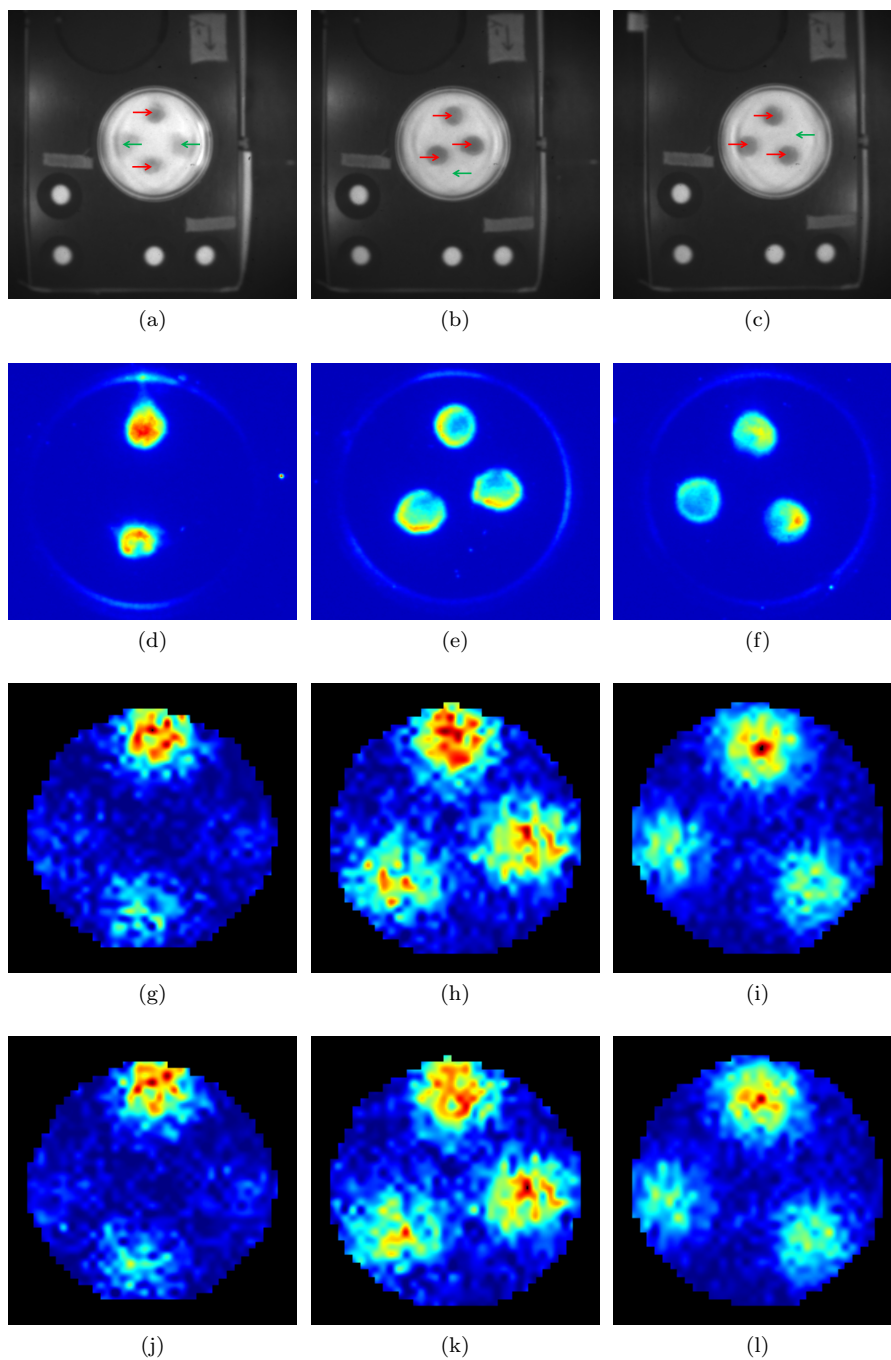


Fig. 5: Row 1: Brightfield images of the three Petri dish configurations, where right oriented red arrows show BCC grown locations (tumor) and left oriented green arrow show grown HFF locations (spot or homogeneous background), (a) Petri dish 1, (b) Petri dish 2 and (c) Petri dish 3. Row 2: Reference bioluminescence images D_I^{ref} of (d) Petri dish 1, (e) Petri dish 2 and (f) Petri dish 3. Row 3: reconstructed images with *projection* method for 3 second sampling rate (no undersampling). Row 4: reconstructed images with *projection* method for 1 second sampling rate.

acquisition time leads to good statistical data consistency in laboratory conditions, but this might not be feasible in intraoperative conditions. In order to observe the impact of lower acquisition times on data quality, raw data were undersampled by a selected factor (see input parameters section 2.3.5). Below, D_I^M is used to represent each configuration of multi-sampled raw positron surface data.

2.3.2 Surface activity reconstruction method F_{M_i}

To perform reconstruction, raw data D_I^M were sampled into a 2D rectangular grid (50 x 50, 1 mm isotropic pixels), for a fixed scanning depth value. Also, the measurement vector \mathbf{m} was created from D_I^M . Here, objective of reconstruction is to define each pixel value of the image, within the rectangular 2D grid. When the basic approach *activity projection* (proj) is considered, pixel values are defined using a nearest neighbor interpolation over measurement vector \mathbf{m} for each position in the grid. By contrast, when acquisition models (presented below) are used to estimate matrix A , reconstruction is obtained through equation $\mathbf{x} = A^{-1} \cdot \mathbf{m}$, with \mathbf{x} the reconstruction vector, \mathbf{m} the measurement vector and A^{-1} the pseudo inverse of the system matrix. The MLEM algorithm with a selected number of iterations (see section 2.3.5) was used to compute its pseudo inverse. \hat{R}_{M_i} represents the image from the reconstruction vector \mathbf{x} estimated with method F_{M_i} .

