HAL
open science

Nonnegative matrix factorization: a blind spectra separation method for in vivo fluorescent optical imaging

Anne-Sophie Montcuquet, Lionel Herve, Fabrice Navarro, Jean-Marc Dinten, Jerome I. Mars

To cite this version:

Anne-Sophie Montcuquet, Lionel Herve, Fabrice Navarro, Jean-Marc Dinten, Jerome I. Mars. Nonnegative matrix factorization: a blind spectra separation method for in vivo fluorescent optical imaging. Journal of Biomedical Optics, 2010, 15 (5), pp.0560091-14. 10.1117/1.3491796 . hal-00583030

HAL Id: hal-00583030

https://hal.science/hal-00583030

Submitted on 4 Apr 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Nonnegative matrix factorization: a blind spectra separation method for *in vivo* fluorescent optical imaging

Anne-Sophie Montcuquet
CEA-LETI, Minatec
17 rue des Martyrs
Grenoble Cedex 9, 38054
France
and
GIPSA-Lab/DIS, CNRS
UMR 5216
BP Saint Martin d'Hères Cedex
France

Lionel Hervé
Fabrice Navarro
Jean-Marc Dinten
CEA-LETI, Minatec
17 rue des Martyrs
Grenoble Cedex 9, 38054
France

Jérôme I. Mars
GIPSA-Lab/DIS, CNRS
UMR 5216
BP Saint Martin d'Hères Cedex
France

Abstract. Fluorescence imaging in diffusive media is an emerging imaging modality for medical applications that uses injected fluorescent markers that bind to specific targets, e.g., carcinoma. The region of interest is illuminated with near-IR light and the emitted back fluorescence is analyzed to localize the fluorescence sources. To investigate a thick medium, as the fluorescence signal decreases with the light travel distance, any disturbing signal, such as biological tissues intrinsic fluorescence (called autofluorescence) is a limiting factor. Several specific markers may also be simultaneously injected to bind to different molecules, and one may want to isolate each specific fluorescent signal from the others. To remove the unwanted fluorescence contributions or separate different specific markers, a spectroscopic approach is explored. The nonnegative matrix factorization (NMF) is the blind positive source separation method we chose. We run an original regularized NMF algorithm we developed on experimental data, and successfully obtain separated *in vivo* fluorescence spectra. © *2010 Society of Photo-Optical Instrumentation Engineers.*
[DOI: 10.1117/1.3491796]

Keywords: fluorescence optical imaging; spectroscopy; source separation; autofluorescence.

Paper 10039R received Jan. 27, 2010; revised manuscript received Jul. 21, 2010; accepted for publication Aug. 16, 2010; published online Sep. 30, 2010.

1 Introduction

Medical diagnostic systems based on fluorescent imaging are envisioned to be noninvasive, easy to use, and cost effective. Fluorescent markers are injected into a patient, and bind specifically to targeted compounds, such as tumors.[1] Several specific markers can be injected at once, and bind to different compounds or organs; that method is used to survey different biological processes or organs, such the evolution of carcinoma, or, for example, to measure blood flow. The region of interest is illuminated with near-infrared (NIR) light; an optimal wavelength range can be defined between 600 and 900 nm, where tissue absorption is lower. The excitation wavelength is thus selected to ease the tissue penetration, and to optimally excite the injected markers. Finally, the emitted back fluorescence signal is measured and the fluorescent source is localized. So far, NIR fluorescence imaging is mainly used on small animals where some markers are available for injection, and where the layer of biological tissues to explore does not exceed a few centimeters. In such a case, a biological tissues intrinsic fluorescence, called autofluorescence,[2] exists, but is insignificant compared to the fluorescent marker signal. To investigate thick media for medical diagnostic applications ($\cong 4$ cm), the autofluorescence must be considered. As the fluorescence signal becomes exponentially weak with the distance light travels, an autofluorescence signal remains constant and becomes a limiting factor. The analysis of a fluorescence signal impaired by autofluorescence may lead to a wrong localization of the markers; the signal must be preprocessed to remove autofluorescence.

Many algorithms to unmix fluorescence spectra and to filter autofluorescence have already been developed and tested on small animal examination equipments: methods such as nonlinear least squares,[3] spectra subtraction,[4] principal component analysis (PCA), independent component analysis, (ICA) and singular value decomposition (SVD) were proposed.[5–8] In particular, the ICA method requires sources to be statistically independent, which is not appropriate hypothesis for our fluorescence unmixing problem, and SVD considers orthogonal sources and allows negative values. Much *a priori* knowledge about the nature of the sources, such as sparsity or independence, is taken into account in these methods. But a principal that must be considered is nonnegativity. Many real-world data are nonnegative, as fluorescence data are, and the resulting unmixed fluorescence spectra have a physical meaning only when nonnegative.

In 1987, Henry[9] raised the issue of the nonexistent nonnegativity constraint in the factor analysis algorithms (SVD, PCA, etc.). In light of this observation, many original algorithms were developed. One of them was the positive matrix factorization (PMF), developed by Paatero and Tapper[10] in

Address all correspondence to Anne-Sophie Montcuquet, Tel: +33 438 782 758; E-mail: anne-sophie.montcuquet@cea.fr

1083-3668/2010/15(5)/056009/14/$25.00 © 2010 SPIE

Downloaded from SPIE Digital Library on 04 Apr 2011 to 193.48.255.141. Terms of Use: http://spiedl.org/terms

