Sensor optimization for fluorescence optical tomography by experimental design methods

M. Freiberger[*]   H. Scharfetter[*]

Abstract

With the increasing importance of molecular imaging fluorescence based methods are continuously gaining im- pact. In fluorescence optical tomography excitation light is injected into the tissue where the fluorophore converts it to radiation of another wavelength. From the emitted light reaching the boundary the 3-D distribution of the fluorophore is reconstructed. This paper aims at finding the optimal spatial distribution of optodes in order to keep their number (hardware costs) low while gaining maximum information from the target object.

The implemented algorithm starts with an arbitrary pool of feasible optodes. The optimal subset is searched by minimizing the mutual information between the different measurements. This goal is reached by subsequently removing those sources and detectors which add the least independent information until a stopping criterion is reached.

Mutual information is estimated by calculating the inner products between the rows of the sensitivity matrix i.e. the first derivative of the forward mapping with respect to the optical parameters to be reconstructed. We assembled this matrix with a finite element implementation of the diffusion approximation of light propagation in scattering tissues. When starting with an initial pool of 96 optodes regularly spaced on a cylindrical surface and focusing on different target regions within the cylinder, the algorithm always converged towards physically reasonable optimal sets. Optimal source/detector patterns are be presented graphically and numerically.

Keywords: fluorescence diffusion optical tomography, optimization, hardware config- uration, optimal design, finite element method

1   Introduction

Fluorescence diffusion optical tomography (FDOT) is one of the newer imaging techniques with promising ap- plication potential in medicine. FDOT provides the possibility of functional imaging, i.e it not only visualizes anatomical structures but also provides information about physiological states and processes. FDOT utilizes the ability of so-called fluorophores to absorb light in a certain wavelength range

[*]Institute of Medical Engineering, Graz University of Technology, Stremayrgasse 16, 8010 Graz, Austria

and to emit photons at a higher wavelength. The excitation light is injected into the sample through a set of sources. A source can either be in contact with the sample's surface (e.g. a waveguide) or it delivers the light in a contactless manner using collimated or divergent light-beams. The excitation light is scattered and absorbed while spreading in the tissue. At sites where a fluorophore is present and active (e.g. inside a tumor), a part of the absorbed light leads to re-emission at another wavelength. This secondary light is again scattered through the tissue and the part which reaches the boundary can be measured by photo detectors. Due to the diffuse propagation of the photons in tissue [1] light emerging from the fluorescent dye widely spreads before it reaches the boundary. This is in contrast to other established imaging techniques like X-rays where the rays travel through the sample of interest in nearly straight lines. The photon diffusion has to be considered in a suitable forward model which is the basis for the reconstruction algorithm that seeks to determine the distribution of the fluorophore from boundary measurements. The reconstructed results usually improve when increasing the number of sensors. However, this is only true up to a certain extent as the diffuse nature of the photon propagation inherently limits the independence of information of different sensors and hence the obtainable resolution. Contrary the computation time and the memory needed for reconstruction increases with the amount of data gathered. The goal is to find a good compromise between image quality and computational effort. Graves et al [2] have performed some investigations on how the amount of sources and detectors and their distance, respectively, influences the reconstruction. Later on the method was extended by Lasser et al [3] who applied it to $360 \deg$ projection tomography. This paper presents a different approach to optimize the optode configuration of a fluorescence tomography system in the sense that it does not just compare different optode configurations but provides an information measure for every single optode offering greater flexibility, therefore. Furthermore, it will be shown how the optimization can be modified such that the reconstruction is focused to a given region of interest.

2 Methods

2.1 Forward model

One of the most accurate ways to model light propagation is to utilize Boltzmann's transport equation for kinetic gases. The photons can then be treated like independent gas particles leading to the radiation transfer equation. Unfortunately, the photon intensity in the radiation transfer equation is a field dependent on the spatial coordinates and the direction (i.e. two angles) into which the photons travel. This leads to a discretization with a huge amount of degrees of freedom and requires extensive computing power and memory.