Solid angle model (SA) As baseline to define each contribution of suxel x_j to measurement m_i , the *solid angle* model (SA) introduced by [17] considers the geometrical attenuation based on the relative position and orientation of the probe to the source. An implementation of the SA model, used in [22] and defined in equation 2, computes each attenuation coefficient a_{ij} , using the following parameters: d_{ij} the distance between the detector and the suxel x_j for measurement m_i (probe to source distance), r the detector radius and α_{ij} the tilt angle between the detector axis and the detector to source axis (see figure 6).

$$a_{ij} = f(d_{ij}, \alpha_{ij}) = \frac{1}{2} \cdot \left(1 - \frac{1}{\sqrt{(r^2/d_{ij}^2) + 1}} \right) \cdot \cos \alpha_{ij} \quad (2)$$

Look-up-table model (LUT) In the previous analytical model, each coefficient of the system matrix A is directly computed from the measurement vector m_i . An alternative solution consists in selecting precomputed a_{ij} coefficients stored in a *look-up table* (LUT). As proposed in [24], attenuation coefficients a_{ij} are defined by equation 3 for respective LUT positions in the FOV.

$$a_{ij} = L(i, j) \quad (3)$$

Here, L refers to the LUT stored in memory and the position in the FOV is represented by the pair (i, j) , shown in figure 7a. Implicitly, a mapping function is used to relate the source-to-detector distance d_{ij} and the tilt angle α_{ij} to the LUT coefficients. This model is based on coefficients stored in memory.

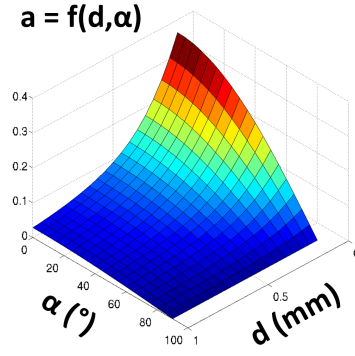


Fig. 6: Plot of the solid angle model function $a_{ij} = f(d, \alpha_{ij})$ with d the probe to source distance and α_{ij} the tilt angle.

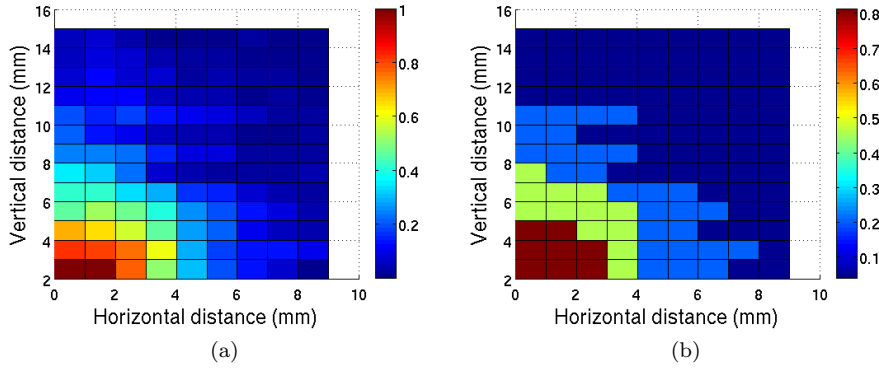


Fig. 7: (a) Representation of the LUT proposed in [24], where a_{ij} are represented by a color-map, according to their position in the FOV. Here, distance from center of the source (X axis) and depth from the source to the detector (Y axis) are defined through d_{ij} and α_{ij} , (b) Partition model generated from $k = 4$ clusters with a_p coefficients represented along distance from center of the source (X axis) and depth from the source to the detector (Y axis): a_1 in dark red, a_2 in green, a_3 in blue and a_4 in dark blue.

Partition model (PM) Although the LUT model yields good results in defining structures better than the SA model when comparing with a groundtruth [24], looking up the a_{ij} in the LUT for each (i, j) pairs involves additional computational time for defining the whole system matrix A ; possibly using nine neighbor averaging interpolation. A *partition model* (PM) is proposed to overcome this constraint, based on the division of the FOV into sectors where particular homogeneous coefficients a_k are assigned for each partition k [15]. PM model is analytically defined in equation 4 and example of a partitioning scheme is shown in figure 7b. Here, a_k factors are defined through a k-means algorithm performed over the LUT, where the number of cluster k corresponds to the wanted number of partition k . Thereby,

values of the LUT defined within each cluster k are averaged to obtain a_k attenuation coefficients. In the same way as the LUT model, a_k coefficients are assigned according to respective (i, j) positions in the FOV.

$$a_{ij} = \begin{cases} a_k & \text{source within partition } k, \\ 0 & \text{source not within any partition.} \end{cases} \quad (4)$$

2.3.3 Error \hat{E}_{M_i} of reconstructions

The reconstruction process involves estimating the pseudo-inverse of the matrix A (computed via acquisition model F_{M_i}) by the MLEM iterative method. Thus, reconstruction error \hat{E}_{M_i} , correlated with iteration number, was defined as the norm of residual given $\hat{E}_M = \|A.\hat{c} - g\|$ with \hat{c} the final estimated reconstruction vector, A the coefficient matrix and g the probe reading vector.

2.3.4 Reference method F_{ref} and associated error \hat{E}_{ref}

In oncology, presence of residual tumor is generally observed with the radiotracer uptake on functional PET images, combined with a CT scan or an MRI when available (high performance) for better localization using co-registered images (CT in this case is primarily needed for attenuation correction). However, this modality requires long acquisition time (30 min. for a PET scan). In this study, we used bioluminescence imaging as a reference for determining the tumor location in Petri dishes [18], without affecting initial cell growth behavior [27]. This method gives high resolution functional images, where accuracy is assumed better than reconstruction with methods F_{M_i} . Reference images were obtained by cropping bioluminescence images D_I^{ref} following a square (size 180 x 180 pixels), fitted and tangential to the Petri dish edges. The symbol \hat{R}_{Ref} is used to represent reference bioluminescence images for each configuration. In addition, the associated error \hat{E}_{ref} is related to the loss of background signal in bioluminescence image (absence of peri-tumoral contrast, see figure 5), which cannot be quantified.