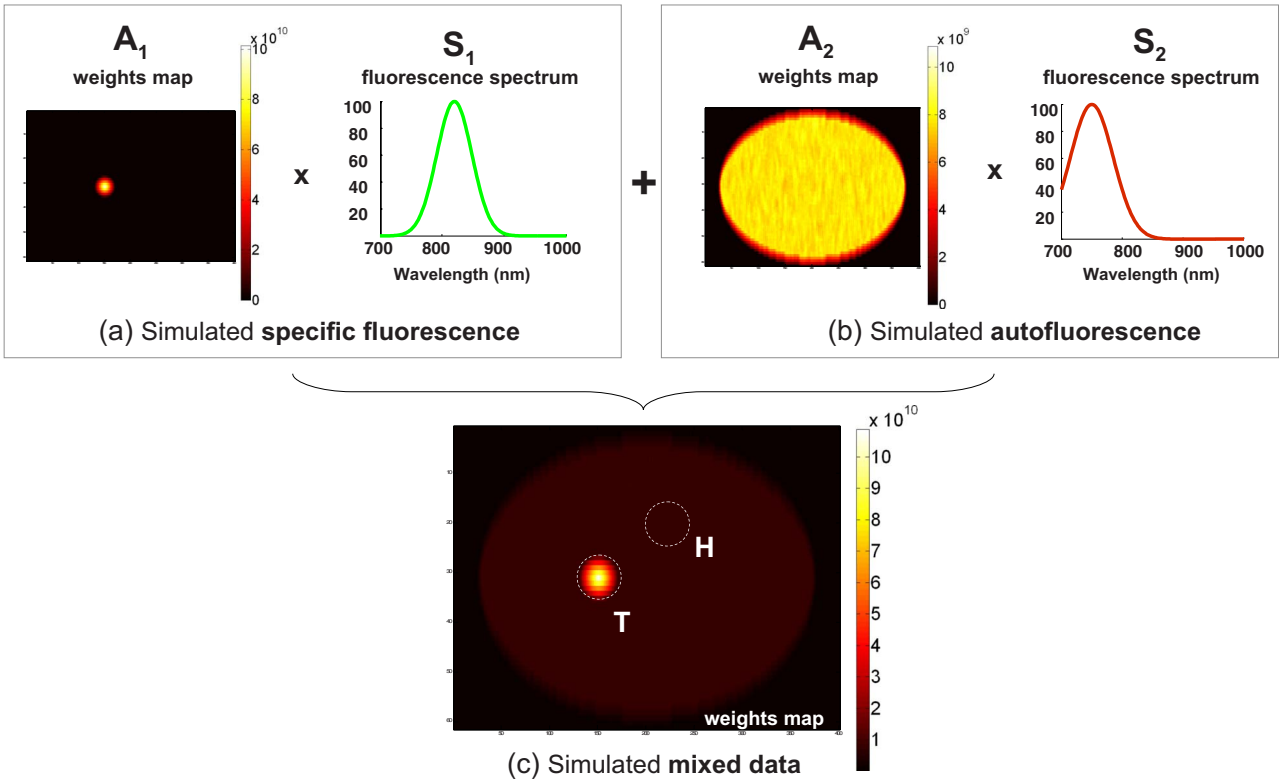

Fig. 1 Sum of (a) a specific fluorescence signal and (b) an autofluorescence signal leads to (c) simulated mixed data.

1994, which uses alternative least squares (ALS) to minimize a chosen cost function. Expectation maximization[11] (EM) also minimizes a cost function by the use of an auxiliary function. Finally, from those methods the nonnegative matrix factorization (NMF) was forged, notably investigated by Paaetero and Tapper, which gained popularity in 2001 through the works of Lee and Seung.[12]

NMF is a useful matrix decomposition for multivariate data, that differs from the methods already cited (SVD, PCA) in that it forces the matrix factors to be nonnegative. Thanks to this distinctive feature, NMF plays a major role in a wide range of applications, such as in chemometrics (commonly known as self–modeling curve resolution[13]), bioinformatics,[14] neuroscience,[15] text mining,[16] or spectral analysis.[17] In 2006, Gobinet et al.[18] used the NMF decomposition on spectroscopic data to unmix several pure fluorescence spectra, and later, as a preprocessing step for Raman spectra analysis of biomedical samples.[19] Since 2008, Xu et al.[17] used and Xu and Rice[20] have a multivariate curve resolution (MCR) method to separate different fluorescence spectra, including the autofluorescence, based on multispectral imaging data. The MCR model is also based on matrix factorization, and constrained to nonnegativity; Xu et al. couple this method to the ALS optimization approach. Their method is probably the closest to our work, but sill differs from experimental setup used (selection of few emission wavelengths, and often several excitation wavelengths while, we use unique excitation wavelength and use the total emission spectrum from 700 to 1000 nm, thanks to an imaging spectrometer) to optimization method suggested (ALS versus multiplicative update rules). As blind source separation problems are often depicted and solved under a matrix factorization form, the result strongly depends on the *a priori* knowledge we introduce. Work on matrix factorization and spectroscopy run by Gobinet et al.[18] inspired our work, but fluorescence spectroscopy introduces fluorescent markers and thus *a priori* knowledge that was not taken into account in Gobinet's work. Xu et al.[17] work on *in vivo* small-animal imaging is close to our problem, but still different from deep imaging problem where autofluorescence may have same intensity level than specific fluorescence signal. Finally, as already underlined, both teams use different resolution methods.

For *in vivo* fluorescence spectroscopy, the unmixing problem is referred to as a blind source separation problem since fluorescence spectra may vary according to the fluorescent dye biological environment. Fluorescence spectra to separate are also supposed statistically dependent, which filters out many methods (such as ICA). Finally, the NMF algorithm seems to be, in many ways, more suitable in blind positive spectra separation than other separation methods. We propose to test this method on spectroscopic data and to define a new regularized NMF algorithm that may better suit fluorescence spectroscopy data in particular cases.