　　Therefore, it is common to use an approximation of the transfer equation known

as diffusion equation [4]. Including a spatially variable flurophore concentration c the diffusion equation reads

$$- \nabla(\kappa(x)\nabla\phi) + \left(\mu_{a,i}(x) + c(x)\epsilon + \frac{i\omega}{\nu}\right)\phi(x) = q(x) \tag{1}$$

together with the boundary condition

$$\phi(x) + 2R\kappa(x)\frac{\partial\phi(x)}{\partial n} = 0 \tag{2}$$

where ϕ is the photon density field, $\kappa = (3(\mu_{a,i}(x) + c(x)\epsilon + \mu'_s))^{-1}$ is the diffusion coefficient, μ_a and μ'_s are the intrinsic absorption and reduced scattering coefficient, respectively, ϵ is the fluorophore's extinction which relates the flurophore concentration to photon absorption. ω is the modulation frequency of the light source, ν is the speed of light and q is the source term. R is a factor to incorporate reflections at the boundary whose normal is denoted by n. Equation 1 can be solved efficiently using a finite element discretization.

For fluorescence applications, two diffusion equations — one describing the propagation of the excitation photons, another one to describe the emission field — can be coupled. We prefer to write this in an operator (or matrix-like) notation, where A_{ex} and A_{em} describe the propagation of the excitation and emission field, respectively.

$$\begin{aligned} A_{ex}(x)\phi_{ex} &= q \\ A_{em}(c)\phi_{em} &= B(c)\phi_{ex} \end{aligned} \tag{3}$$

The operator B is an operator converting the photon density from the excitation wavelength to the emission wavelength at sites where a fluorophore is present and serves as a source for the emission field, thus. It definition reads

$$B(c)\phi_{ex}(x) = Qc(x)\epsilon ex\phi_{ex}(x), \tag{4}$$

whereby Q is the quantum yield [5].

Although more elaborate detector models (e.g. [6]) could be used, in this paper a measurement d is defined as the number of photons leaving the sample at a certain point x_D per unit time:

$$d := -\upsilon\int_{\partial\Omega}\delta(x - x_D)\kappa_{em}\frac{\partial\phi_{em}(x)}{\partial n}dx \stackrel{=}{_2} \nu\int_{\partial\Omega}\delta(x - x_D)\phi_{em}(x)dx. \tag{5}$$

If more than one source and one detector are present, we denote the measurement made with the i-th pair of source/detector by d_i.

2.2 Sensitivity

In order to solve the inverse problem, i.e. the reconstruction of the distribution of the fluorophore's concentration c from measurements on the boundary, it is necessary to know the influence of a change in the concentration distribution on the measurements. In other words the so-called *sensitivity*, the derivative of the system 3 with respect to c, is needed. Assuming the existance of the inverses of the operators A_{ex}, A_{em} and B as well as their Fréchet derivatives A'_{ex}, A'_{em} and B' one can derive the expression

$$dd_i = \int_{\partial\Omega} \frac{1}{2R}\delta(x - x_D)A_{em}^{-1}\left[B'(\delta c)\phi_{ex} - BA_{ex}^{-1}A'_{ex}(\delta c)\phi_{ex} - A'_{em}(\delta c)\phi_{em}\right]dx \quad (6)$$

which relates the change of the concentraton δc to a change in the measurement dd_i.

In a finite element context, the discretization of the concentration using piecewise-constant ansatz functions in equation 6 leads to the Jacobian or sensitivity matrix which is denoted by J in this paper. The element J_{ij} describes the effect of a concentration change in the j-th finite element on the i-th measurement.

In certain applications it is feasible to operate with difference measurements. A measurement d_0 is made with a base-line concentration c_0 and a second measurement d_1 is performed after the concentration distribution has changed to c_1. If the difference in concentrations is small, the following linearization can be used for reconstruction:

$$\Delta d = J(c_0)\Delta c, \quad (7)$$

where Δd is a vector of difference measurements and Δc the vector of concentrations in the finite elements. This formulation will be used throughout this paper. However, the Δ is neglected from now on and we understand all measurement and concentrations as differences to a base-state.