2.3.5 Input parameters P_I

In this section, parameters that can possibly influence the performance of reconstruction are elaborated. According to the model of reference-based validation [10], parameters related to data acquisition and the reconstruction method are considered.

Sampling raw data D_I^M The variation of this data-related parameter allows the observation of data quality over different simulated acquisition times. The range [0.1; 3] seconds has been selected for simulating scans between a quick scan (0.1 sec per position) and a more detailed one (3 sec per position), corresponding to 3% and 100% of probe readings considered per position, respectively. In figure 5, row 3 and row 4 show reconstructed images generated with the *projection* method using 3 sec. and 1 sec. acquisitions, respectively.

MLEM iteration number According to [6], reconstruction performed with large number of iteration (e.g. 100) increases artifact formation, in SPECT imaging systems. The variation of the number iteration in the range [1; 50] allows us to observe the influence of this method-related parameter on reconstruction.

Number of clusters in PM Finally, in the *partition model* PM proposed to reduced computational power by dividing the LUT FOV, according to a k-means algorithm. The variation of the number of clusters k in the range [2; 6], allows us to observe influence of this method-related parameter on the computation time in comparison with other reconstruction methods. In the following, we use the symbol PM- k (PM-2, PM-3, ..., PM-6).

2.3.6 Normalization function of reconstruction F_{NRM_i}

In order to compare reconstructions with their respective reference \hat{R}_{Ref} , \hat{R}_{M_i} images are resized to fit reference image spatially (180 x 180 pixels) and amplitudes are normalized within the range [0; 1]. Thus \hat{R}_{NM_i} represents \hat{R}_{M_i} images normalized through F_{NRM_i} function and \hat{E}_{NM_i} represents visual blurring errors due to the interpolation method used in the resizing step.

2.3.7 Normalization function of reference F_{NRref}

Due to the different modalities used for generating the validation datasets D_I^{ref} and D_I^M , a scale shift exists between bioluminescence reference images \hat{R}_{Ref} and reconstructions \hat{R}_{M_i} , respectively acquired with a camera and the PSI system. For each Petri dish configuration, the reconstructed image \hat{R}_{M_i} generated through the *projection* method F_{M_i} , parametrized with the minimum sampling rate (3 sec per position, no undersampling) is considered. The visual scaling shift was evaluated by overlapping reference \hat{R}_{ref} with its respective considered reconstruction. Experimentally, a scale of 1.1 is observed for reconstructed image by comparison with bioluminescence images, in each case (see figure 8, upper row). Then, a 2D affine spatial transformation (with 1.1 scaling parameter) was applied to cropped bioluminescence images (figure 8 lower row). A binarization step was performed by thresholding images with a determined level. Heuristically, a 0.3 level value gives the global size-position (see figure 9). Finally, remaining Petri dish edges were removed with a morphological opening operation. \hat{R}_{Nref} represents segmented bioluminescence images in the SOI coordinate system obtained with both level values, and \hat{E}_{Nref} represents the qualitative error corresponding to the possible mismatch between Petri dish edges (shown in figure 8, lower row).

2.3.8 Validation criterion and comparison function F_C

From normalized outputs \hat{R}_{Nref} and \hat{R}_{NM_i} , we estimated the following validation criterion: reconstruction efficiency, computation speed, spatial accuracy and tumor characteristics (contrast and size).

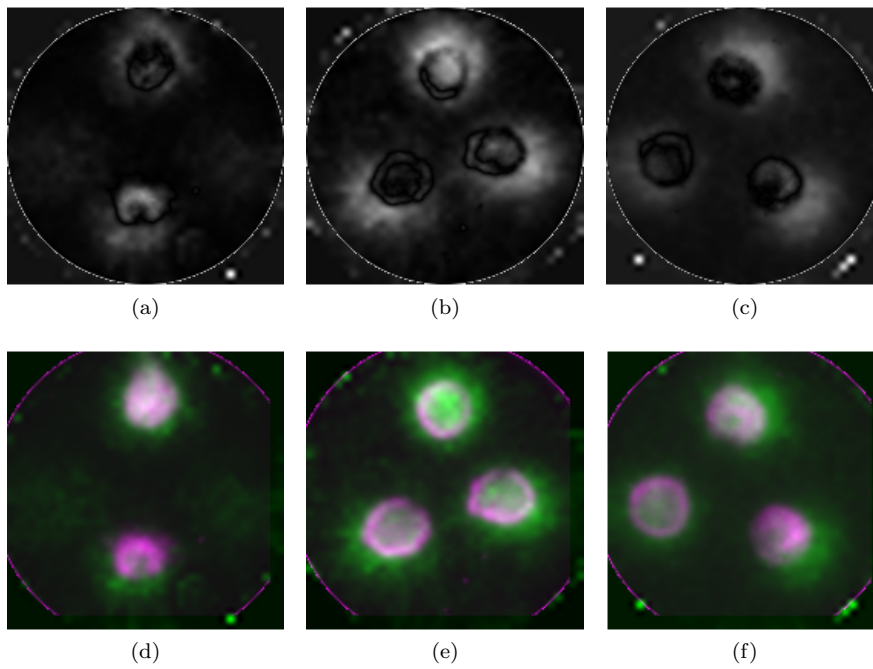


Fig. 8: Upper row: Difference of cropped bioluminescence and raw images, (a) Petri dish 1, (b) Petri dish 2 and (c) Petri dish 3. Lower row: bioluminescence images in the SOI coordinate system \hat{R}_{ref} (d) Petri dish 1, (e) Petri dish 2 and (f) Petri dish 3.