First, we introduce the spectra unmixing problem: from a mixed fluorescence signal composed of a known number of sources, we want to obtain separated contributions for each fluorescence source; the NMF decomposition is proposed to unmix fluorescence spectra, and the method is explained. Several NMF algorithms exist, based on diverse criteria to minimize and optimization methods: we present in the second part a classical NMF algorithm that deals with multiplicative update rules.[12] As with all blind source separation methods, it is

Downloaded from SPIE Digital Library on 04 Apr 2011 to 193.48.255.141. Terms of Use: http://spiedl.org/terms

impossible to find a unique NMF decomposition. In a second part we study the nonuniqueness issue and suggest regularizations and *a priori* knowledge considerations to refine the solution set. Two axes are examined: the influence of initialization of matrices and regularization on initialization. Both issues are studied on a breast simulation case; a regularized NMF algorithm that selects appropriate initialization or takes into account some *a priori* knowledge on the fluorescence spectra we want to unmix is presented; associate original update rules are also detailed. Finally, we ran our NMF algorithm on experimental *in vivo* data with success. Two fluorescent markers—ICG loaded into nanoparticules (ICG-LNP) and Alexa 750—were placed on mice to simulate marked targets. Adding the autofluorescence signal, the NMF algorithm achieved the separation of three overlapping fluorescent sources. A second example deals with detection of a multi-depth target: the same marker (ICG-LNP) at different depths is placed on a mouse, and deeper markers are mixed up with autofluorescence signal. The results of both experiments are presented in the last section.

2 NMF

2.1 *NMF and Spectroscopy*

Let us consider a fluorescence spectrum $\mathbf{m}$ measured in a medium that for example contains two kinds of fluorescence sources: fluorescence spectra $\mathbf{s}_1$ and $\mathbf{s}_2$ may overlap, but have distinct emission peaks. Thus, the measured fluorescence $\mathbf{m}$ is a mixture of both sources of the medium; if we call a_1 and a_2 the amount of respectively spectra $\mathbf{s}_1$ and $\mathbf{s}_2$ in $\mathbf{m}$, all those quantities being nonnegative, we can write

$$\underset{(1\times N_\lambda)}{\mathbf{m}} = a_1\mathbf{s}_1 + a_2\mathbf{s}_2 = \underset{(1\times 2)}{\begin{pmatrix} a_1 & a_2 \end{pmatrix}} \underset{(2\times N_\lambda)}{\begin{pmatrix} s_{1,1} & \dots & s_{1,N_\lambda} \\ s_{2,1} & \dots & s_{2,N_\lambda} \end{pmatrix}} = \mathbf{aS}. \tag{1}$$

It can be easily generalized to a mixing model with P distinct sources present in the medium. The fluorescence spectrum $\mathbf{m}$ is now a linear combination of the P sources spectra $s_i, i \in (1,P)$, in quantities $a_i, i \in (1,P)$. Vector $\mathbf{a} = (a_1, \dots, a_P)$ is considered as the weight vector, and S as the spectra matrix, both containing as much elements P as fluorescent sources to separate. The following equation depicts the decomposition obtained for P sources:

$$\underset{(1\times N_\lambda)}{m} = a_1\mathbf{s}_1 + a_2\mathbf{s}_2 + \dots + a_P\mathbf{s}_P = \underset{(1\times P)}{\overbrace{\begin{pmatrix} a_1 & a_2 & \dots & a_P \end{pmatrix}}^{\text{weight vector}}} \underset{(P\times N_\lambda)}{\overbrace{\begin{pmatrix} s_{1,1} & \dots & s_{1,N_\lambda} \\ & \vdots & \\ s_{P,1} & \dots & s_{P,N_\lambda} \end{pmatrix}}^{\text{spectra matrix}}} \tag{2}$$

We now finally enlarge the example to a set M of N_s spectra acquired on a N_λ wavelength range, for P sources present in the medium. We have thus:

$$\underset{(N_s\times N_\lambda)}{\mathbf{M}} = \left.\begin{matrix} \underset{(N_s\times P)}{\overbrace{\begin{pmatrix} a_{1,1} & \dots & a_{1,P} \\ \vdots & & \vdots \\ a_{N_s,1} & \dots & a_{N_s,P} \end{pmatrix}}^{\text{weight matrix } \mathbf{A}}} & \underset{(P\times N_\lambda)}{\overbrace{\begin{pmatrix} s_{1,1} & \dots & s_{1,N_\lambda} \\ \vdots & & \vdots \\ s_{P,1} & \dots & s_{P,N_\lambda} \end{pmatrix}}^{\text{spectra matrix } \mathbf{S}}} \end{matrix}\right\} P \text{ lines} = \mathbf{AS} \tag{3}$$

This reasoning led to a matrix factorization, that only deals with nonnegative components, which brings us close to the NMF definition. NMF proposes to find a couple of matrices $(\mathbf{A},\mathbf{S})$ with nonnegatives coefficients, whose product approaches the best possible the initial non-negative data $\mathbf{M}$. Indeed, the classical NMF definition says[12]

Given a nonnegative matrix $\mathbf{M} \in \mathbb{R}^{N_s\times N_\lambda}$, find nonnegative matrix $\mathbf{A} \in \mathbb{R}^{N_s\times P}$ and $\mathbf{S} \in \mathbb{R}^{P\times N_\lambda}$ such that

$$\mathbf{M} \simeq \mathbf{AS}, \tag{4}$$

where nonnegative matrices are matrices whose all factors are nonnegative.

To find the particular matrices $\mathbf{A}$ and $\mathbf{S}$ that satisfy Eq. (4), two distinct steps must be considered. First, a distance between $\mathbf{M}$ and model $\mathbf{AS}$ is defined. Then, in a second step, optimizing this distance under the nonnegativity constraint

Downloaded from SPIE Digital Library on 04 Apr 2011 to 193.48.255.141. Terms of Use: http://spiedl.org/terms

would lead to the a factorization solution. Several "distances" (the Eunclidean distance, the Kullback-Leibler divergence,[10,12] etc.) and different optimization methods (ALS, update rules) may be chosen to obtain the NMF decomposition.