3 Optode Optimization

3.1 Redundancy reduction

The optimization approach implemented is based on the idea that the different measurements should be as independent as possible, i.e. every measurement should result in new information which can be used for the inverse problem. The method was originally developed by Michelini and Lomax and published by Curtis et al [7] One way to quantify the independence of two measurements is to calculate the inner product and the angle, respectively, between the respective rows of the sensitivity matrix J. Then the algorithm has to find that set of measurements which is closest to an orthogonal set.

Table 1: Values of optical parameters used for the forward simulation [5, 9, 10]

	$\mu_s'/\,\mathrm{mm}^{-1}$	$\mu_{a,i}/\,\mathrm{mm}^{-1}$	$c\epsilon/\,\mathrm{mm}^{-1}$	R
excitation	0.275	0.036	$0.84 10^{-3}$	2.51
emission	0.235	0.029	$0.25 10^{-3}$	2.51

Let $\mathcal{M}$ be the set of all measurement indices, i.e. each element of $\mathcal{M}$ uniquely defines one pair of source/detector. Further, let $S_i \subset \mathcal{M}$ denote the indices of those measurements which are made with the i-th source. The cosine of the angle between two measurements, one made with source i and one made with another source, is given by the term

$$\frac{|(J_m, J_n)|}{\|J_m\|\|J_n\|}, \quad m \in S_i, \quad n \in \mathcal{M} \setminus S_i \tag{8}$$

which is a real number from the interval $[0, 1]$. If the measurements are orthogonal, the cosine will be zero and the more they depend on each other, the more the cosine will approach 1.

Now, consider the average cosine between all measurements made with source i and those measurements made with another source. This will lead to the expression

$$r_i := \frac{1}{|S_i||\mathcal{M} \setminus S_i|} \sum_{m \in S_i} \sum_{n \in \mathcal{M} \setminus S_i} \frac{|(J_m, J_n)|}{\|J_m\|\|J_n\|} \tag{9}$$

which is again from $[0, 1]$. If the value is close to 1, the information gained using source i can also be gained using other sources, which means that the source i is redundant. Therefore, we call ri the redundancy of source i.

The quantity $q_i := 1 - r_i$ is used as a quality criterion in the optimization algorithm. An analog expression can be derived for the detectors. This is achieved by replacing Si in [9] by Di , the set of indices belonging to the i-th detector.

The optimization algorithm starts with a set of feasible optodes. It then calculates the quality measure for every source and removes the one with the lowest measure from the optode pool. In the next iteration the same is performed for the detector optodes. This is done until a previously defined amount of sources and detectors is left in the pool, which is considered to be the optimal measurement configuration.

3.2 Focusing

Sometimes it is necessary to focus the optode arrangement so as to reach e.g. higher sensitivity or a higher resolution in a specified region than in the rest of the sample.

A rather simple approach is to multiply the sensitivity matrix J with a weighting mask f from the right:

$$J_F = J \,\mathrm{diag}(f), \tag{10}$$

whereby J_F denotes the focused sensitivity matrix. This resultant matrix will then be used in the optimization algorithm described above. In the simplest case f is a binary vector being one in the region of interest and zero everywhere else. Also smooth variations of the mask are possible. Generally we can assume that $0 \leq f_i \leq 1$.

4   Results

The optode optimization was performed on a cylinder with a height of $90\,\mathrm{mm}$ and a radius of $30\,\mathrm{mm}$, which should mimic a small animal. The values of the optical properties are found in table 1. The fluorophore concentration was assumed to be spatially constant and rather low. This should model a cylinder without fluorescent dye but with a certain amount of auto-fluorescence. Equations and estimates of these parameters can be found in references [5,9,10]. The absorption due to auto-fluorescence was chosen to be 10^{-2} of the fluorophore absorption found in the thesis by Joshi [5].

As initial pool of feasible optodes position a regular grid with 48 source and 48 detector nodes was specified. The optodes were arranged in a zig-zag like pattern on six rings with a spacing of 10 mm (see figure 4). The optimization algorithm needs the desired amount of sources and detectors as stopping-criterion. It was chosen to search for an optimal configuration consisting of eight sources and eight detectors.