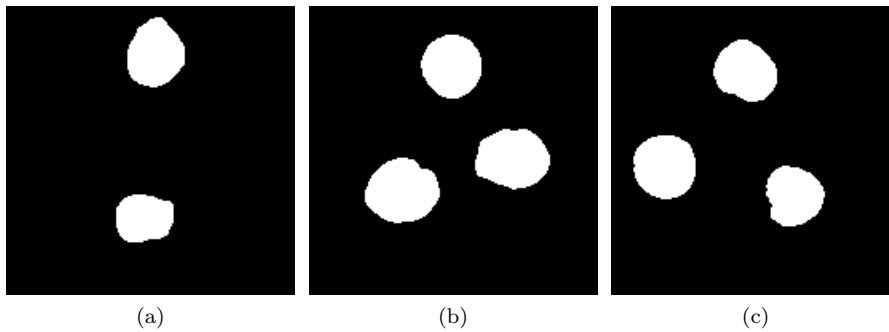


Fig. 9: Normalized reference, giving global size and position of tumors. (a) Petri dish 1, (b) Petri dish 2, (c) Petri dish 3.

Efficiency For each reconstructed image, compared to reference image, we calculated the number of false negative (FN), false positive (FP), true positive (TP) and true negative (TN) pixels. Then, sensitivity-specificity curves were computed using these metrics. Sensitivity has been defined as the true positive rate ($TP/(TP + FN)$), which represents the ability of our system to detect tumor when a tumor is present in the reference. Specificity has been defined as the true negative rate ($TN/(TN + FP)$), which represents the ability of our system to detect a tumor when a tumor is not present in the reference. Efficiency of reconstruction $O_D = F_C(\hat{R}_{Nref}, \hat{R}_{NM_i})$ can be quantified with the optimal point observed on sensitivity-specificity curves (i.e. optimal point obtained by maximizing sensitivity and specificity).

Computation speed The computation time of each reconstruction performed with acquisition model was determined. The speed performance is a relevant criterion for an efficient intraoperative use.

Spatial accuracy Correlation of reconstructions \hat{R}_{NM_i} to respective reference \hat{R}_{Nref} were estimated. Due to the nature of \hat{R}_{NM_i} and \hat{R}_{Nref} outputs, the Dice coefficient (DSC) was used as comparison function F_C , shown in equation 5.

$$DSC(A, B) = \frac{2 |A \cap B|}{|A| + |B|} \quad (5)$$

Tumor contrast Similarly with PET scan images, the functional information over reconstructions is provided by the tumor to background ratio (T/B). From normalized reconstructed images \hat{R}_{NM_i} , mean T/B was computed over normalized reconstruction \hat{R}_{NM_i} using normalized reference \hat{R}_{Nref} as binary mask. Based on the boolean values in the mask definition, each pixel with a 'true' value was considered part of the tumor region of interest (tROI) and each 'false' value part of the background region of interest (bROI). Each ROI was averaged by its number of pixels and thus T/B ratio was computed, defined as result O_D .

2.3.9 Quality index O_{QI} and hypothesis testing F_H

From comparison functions F_C defined above according to the validation criteria, a quality index can be deduced for characterizing the overall performance of acquisition models. In each sensitivity-specificity curve, Area Under the Curve (AUC) was computed for each acquisition model using the trapezoidal numerical integration method (F_{QI}), representing the quality index O_{QI} . In addition, no statistical evaluation F_H was performed in this scheme, due to insufficient amount of data. Visual comparisons have been done between reconstruction and related reference.

3 Results

Two independent studies were performed using comparison function F_C along two input parameters P_I . The first comparison was made by varying the sampling rate while fixing the number of MLEM iteration at 25, an empirical value used in previous work on PSI systems [23, 24]. The second one was made by varying

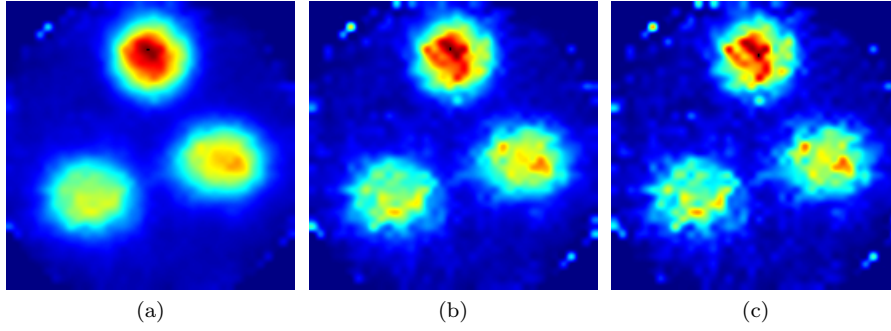


Fig. 10: Reconstructed images varying MLEM iteration number with SA model for sampling time fixed at 3 seconds, in Petri dish 2. (a) 10 iterations, (b) 30 iterations, (c) 50 iterations.

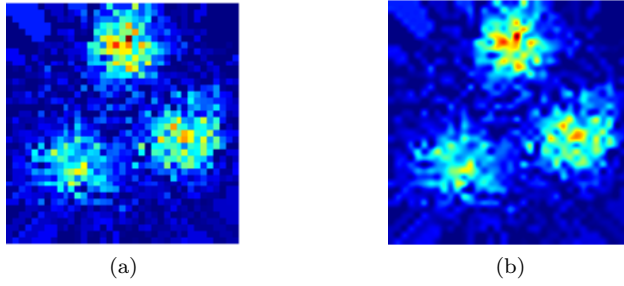


Fig. 11: Blurring effect of normalization function F_{NRM_i} , example with projection method, optimal input parameters P_I on Petri dish 2 (a) Reconstructed image \hat{R}_{M_i} (50 x 50 pixels), (b) Normalized reconstructed image \hat{R}_{NM_i} (180 x 180 pixels).

number of iterations of the MLEM algorithm against a constant sampling time of 3 seconds, corresponding to the initial acquisition time of data. In figure 10, exemplary reconstructed images with the SA model are shown for Petri dish 2, for various numbers of MLEM iterations.