2.2 *NMF Algorithm*

Many NMF algorithms have been proposed since Paatero and Tapper,[10] but in 2001, NMF popularity increased after Lee and Seung published two original NMF algorithms,[12] based on the use of multiplicative update rules to minimize square of the euclidean distance and the Kullback-Leibler divergance between and $\mathbf{M}$ and $\mathbf{AS}$.

2.2.1 *Cost function definition*

Here, we chose our cost function F as the square of the Euclidean distance between $\mathbf{M}$ and $\mathbf{AS}$ (Refs. 12 and 21), and defined as

$$F=\sum_{n=1}^{N_s}\sum_{\lambda=1}^{N_\lambda}\left(m_{s\lambda}-\sum_{p=1}^{P}a_{sp}s_{p\lambda}\right)^2=\|\mathbf{M}-\mathbf{AS}\|^2. \tag{5}$$

The cost function F is lower bounded by 0.

2.2.2 *Optimization*

The following optimization problem is thus considered:

Problem 1. Find couple $(\mathbf{A},\mathbf{S})$ such as $(\mathbf{A},\mathbf{S})=\mathrm{argmin}_{(\mathbf{A},\mathbf{S})\geqslant 0}\|\mathbf{M}-\mathbf{AS}\|_2^2$.

Many methods can be implemented to solve this problem. The gradient descent is probably the simplest, and most famous one, but it is also known for a slow convergence. Other faster methods, such as the conjugate gradient, are usually more complicated to implement.[22] ALS is also commonly used in such problems. In 2001, Lee and Seung[12] proposed multiplicative update rules to minimize F: it offers a good compromise between speed and ease of implementation to solve Problem.

Theorem 1. The distance $\|\mathbf{M}-\mathbf{AS}\|^2$ is nonincreasing under the update rules:

$$\mathbf{S}_{p\lambda}\leftarrow\mathbf{S}_{p\lambda}\frac{(\mathbf{A}^t\mathbf{M})_{p\lambda}}{(\mathbf{A}^t\mathbf{AS})_{p\lambda}}\quad \mathbf{A}_{np}\leftarrow\mathbf{A}_{np}\frac{(\mathbf{MS}^t)_{np}}{(\mathbf{ASS}^t)_{np}}, \tag{6}$$

where $\mathbf{X}^t$ is the transpose of a matrix $\mathbf{X}$. The proof of this theorem is given in Lee and Seung's publication.[12]

We became interested in these update rules precisely because of their ease of implementation and speed, for which they were initially created. Finally, they are easily convertible if some regularization is required. We thus defined original regularized update rules adapted to our fluorescence data which take into account *a priori* knowledge on the fluorescence spectra.

2.3 *Nonuniqueness of the NMF Decomposition*

The chosen cost function F is not jointly convex in matrices $\mathbf{A}$ and $\mathbf{S}$: there are numerous local minima to the function and nonuniqueness of the NMF factorization. Let us assume a factorization of $\mathbf{M}$ by the product of some matrices $\mathbf{A}$ and $\mathbf{S}$ exists. If we now consider any invertible matrix $\mathbf{T}$ of size $P\times P$, then a new couple $(\mathbf{A},\mathbf{S})$ of solutions is easily found:

$$\mathbf{M}=(\underbrace{\mathbf{AT}^{-1}}_{\tilde{\mathbf{A}}})(\underbrace{\mathbf{TS}}_{\tilde{\mathbf{S}}}) \tag{7}$$

Without any constraint and *a priori* information brought to the optimization step, there is an infinity of factorizations of matrix $\mathbf{M}$. The solution range may nevertheless be imposing regularization to the algorithm, as for example non-negativity imposed to $\mathbf{A}$ and $\mathbf{S}$ coefficients.[23]

Techniques to moderate ambiguity of factorization are required. Adding constraints to original cost function F and incorporating *a priori* information restrains the solutions set, and may be sufficient to solve the NMF problem uniquely.[24] In next section, we propose to study breast simulations of fluorescence detection and autofluorescence removal by NMF, before to test our algorithm on *in vivo* data in last part of this paper. Studies will focus on initialization problem and use of the *a priori* on fluorescence sources.

3 Simulation Studies

We chose to run simulations based on breast data: many clinical experiments on this organ enabled us to design a computer breast model with realistic optical properties of tissues. A fake marked tumor is introduced to the model, and consistent modeling fluorescence acquisitions of the simulated breast are obtained, with modified depth of the marked tumor.

3.1 *Definition of Simulated Matrices* $\mathbf{A}$ *and* $\mathbf{S}$

The breast fluorescence acquisition model is composed of an homogeneous autofluorescence distribution and of a specific fluorescence contribution (to mimic the tumor pointed out by an injected fluorescent marker). Signal ratio between healthy tissue and tagged tumor depends on biomarkers. From bibliography, and from experience, ratios from 3 to 15 (Ref. 25) (for more specific-to-tumor markers) are usual, while new generation of activatable fluorescent markers reach ratios from 24 to 180 (Refs. 26–29), depending on wavelength range, and conditions of experimentation (*ex vivo* or *in vivo*, and site of tumor). We chose for our simulation study a ratio between tumor and healthy tissue approximately equal to 10 when tumor is 1 mm deep in tissues.

The specific fluorescence part is the product of a weight vector $\mathbf{A}_1$ by a fluorescence spectrum $\mathbf{S}_1$ [see Fig. 1(a)]. *A priori*, the autofluorescence part is the product of the weight vector $\mathbf{A}_2$ by a fluorescence spectrum $\mathbf{S}_2$ [see Fig. 1(b)]. Simulated spectra are Gaussian models chosen close to usually used and observed fluorescence spectra of autofluorescence and fluorescent markers (indocyanine green, for example). Finally the total simulated acquisition is obtained by adding the specific fluorescence and the autofluorescence parts [Fig. 1(c)].