Three different focus regions were chosen to demonstrate the optimization. Region A consists of the voxels in the cylinder slice given by $10\,\mathrm{mm} < z < 20\,\mathrm{mm}$, region B is another slice defined by $-7.5\,\mathrm{mm} < z < 7.5\,\mathrm{mm}$ and region C is the half-cylinder slice $-20\,\mathrm{mm} < z < -10\,\mathrm{mm}$ and $x > 0\,\mathrm{mm}$. The optimization was performed on a finite element mesh with approximately 30,000 tetrahedral elements.

As can be seen in the plots 2(a)–2(c), the optimization algorithm tends to concentrate the optodes near the focused region. The source optodes are very close to the region of interest and so are the detectors in 2(c). However, the detector optodes when focusing on larger slices as in 2(a) and 2(b) are more spread.

Figures 3(a)-3(c) showes the rank of the optodes, i.e. the iteration in which they were removed from the pool of feasible optodes. One can see clearly that the optodes far away from the focus region, which is drawn in orange, are removed first while those near or in the region remain longer in the feasible set.

As a quality measure for the the resultant sensitivity matrix, its singular values and its condition number, respectively, can be used. The largest and smalles singular values together with the ratio between them can be found in 2. The full optode configuration has a rather high ratio of $6\ 10^{12}$ and is rather ill-conditioned, thus. The focused designs show a SV ratio which is reduced by a factor of at least 10^6.
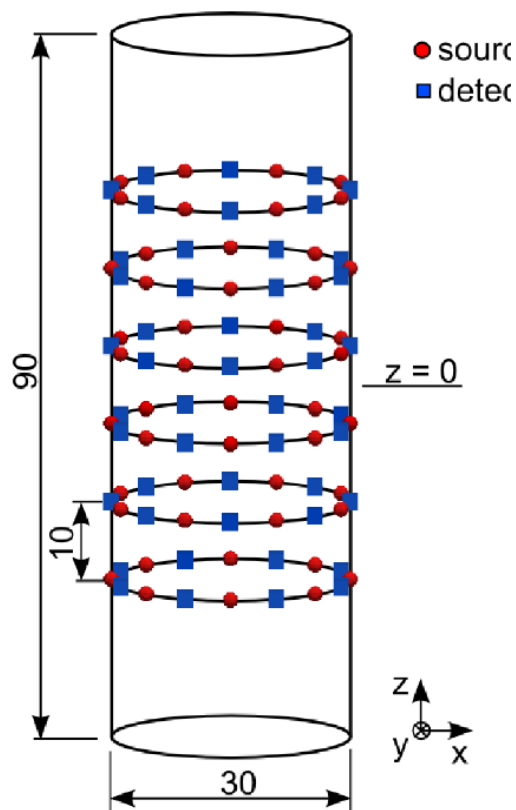
Figure 1: Geometry of the simulation model used for optode optimization together with the initial pool of feasible optodes which are arranged in a zig-zag pattern on six rings with 10mm spacing. All measures are in mm.
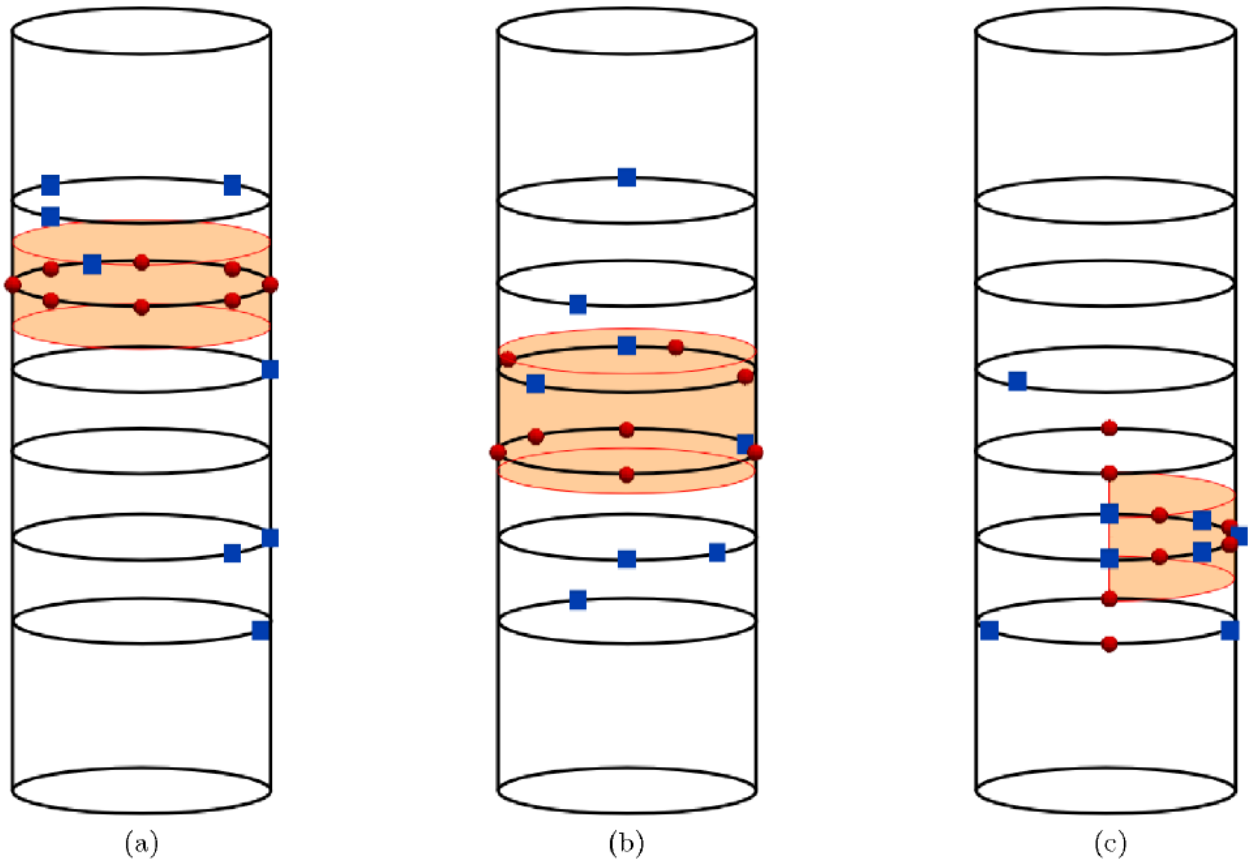
Figure 2: The result of the optode optimization for three different focus regions which are drawn in orange.
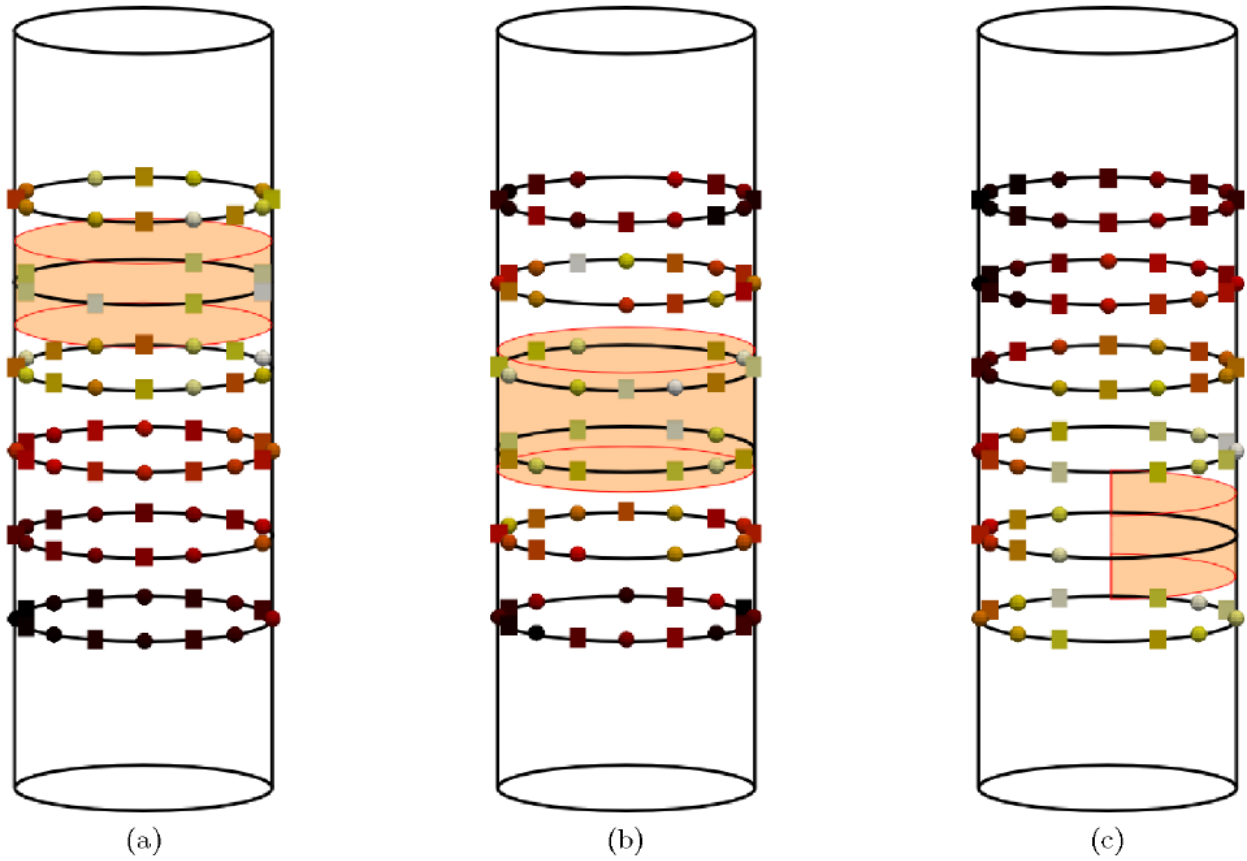
Figure 3: The rank of the removed optodes encoded in colour for three different focus regions (sketched in orange). Optodes drawn in black or red were eliminated early during the optimization while those in yellow and white remained longer in the pool. The optimal optodes have been omitted.