Reconstruction errors, described in section 2.3.3, were obtained using the MLEM method for each acquisition model. Error values reached a range of $[10^{-7}; 10^{-6}]$ by varying the sampling rate and the number of MLEM iterations respectively. The loss of peri-tumoral tissue signal in bioluminescence images, represented by error \hat{E}_{Ref} , is observed in figure 5 lower row. This total lack of background signal was due to the bioluminescence process, which shows tumoral tissues only. Error \hat{E}_{NM_i} represents the blurring effect of the normalization function F_{NRM_i} applied on reconstructions. An example is shown in figure 11 for the projection method in the Petri dish 2.

Misalignment error \hat{E}_{NRef} generated by the normalization function of reference F_{NRref} is observed over figure 8 lower row. Results show a satisfying match between Petri dish edges from reconstruction \hat{R}_{M_i} and from reference \hat{R}_{Ref} , in each

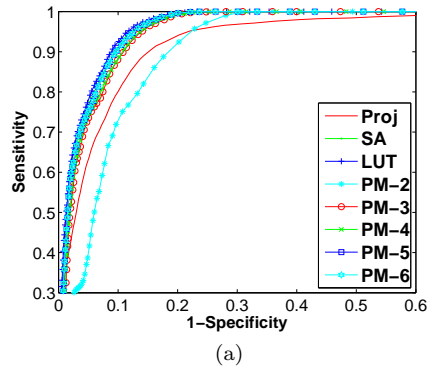


Fig. 12: Mean ROC curves of reconstruction methods for optimal sampling and MLEM iteration number parameters.

Table 1: Optimal points extracted from mean ROC curves, computed for the different Petri dish configurations.

Method	Proj	SA	LUT	PM-k ³				
				2	3	4	5	6
1-Spe ¹	0.16	0.10	0.10	0.25	0.13	0.13	0.11	0.11
Se ²	0.9	0.93	0.94	0.95	0.93	0.94	0.93	0.93

¹Specificity. ²Sensitivity. ³Partition model with $k \in [2; 6]$ clusters selected.

configuration. Visually, performance of acquisition models can be observed over sensitivity-specificity curves derived from reconstructions. In figure 12, mean ROC (Receiver Operating Characteristic) curves over Petri dish configurations were defined by considering optimal sampling and MLEM input parameters P_I (MLEM = 25 iterations and sampling = 3 seconds). In ROC representation, the best curve is found by minimizing the 1-specificity term while maximizing the sensitivity. Values of optimal points are presented in table 1. Here, two acquisition models outperformed the *projection* method: *SA* and *LUT* models.

In figures 13a and 13b, AUC values are shown for ROC curves varying the sampling rate and MLEM iterations respectively. We observed that the projection method results are nearly constant along the sampling rate (AUC around 0.74) and invariant along the number of MLEM iterations. From 1-sec sampling onwards, each acquisition model tends to 0.77 mean AUC value, except for method *PM-2* (mean AUC = 0.67). Moreover, mean AUC values tend to decrease linearly along MLEM iterations, with a steeper decrease when using the *PM-2* model: from a 0.86 starting AUC value for one MLEM iteration down to 0.79 around for 50 MLEM iterations. Globally, both AUC curves emphasize three efficient and stable acquisition models by comparison with the *projection* method: *SA*, *LUT* and *PM-5* models.

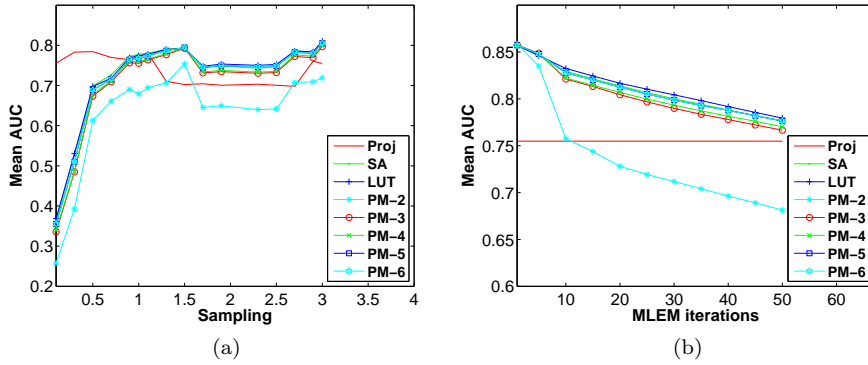


Fig. 13: Mean AUC values along input parameters P_I : (a) Varying sampling rate and MLEM = 25 iterations, (b) Varying MLEM iterations, sampling = 3 seconds.

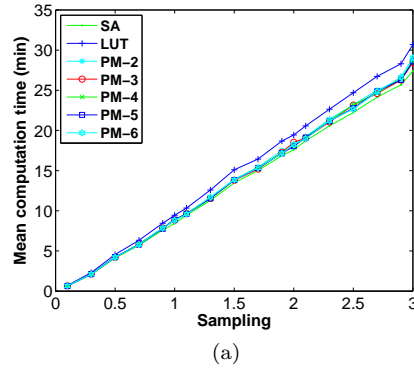


Fig. 14: Mean computation time observed for reconstruction methods varying sampling rate.

Considering computation time required by the *projection* method (3 ms independently of sampling rate), the reconstruction process was hardly time consuming. This computation time of reconstruction methods is linearly dependent on the sampling rate of data, as shown in figure 14. The *LUT* was the most time consuming method with a maximum of 31 minutes observed. Other methods reached a maximum around 28 minutes of computation required to perform reconstructions.

The quantification of spatial accuracy of each reconstruction method is given by the Dice score metric. In figure 15a, the *projection* method yields the most stable Dice score, around 0.6, overtaking the *PM-2* model. Starting with the 0.5 sec sampling value, all other reconstruction models outperformed the *projection* method, where higher Dice values are obtained with the *LUT* model. The mean Dice score variation along MLEM number of iterations is shown in figure 15b. Results show a slight variation of the behavior of the Dice value for each reconstruction method

starting with 5 MLEM iterations. Two methods are emphasized here: *LUT* and *SA* models.