3.2 *Contrast*

We introduce straight away the contrast $c_{T,N}$, measured between tumorous area T and normal tissues area N: it characterizes the improvement of tumor detection after autofluores-

Downloaded from SPIE Digital Library on 04 Apr 2011 to 193.48.255.141. Terms of Use: http://spiedl.org/terms

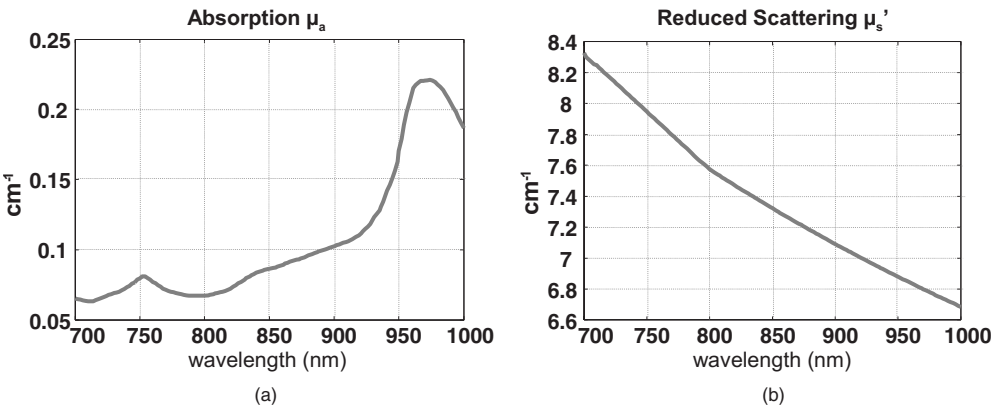

Fig. 2 Average values chosen for breast tissues (a) absorption and (b) scattering (inspired by Ref. 30).

cence removal, on simulation and experimental results. Average intensity of fluorescence signal is measured on both concerned regions of interest (ROIs); $\bar{T}$ and $\bar{N}$ are, respectively, the average intensities in photons per pixel of areas T and N defined in Fig. 1:

$$c_{T,N} = \frac{\bar{T} - \bar{N}}{\bar{T} + \bar{N}}. \tag{8}$$

The closer to one the contrast value gets, the better will be the detection.

3.3 *Optical Parameters*

Clinical studies allowed us to have measurement of breast tissues optical parameters at our disposal. In 2001, Cerussi et al. conducted clinical measurements and obtained average reduced scattering and absorption coefficients of breast, depending on wavelength.[30] From those results, we defined average simulation values of absorption coefficient μ_a (in inverse centimeters) and reduced scattering coefficient μ_s' (in inverse centimeters) on wavelength range 700 to 1000 nm (see Fig. 2).

3.4 *Light Propagation*

Propagation of light in turbid media has been extensively discussed. To simulate decrease of intensity emitted by fluorescent markers embedded in diffusive tissues, and to estimate evolution of contrast between specific fluorescence and autofluorescence depending on depth of fluorescent markers in tissues, classical diffusion approximation is considered. Thus, for homogeneous medium and continuous illumination, the photon density ϕ (in W m^{-2}) satisfies the following derivative equation:

$$\Delta\phi(r) - k(\lambda)^2\phi(r) = -\frac{S(r)}{D(\lambda)} \tag{9}$$

where r is the marker position in medium, the diffusion term D is defined as $D=1/(3(\mu_a+\mu_s)$, k is defined as $k=(3\mu_a\mu_s)^{1/2}$, and $S(r)$ is the isotropic source term. Solution θ of this equation, assuming an infinite medium, is known. Even if the infinite medium hypothesis does not correspond to our case, it has the benefit of having an analytic solution and it enables us to easily describe the light propagation model and changes on spectra. Such a hypothesis would be annoying for a more precise study, especially concerning side effects. For now, however, this approximation is sufficient, and the solution θ to the diffusion equation is used:

$$\theta = \frac{\exp[-k(\lambda)r]}{4\pi D(\lambda)r}. \tag{10}$$

Wavelength-dependent absorption and reduced scattering coefficients of breast tissue are responsible for modification of markers emission spectrum. After fluorescent markers are excited, the emitted photons propagate back in tissues following Eq. (9) to finally reach detectors, as depicted in Fig. 3. Thus, the emission spectrum of detected fluorescent markers $S_d(\lambda)$ varies from *ex vivo* fluorescence spectrum $S(\lambda)$:

$$S_d(\lambda) \propto S(\lambda)\frac{\exp[-k(\lambda)r_{\mathrm{md}}]}{4\pi D(\lambda)r_{\mathrm{md}}}, \tag{11}$$

where r_{md} is the distance from fluorescent markers to detectors.

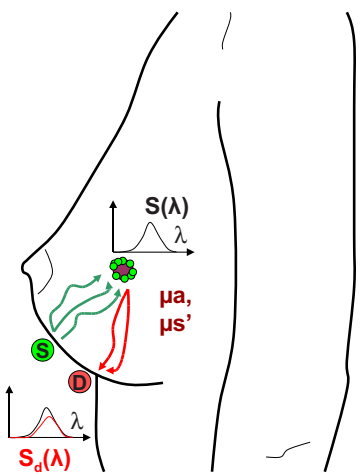

Fig. 3 Schemalic of light propagation after excitation by source *s* in breast tissues, and detection of photons emitted back from fluorescent markers on detectors *d*.