Table 2: Singular value (SV) analysis for the full sensitivity matrix and the optimized configurations with focusing to different regions. The table lists the largest and smalles singular value as well as the ratio between them which is a measure of stability for matrix inversion.

Design	max SV	min SV	ratio
Full pool of optodes	60.21e-3	9.972e-15	6.066e12
Focus on region A	20.23e-3	1.929e-8	1.049e6
Focus on region B	19.37e-3	1.143e-6	16.94e3
Focus on region C	27.72e-3	2.500e-6	11.09e3

This is exactly what is intended by the redundancy minimization algorithm as the removal of non-orthogonal rows from the sensitivity matrix improves conditioning. The ratio of the design focusing on region A (the uppermost cylinder slice) is higher by a factor of approximately 100 than the other two designs. This is probably due to the rather large spread of the optimal detectors (figure 2(a)).

5   Discussion

The location of the sensors and detectors is a critical design parameter for FDOT hardware. A configuration determines the sensitivity of the measurements in a given region and also the obtainable resolution. The method presented herein is based on a simulation model and can be used prior to building hardware, thus. Comparisons of different hardware configurations for fluorescence tomography that we are aware of were previously reported by Graves et al [2] and Lasser et al [3]. Their approaches were based on a singular value (SV) analysis of the so-called weight matrix. This matrix is essentially the sensitivity matrix used in this paper but was obtained using the normalized Born approximation [11].

In contrast to the previous methods, the optimization algorithm used herein operates on the complete deriva- tive of the diffusion approximation given by the system 3. Therefore, also the change of the excitation field due to a perturbation in the fluorophore distribution is considered. This is neglected, if a first-order Born approximation is used.

The SV analysis implemented by Graves and Lasser requires a singular value decomposition (SVD) of the sensitivity matrix. This demands a large amount of computation time. As an example, the calculation of a single SVD for a matrix of size $48 \times 48 \times 30000$, takes about 50 minutes on an Intel Core2Duo with two processors. Therefore, the SV optimization is limited to 2D applications or to rather simple 3D geometries, which can be modelled with fewer finite elements. The redundancy

reduction algorithm does not suffer from these limitations as it only needs the sensitivity matrix on which it performs basically inner products. The assemby of the sensitivity matrix as well as the inner products can be parallelized with moderate effort (the basic principle can be found in reference 12, for example) which allows a further speed-up.