Finally, variations of the tumor contrast along the sampling rate and number

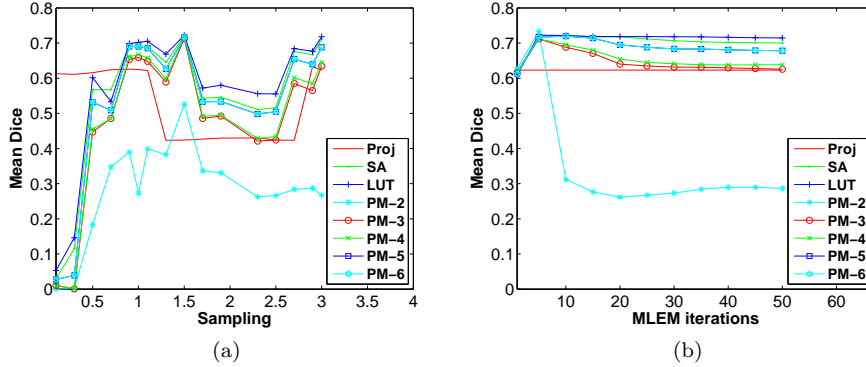


Fig. 15: Mean Dice values along input parameters P_I : (a) Varying sampling rate and MLEM = 25 iterations, (b) Varying MLEM iterations, sampling = 3 seconds.

of MLEM iterations are shown in figure 16a and 16b respectively. The *projection* method yields constant mean T/B contrast rate over varying sampling rates or number of MLEM iterations. For reconstruction methods, T/B values outperformed the *projection* method, increasing slightly with the sampling, starting at 0.5 (figure 16a). On the contrary, mean T/B value decreases with increasing MLEM number of iterations (figure 16b), but remains better than the *projection* method. *LUT* and *SA* models yields the highest T/B values for a sampling rate of 3 combined with 5 MLEM iterations.

4 Discussion and conclusion

4.1 Discussion

We have compared detection models by assessing respective reconstructed images from a chosen dataset, relying on a validation framework. We have outlined benefits and drawbacks of each acquisition model according to selected input parameters P_I (i.e. section 2.3.5) and validation criteria (i.e. section 2.3.8). Images reconstructed using acquisition models yield global detection improvement along chosen comparison functions and quality index (i.e. section 2.3.9) with respect to the basic *projection* method.

According to the clinical applicability involved here in the selection of an adequate acquisition model, the trade-off between reconstruction quality and reconstruction time has been outlined and influences this selection. The quality of reconstruction has impact on decision making of the surgeon on the detection of residual tumor. Reconstruction time is important for respect of the surgical workflow. Additional

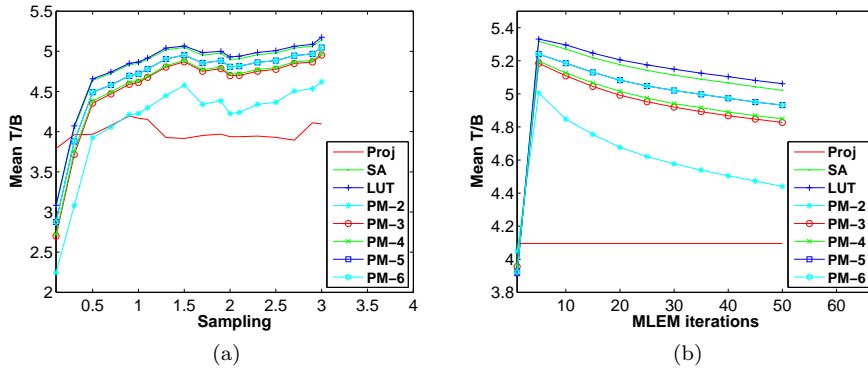


Fig. 16: Mean T/B values along input parameters P_7 : (a) Varying sampling rate and MLEM = 25 iterations, (b) Varying MLEM iterations, sampling = 3 seconds.

time for PSI acquisition needs be as short as possible in order to avoid potential complications or infections [3].

Our previous work [22], involved long acquisition times. Although long acquisitions provide statistically good data, it is not realistic to expect such data in clinical conditions. In addition, acquisition time and scan quality depend on the operator (i.e. neurosurgeon). In this work, we simulated the acquisition time variations based on the sampling rate input parameter. Results show that this variation slightly affects data quality. However, transient states can be observed in figures 13a, 15a and 16a in the range of [0.1; 0.7] seconds before reaching a stable value. This, suggests a threshold value of 0.7 second per location that should be considered as a measure of the data acquisition speed. A sampling rate above this value does not improve the data quality significantly, but still increases acquisition times.

Empirically, 25 iterations for the MLEM reconstruction algorithm were fixed to perform reconstructions with freehand PSI systems [23, 24], close to values used in freehand SPECT systems [13, 28] (i.e. 20 iterations). We investigated this value by varying it as an input parameter and looking at the respectively generated reconstructions. Increasing the MLEM number of iterations has slightly digressive effects on validation criteria. Peaks were observed at five iterations in figures 13b, 15b and 16b. Based on our experiments, we propose to use this value in future experiments for optimal validation criteria values.

Intraoperative use involves high performance in terms of computational speed. One factor that affects it is the sampling rate, which is directly correlated with the amount of data acquired per location. In figure 14, the linearity between the computational speed and sampling rate can be observed. Another factor that affects computational speed is the choice of the acquisition model. In this context, the *partition* model was introduced as a faster *LUT* model. In this work, we have proposed using a range of k clusters in the validation of the *partition* model (where similar reconstruction results as the off-the-shelf *LUT* have been expected). In figure 14, each generated PM- k model falls under the *LUT* experimental curve, i.e. is better in terms of computational speed. Results given by each PM- k were slightly faster than other methods, especially with the PM-2 method.