Downloaded from SPIE Digital Library on 04 Apr 2011 to 193.48.255.141. Terms of Use: http://spiedl.org/terms

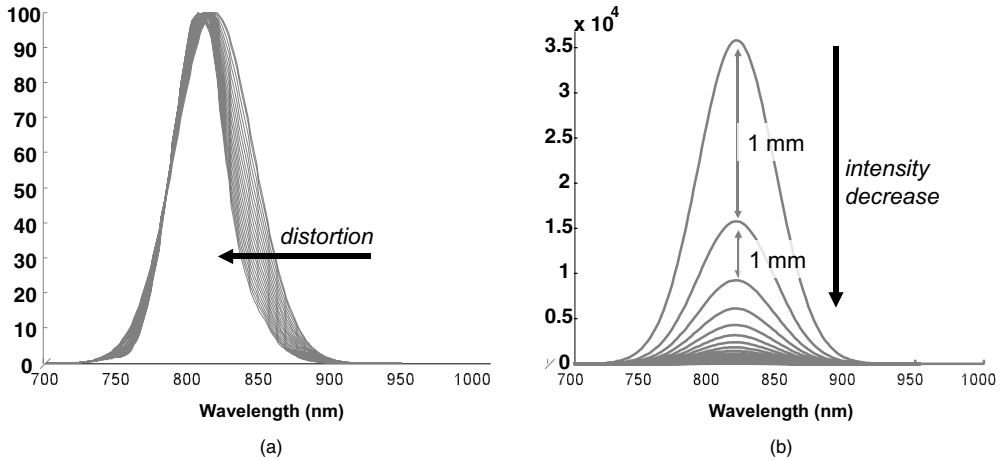

Fig. 4 Normalized fluorescence spectra of simulated markers: (a) spectrum distortion and (b) intensity loss observed for markers moved from the surface to 10 cm deep in simulated breast tissues.

We can now simulate detected fluorescence signal, depending on depth r of markers in breast tissues. Decreasing intensity and spectral distortions resulting from the emitted markers light travel in tissues are shown in Fig. 4 for markers moved from the surface to 10 cm deep in breast tissues.

3.5 *Source Number Determinacy*

When running NMF on our spectroscopic data, we assume that the number of sources we are looking for is equal to the number of the different specific markers injected, plus one for the autofluorescence contribution. Distortion of fluorescence spectra could lead to an increase in the number of sources to unmix; in a case where a same marker is present at different depths in a medium, fluorescence spectrum emitted by the deepest markers appear distorted compared to the less deeply embedded markers.

When the number of sources to unmix is not empirically chosen, a method to define it is to compute the SVD of initial data $\mathbf{M}$:

$$\mathbf{M} = \mathbf{U}\Sigma\mathbf{V}, \tag{12}$$

where matrix $\mathbf{U}$ contains spatial information of fluorescence sources, matrix $\mathbf{V}$ contains spectral information, and Σ gives the ordered singular values. We thus assume our data can be expressed as a separable set of orthonormal spatial and wavelength components. By looking at the singular values, we can define number of sources in the mixed data by selecting the first nonnegligible singular values.

We ran the NMF algorithm on our simulated breast data, and give the SVD of data for each depth of markers tested, from 0.1 to 10 cm. The results are presented in Fig. 5. While spectrum distortion should lead us to consider a third source in our unmixing model, the exponential loss of marker signal intensity does actually not allow time for spectra to become distorted, and SVD only finds two sources.

Since spectrum distortion is insignificant in front of intensity loss of deep embedded markers, looking for an average spectrum S for a same family of markers, but at different depths, is sufficient in diffusive optical imaging. In last section, an *in vivo* unmixing example will confirm that result.

4 *A Priori* Information and Regularization

Many hybrid NMF algorithms, most often dealing with sparsity and smoothness constraints, were developed in the last 5 y, most of them trying to directly adapt from Lee and Seung's multiplicative update rules.[31,32] Kim and Park[33] became interested in sparse NMFs by L_1-norm constraint term minimization; Cichocki et al.[34] presented cost functions no longer based on the Kullback-Leibler divergence but on Csiszár's φ-divergence; while other approaches use alternative cost functions formulations.[35,36] We propose in this section to study the effect of initialization of matrices $\mathbf{A}$ and $\mathbf{S}$ on NMF solutions, and we present regularized NMF update rules, adapted to spectral imaging, that deal with the *a priori* knowledge on the fluorescence spectra considered.

4.1 *Influence of Initialization on NMF Decomposition*

Choice of initialization is once more fundamental and NMF decomposition directly depends on the initial guess on matrices $\mathbf{A}$ and $\mathbf{S}$. In that part, we study the influence of initializa-

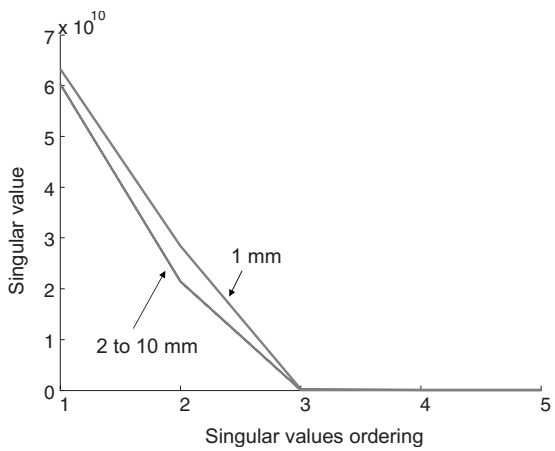

Fig. 5 From 1 mm to 10 cm in breast tissues; even if fluorescence spectra of markers are distorted by traveling tissue, only two fluorescence sources (markers and autofluorescence) are considered in the unmixing model.

Downloaded from SPIE Digital Library on 04 Apr 2011 to 193.48.255.141. Terms of Use: http://spiedl.org/terms

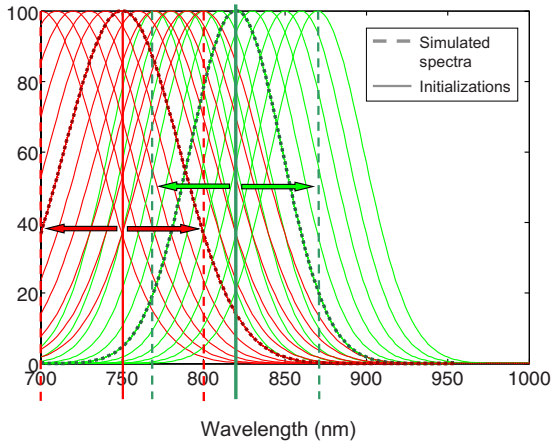

Fig. 6 Influence of initialization on NMF decomposition is studied on simulated example: initial spectra of matrix $\mathbf{S}_0$ are translated (simultaneously for this example) in a range of 100 nm on both sides of expected spectra.

tion on our simulated example. Gaussian spectra similar to simulation spectra of matrix **S** are chosen to initialize the NMF algorithm. We observe the influence of wavelength translation of initialization spectra on the NMF decomposition. For this specific example, initialization spectra for matrix $\mathbf{S}_0$ are translated on a range of 100 nm, on both sides of simulation spectra, as depicted Fig. 6.