Another advantage of redundancy minimization is provides a quality measure for any single optode rather than comparing complete configurations. This offers the possibility to choose a superior configuration first, which could even be obtained with a completely different method, and to optimize the arrangement further through removal of optodes exhibiting a poor quality measure.

The optimization algorithm presented works top-down i.e. it starts with a full optode arrangement and iteratively discards the optodes having a low quality measure. Such a strategy is easy to implement but there is no guarantee that it reaches the global minimum. This is due to the fact that the algorithm cannot determine the final result if in a certain iteration the optode with the worst quality was kept and another one was thrown out instead. To ensure finding the global minimum, all possible optode configurations needed to be tried which is in the order of 10^{18} for the initial configuration presented in here and computationally not feasible, thus.

The large spread of the detector optodes as it is visible especially in figure 2(a) may appear counter-intuitive at first sight. However, still unpublished observations show that the intrinsic absorption $\mu_{a,i}$ influences the optimalconfiguration. If this parameter is increased, the optimal optodes cluster more around the focused region.

A drawback of this algorithm is that the desired number of optimal sources and detectors has to be specified beforehand. However, other stopping criteria as the minimum resolution or the minimum contrast-to-noise ratio could be preferable. A more dynamic approach would be to monitor a certain quality measure (resolution or CNR) during the removal of optodes from the feasible set. Then the algorithm could stop when the quality measure shows a steep decrease, i.e. it would remove only those sources and detectors which do not contribute too much to the quality measure chosen. Further investigations have to be done to include those quality measures into our method.

6   Conclusion

A quite flexible algorithm for optimizing the illumination and detection pattern in fluorescence diffusion optical tomography has been presented. The algorithm is based on the maximization of independence between the individual measurements. It has been tested with synthetic data and provided reasonable results.

7   Acknowledgements

References

[1]   Mobley, J. and Vo-Dinh, T.   Optical properties of tissue   In *Biomedical Photonics Handbook], Vo-Dinh, T., ed., 2/1-2/75*, CRC Press LLC, Boca Raton, USA (2003).

[2]   Graves, E. E., Culver, J. P., Ripoll, J., Weissleder, R., and Ntziachristos, V. Singular-value analysis and optimization of experimental parameters in fluorescence molecular tomography J. Opt. Soc. Am. A 21, 231–241 (2004).

[3]   Lasser, T. and Ntziachristos, V. Optimization of 360 projection fluorescence molecular tomography Medical Image Analysis 11, 389–399 (2007).

[4]   Arridge, S. R. Optical tomography in medical imaging Inverse Problems 15, R41–R93 (1999).

[5]   Joshi, A. Adaptive finite element methods for fluorescence enhanced optical tomography PhD thesis, Texas A&M University (2005).

[6]   Schulz, R. B., Peter, J., Semmler, W., D'Andrea, C., Valentini, G., and Cubeddu, R. Comparison of noncontact and fiber-based fluorescence-mediated tomography Optics Letters 31, 769–771 (2006).

        Curtis, A., Michelini, A., Leslie, D., and Lomax, A. A deterministic algorithm for experimental design applied to tomographic and microseismic monitoring surveys Geophysical Journal International 157, 595–606 (2004).

[8]   Ahmed, N. A. and Gokhale, D. V. Entropy expressions and their estimators for multivariate distributions IEEE Transactions on Information Theory 35, 688–692 (May 1989).

[9]   Alexandrakis, G., Rannou, F. R., and Chatziioannou, A. F. Tomographic bioluminescence imaging by use of a combined optical-pet (opet) system: a computer simulation feasibility study Physics in Medicine and Biology 50, 4225–4241 (2005).

[10]   Keijzer, M., Star, W. M., and Storchi, P. R. M. Optical diffusion in layered media Applied Optics 27, 1820–1824 (1988).

[11]   Ntziachristos, V. and Weissleder, R. Experimental three-dimensional flu-
         orescence reconstruction of diffuse media by use of a normalized born
         approximation Optics Letters 26, 893–895 (2001).

[12]   Haase, G., Kuhn, M., and Langer, U. Parallel multigrid 3d maxwell solvers
         Parallel Computing 27, 761–775 (2001).