According to our experiments, one acquisition model outperformed others: the *LUT* model. This model offers the best result in terms of efficiency (1-Spe: 0.1; Se: 0.94), spatial accuracy (max Dice: 0.77) and tumor contrast (max T/B: 5.2), but requires the highest computational speed among all models (max computation time: 31 min). However, considering the stability of data quality over the sampling time threshold (i.e. 0.7 second per location), the *LUT* model seems appropriate for intraoperative use. Moreover, the current *LUT* model developed in [24] has been generated from planar acquisition. In real cases, the SOI would represent the resection cavity as a 3D surface. A realistic model would consider the inclination of the probe to the surface axis, the tilt angle.

In our validation framework, qualitative errors have been identified in the output of the reference method F_{Ref} and the normalization functions F_{NRM_i} and F_{NRref} . Such errors inducing possible bias in the results of our comparison functions F_C should be taken into account. According to [10], bias can be studied with respect to two criteria: accuracy and metric realism. Firstly, \hat{E}_{Ref} has been identified as the loss of background signal in bioluminescence images. No difference can be observed between inner and outer contrast of the Petri Dish (background in blue, figure 5, row 2) by comparison with reconstructed images (outer background in red and inner background in blue, figure 5, row 3 and 4). Thus, bioluminescence images suffer from a low realism due to radiotracer uptake in both tumoral and healthy tissue. In terms of accuracy, this can induce a few errors in the ROC curve creation for the efficiency validation criteria, without impacting the global behavior of curves. Secondly, \hat{E}_{NM_i} corresponds to the blurring effect of the interpolation performed on reconstructions. The smaller the size of the reconstructed images, the more the blurring effect increases, leading to a high bias in comparison functions. A possible impact of this is the overestimation of residual tumor size. And thirdly, \hat{E}_{NRref} gives the overall idea of the possible misalignment after the normalization method of references. In a real case, the imaging devices used for the different modalities that give the actual acquisition and the groundtruth reference, do not have the same coordinate system. Here, since edges of Petri dish are visible in both images (reconstruction and bioluminescence), it is possible to perform circle fitting to estimate the difference between the radius of the actual and reference images, which could subsequently be used as an estimate of the misalignment error \hat{E}_{NRref} . While a Petri dish radius larger in reference than in reconstruction will lead to an underestimation of the real tumor size, a larger radius in reconstruction than in reference will lead to an overestimation.

We have introduced several validation criteria for selecting an optimal acquisition model. They are strongly dependent to the nature of the reference dataset. The Dice score metric is adapted for validating the global localization of tumor spots with a reference extracted for an imaging modality of a different nature than the one used (i.e. bioluminescence imaging). In contrast, for a reference from an emission-based imaging modality (PET, SPECT, etc.), the normalized cross correlation metric (NCC) would give more information in checking resection margins. In addition, the tumor size over reconstruction (in cm^2) would provide information on the under/over estimation of the real tumor size in reference.

4.2 Conclusion

For the first time, we have compared acquisition models, dedicated to Positron Surface Imaging based on a dedicated reference-based validation scheme. Associated to our input parameters, new validation criteria can be integrated, considering the nature of reference datasets. Combined with AUC values, quality criteria such as Youden index can also be used for the characterization of ROC curves in terms of detection. In future work, the validation protocol will be extended to 3D realistic reference datasets (e.g. realistic neurosurgical phantom), as a necessary validation step involving realistic conditions for clinical use in neurosurgery.

Acknowledgments

The authors would like to thank CAMPUS FRANCE (PROCOPE Program), Rennes Métropole and Le Ministère des Affaires étrangères et du Développement international for French-Bavaria student exchanges and La Ligue Contre le Cancer for financial support.

Compliance with Ethical Standards

Funding: This study was funded by Campus France (FR), Rennes Métropole (FR), Le Ministère des Affaires Étrangères (FR) and la Ligue Contre le Cancer (FR).

Conflict of Interest: The authors declare that they have no conflict of interest.

Ethical approval: This article does not contain any studies with human participants or animals performed by any of the authors.

Informed consent: This articles does not contain patient data.

References

1. Behbahania M, Martirosyan NL, Georges J, Udovich JA, Kalani MYS, Feuerstein BG, Nakaji P, Spetzler RF, Preul MC (2013) Intraoperative fluorescent imaging of intracranial tumors: A review. *Clinical neurology and neurosurgery* 115(5):517–528
2. Daghighian F, Mazziotta JC, Hoffman EJ, Shenderov P, Eshaghian B, Siegel S, Phelps ME (1994) Intraoperative beta probe: a device for detecting tissue labeled with positron or electron emitting isotopes during surgery. *Medical physics* 21(1):153–157
3. Daley BJ, Cecil W, Clarke PC, Cofer JB, Guillamondegui OD (2015) How slow is too slow? correlation of operative time to complications: an analysis from the tennessee surgical quality collaborative. *Journal of the American College of Surgeons* 220(4):550–558
4. Gerganov VM, Samii A, Giordano M, Samii M, Fahlbusch R (2011) Two-dimensional high-end ultrasound imaging compared to intraoperative mri during resection of low-grade gliomas. *Journal of Clinical Neuroscience* 18(5):669–673