To underline the dependence of NMF result to initialization, three cases are presented. For each case, resulting contrast for different depths of fluorescent markers in tissue (from 0.1 to 4 cm) is obtained.

First we observe healthy tissue/tumor contrast on raw data, without any unmixing processing: contrast and detection are naturally decreasing with depth [see Fig. 7(a)]. Then NMF processing is applied on data but with random initialization (random nonnegative values for matrices **A** and **S**, 30 draws per depth): unmixing processing improves detection [see Fig. 7(b)]. Finally, NMF algorithm with this time Gaussian initialization for **S** (Gaussian models are translated in a 100-nm range) is tested: even with less appropriate initialization (in that case, when both simulation spectra are translated 50 nm up that simulation models for initialization), contrast is improved compared to both prior cases [see Fig. 7(c)].

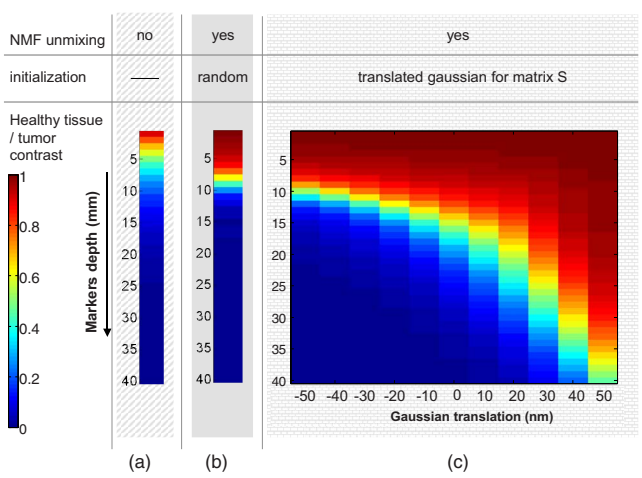

Fig. 7 Influence of initialization on resulting contrast (breast simulation example).

NMF processing without any *a priori* information still improves resulting contrast compared to results on raw data. Then, Gaussian initialization refines results.

But the initialization guess is not obvious. When the initialization spectrum is translated 50 nm from the expected spectra, this initialization leads to better result than if initialization is equal to simulated spectra. One explanation may be that we move away from the crossing area where fluorescence spectra overlap and where indeterminacy between autofluorescence and fluorescence of markers is high. Moreover, the obtained contrast is not symmetric on both sides of the 0-shifted spectra: the fluorescence of markers is spatially restraint compared to the autofluorescence background (see Fig. 1), and moving away to an initialization zone where the fluorescence marker spectrum does not clash with the autofluorescence spectrum emission is advantageous.

From that example, we show initialization choice is very sensitive, but necessary to improve marker detection. Even if an approximate initialization already improves contrast between tumor area and healthy tissue compared to the algorithm with random initialization (see Fig. 7), a more accurate initialization selection may push back the detection limits.

Such selection can be obtained with a multistart initialization step prior to the NMF algorithm we proposed earlier.

4.2 *New Multiplicative Update Rules*

As explained in previous section, *a priori* knowledge of fluorescence sources enables refining the range of solutions. By choosing appropriate initialization, detection of marked tumors can be improved. Another way to restrain the solutions set is to constrain the initial cost function F; we propose in this section to lightly modify the cost function, and find a new regularized algorithm to minimize updated cost function.

4.3 *A Priori Knowledge on the Fluorescence Sources*

In the optical spectroscopy context, injected markers are known and could thus ease the unmixing problem. But whether it refers to the specific markers, or to the autofluorescence of biological tissues, we can actually not define a precise model of the fluorescence spectra. Spectra of the specific markers may vary from the *ex vivo* known spectra once injected *in vivo* and illuminated. In the *in vivo* medium, the markers may create new products that are able, in turn, to emit unknown fluorescence signals. Fluorescence spectra may also vary with the pH values in the medium, and their intensity may decline with time due to the photobleaching phenomenon. Even if chemical modifications appear on fluorescent molecules under illumination or due to the receiver medium components, we usually observe after that a constant emission spectrum (except in the case of distortion of spectra due to depth of tissue, explained later). Finally, the optical parameters of the biological tissues—diffusion and absorption—cause emissions fluorescence spectra to vary with the depth of the fluorescent source. This last problem was treated in simulation and presented in previous section.