5. Grosu A, Weber W (2010) Pet for radiation treatment planning of brain tumours. *Radiotherapy and Oncology* 96(3):325
6. Hartl A, Shakir DI, Lasser T, Ziegler SI, Navab N (2015) Detection models for freehand spect reconstruction. *Physics in medicine and biology* 60(3):1031
7. Hata N, Muragaki Y, Inomata T, Maruyama T, Iseki H, Hori T, Dohi T (2005) Intraoperative tumor segmentation and volume measurement in mri-guided glioma surgery for tumor resection rate control. *Academic radiology* 12(1):116–122
8. Ius T, Isola M, Budai R, Pauletto G, Tomasino B, Fadiga L, Skrap M (2012) Low-grade glioma surgery in eloquent areas: volumetric analysis of extent of resection and its impact on overall survival. a single-institution experience in 190 patients: clinical article. *Journal of neurosurgery* 117(6):1039–1052
9. Jannin P, Korb W (2008) Assessment of image-guided interventions. In: *Image-Guided Interventions*, Springer, pp 531–549
10. Jannin P, Grova C, Maurer Jr CR (2006) Model for defining and reporting reference-based validation protocols in medical image processing. *International Journal of Computer Assisted Radiology and Surgery* 1(2):63–73
11. Kubben PL, ter Meulen KJ, Schijns OE, ter Laak-Poort MP, van Overbeeke JJ, Santbrink Hv (2011) Intraoperative mri-guided resection of glioblastoma multiforme: a systematic review. *The lancet oncology* 12(11):1062–1070
12. Maier-Hein L, Groch A, Bartoli A, Bodenstedt S, Boissonnat G, Chang PL, Clancy N, Elson D, Haase S, Heim E, Hornegger J, Jannin P, Kennigott H, Kilgus T, Müller-Stich B, Oladokun D, Röhl S, dos Santos T, Schlemmer HP, Seitel A, Speidel S, Wagner M, Stoyanov D (2014) Comparative validation of single-shot optical techniques for laparoscopic 3-d surface reconstruction. *Medical Imaging, IEEE Transactions on* 33(10):1913–1930
13. Matthies P, Gardiazabal J, Okur A, Vogel J, Lasser T, Navab N (2014) Mini gamma cameras for intra-operative nuclear tomographic reconstruction. *Medical image analysis* 18(8):1329–1336
14. McGirt MJ, Chaichana KL, Attenello FJ, Weingart JD, Than K, Burger PC, Olivi A, Brem H, Quinoñes-Hinojosa A (2008) Extent of surgical resection is independently associated with survival in patients with hemispheric infiltrating low-grade gliomas. *Neurosurgery* 63(4):700–708
15. Monge F, Shakir DI, Navab N, Jannin P (2016) Partition-based acquisition model for speed up navigated beta-probe surface imaging. In: *SPIE Medical Imaging, International Society for Optics and Photonics*, pp 97,862P–97,862P, location of the conference: San Diego, CA, USA
16. Navab N, Traub J, Wendler T, Buck A, Ziegler S (2008) Navigated nuclear probes for intra-operative functional imaging. In: *Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008. 5th IEEE International Symposium on*, IEEE, pp 1395–1398
17. Oezguer C, Bieniarz J, Lasser T, Ziegler S, Navab N, Wendler T (2009) Phenomenological models for intraoperative positron emission surface imaging using handheld probes. In: *World Congress on Medical Physics and Biomedical Engineering, September 7-12, 2009, Munich, Germany*, Springer, pp 213–216
18. O’Neill K, Lyons SK, Gallagher WM, Curran KM, Byrne AT (2010) Bioluminescent imaging: a critical tool in pre-clinical oncology research. *The Journal of pathology* 220(3):317–327

19. Pauleit D, Stoffels G, Bachofner A, Floeth FW, Sabel M, Herzog H, Tellmann L, Jansen P, Reifenberger G, Hamacher K, Coenen HH, Langen KJ (2009) Comparison of 18f-fet and 18f-fdg pet in brain tumors. *Nuclear medicine and biology* 36(7):779–787
20. Roessler K, Becherer A, Donat M, Cejna M, Zachenhofer I (2012) Intraoperative tissue fluorescence using 5-aminolevulinic acid (5-ala) is more sensitive than contrast mri or amino acid positron emission tomography (18f-fet pet) in glioblastoma surgery. *Neurological research* 34(3):314–317
21. Selbekk T, Brekken R, Indergaard M, Solheim O, Unsgard G (2012) Comparison of contrast in brightness mode and strain ultrasonography of glial brain tumours. *BMC Medical Imaging* 12(1):11, DOI 10.1186/1471-2342-12-11
22. Shakir DI, Navab N, Ziegler SI (2010) Acquisition model for iterative reconstruction of navigated beta-probe surface images. In: *Proceedings of IEEE Nuclear Science Symposium and Medical Imaging Conference (IEEE NSS-MIC)*, Institute of Electrical and Electronics Engineers, Knoxville, TN, USA
23. Shakir DI, Hartl A, Navab N, Ziegler SI (2011) Evaluation of an ad hoc model of detection physics for navigated beta-probe surface imaging. In: *SPIE Medical Imaging, International Society for Optics and Photonics*, pp 796,405–796,405, location of the conference: Orlando, FL, USA
24. Shakir DI, Hartl A, Schneider FR, Pulko J, Ziegler SI, Navab N, Lasser T (2012) Two new ad-hoc models of detection physics and their evaluation for navigated beta probe surface imaging. In: *SPIE Medical Imaging, International Society for Optics and Photonics*, pp 83,162G–83,162G
25. Shepp LA, Vardi Y (1982) Maximum likelihood reconstruction for emission tomography. *Medical Imaging, IEEE Transactions on* 1(2):113–122
26. Šteňo A, Karlík M, Mendel P, Čík M, Šteňo J (2012) Navigated three-dimensional intraoperative ultrasound-guided awake resection of low-grade glioma partially infiltrating optic radiation. *Acta neurochirurgica* 154(7):1255–1262
27. Tiffen JC, Bailey CG, Ng C, Rasko JE, Holst J (2010) Luciferase expression and bioluminescence does not affect tumor cell growth in vitro or in vivo. *Molecular cancer* 9(1):299
28. Wendler T, Hartl A, Lasser T, Traub J, Daghighian F, Ziegler SI, Navab N (2007) Towards intra-operative 3d nuclear imaging: reconstruction of 3d radioactive distributions using tracked gamma probes. In: *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2007*, Springer, pp 909–917