Downloaded from SPIE Digital Library on 04 Apr 2011 to 193.48.255.141. Terms of Use: http://spiedl.org/terms

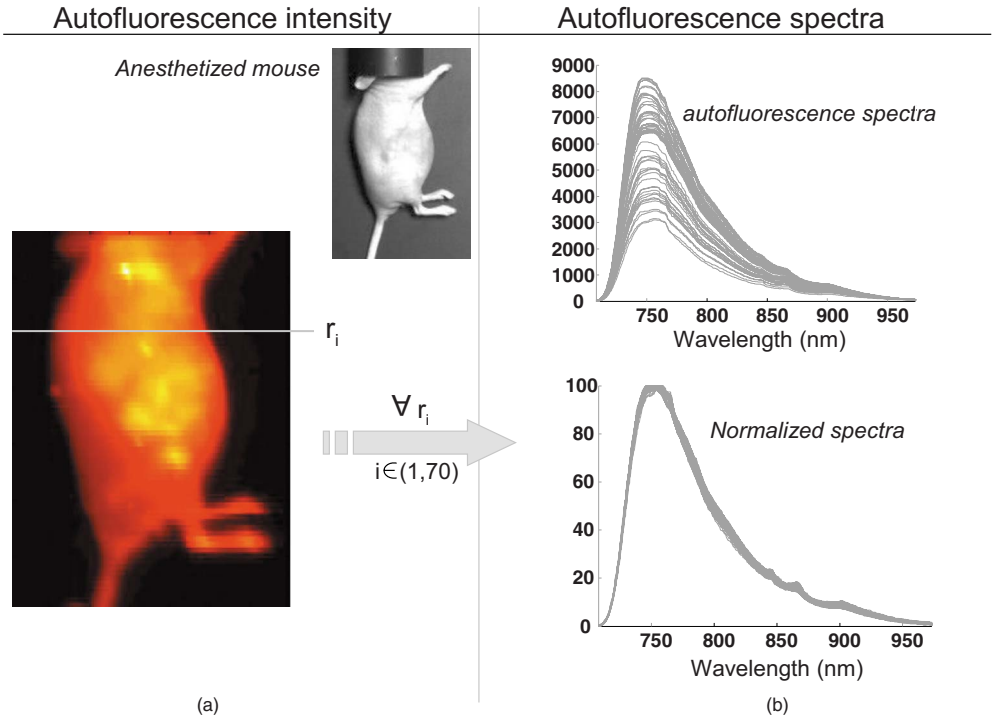

Fig. 8 Autofluorescence acquisition of noninjected mouse: (a) intensity map and (b) associate average (raw by raw) spectra.

Regarding the autofluorescence of tissues, as for the specific markers, the illumination may be responsible for the creation of new products. Furthermore, autofluorescence spectrum may vary according to the pH of analyzed area,[37] or according to the patient.[38–40] But once measured, the *in vivo* autofluorescence spectrum usually not differs across the organism being observed. An example of mouse autofluorescence acquisition is given Fig. 8, where normalized spectra remain constant across the mouse body. We assume this property should also be true for specific human area observed (prostate, breast, etc.).

For slightly different autofluorescence emission spectra, the average spectrum is sufficient for NMF decomposition (a similar case is that for distorted fluorescent markers with depth).

Even if fluorescence spectra may not be initially perfectly defined, we still have some *a priori* information concerning them, from *ex vivo* and empirical measurements. The emission wavelength range, and the shape of the expected spectra may be globally known and used to refine solutions set.

4.4 *NMF Initialization Step*

The initialization step uses *a priori* models to guide the solutions, but lets the algorithm, thanks to its blind specificity, to adapt solutions depending on the original experimental data. We would like the algorithm to take into account that piece of information: we thus define a new cost function to minimize, that directly depends on $\mathbf{S}_0$.

4.5 *New Cost Function and Regularized Update Rules*

The new cost function F_2 is now defined as the sum of the square of the Euclidean distance between $\mathbf{M}$ and $\mathbf{A}\times\mathbf{S}$ plus a regularization term on the initialization $\mathbf{S}_0$, weighted by a regularization vector $\boldsymbol{\alpha}$ of length P:

$$F_2=\sum_{n=1}^{N_s}\sum_{\lambda=1}^{N_\lambda}\left(m_{n\lambda}-\sum_{p=1}^{P}a_{\mathrm{np}}s_{p\lambda}\right)^2+\sum_{p=1}^{P}\boldsymbol{\alpha}_p\times\sum_{\lambda=1}^{N_\lambda}(s_{p\lambda}-s_{0_{p\lambda}})^2, \tag{13}$$

or rewritten in matrix form:

$$F_2=(\mathbf{M}-\mathbf{AS})^t(\mathbf{M}-\mathbf{AS})+(\mathbf{S}-\mathbf{S}_0)^t\mathbf{D}_{\boldsymbol{\alpha}}(\mathbf{S}-\mathbf{S}_0), \tag{14}$$

where $\mathbf{D}_{\boldsymbol{\alpha}}=\mathrm{diag}[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_p]$ is a diagonal matrix containing the P regularization parameters associated to the P spectra of matrix $(\mathbf{S}-\mathbf{S}_0)$.

The regularization term lets $\mathbf{S}$ vary from $\mathbf{S}_0$, according to the confidence we have in the initialization. Then F_2 is minimized by alternatively updating matrices $\mathbf{A}$ and $\mathbf{S}$, with respect to $(\mathbf{A},\mathbf{S})\geqslant 0$. We propose to use original multiplicative regularized update rules, defined as follows.

Theorem 2. The function F_2 is nonincreasing under the update rules:

$$\mathbf{S}_{p\lambda}\leftarrow\mathbf{S}_{p\lambda}\frac{(\mathbf{A}^t\mathbf{M}+\mathbf{D}_{\boldsymbol{\alpha}}\mathbf{S}_0)_{p\lambda}}{(\mathbf{A}^t\mathbf{AS}+\mathbf{D}_{\boldsymbol{\alpha}}\mathbf{S})_{p\lambda}}\quad\mathbf{A}_{\mathrm{np}}\leftarrow\mathbf{A}_{\mathrm{np}}\frac{(\mathbf{MS}^t)_{\mathrm{np}}}{(\mathbf{ASS}^t)_{\mathrm{np}}}, \tag{15}$$

with $\mathbf{D}_{\boldsymbol{\alpha}}=\mathrm{diag}[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\ldots,\boldsymbol{\alpha}_p]$.

The value of $\boldsymbol{\alpha}$ must be set according to the level of confidence in the *a priori* information on the initial spectra. We can also use different degree of confidence on the different spectra by using a vector $\boldsymbol{\alpha}$.

Downloaded from SPIE Digital Library on 04 Apr 2011 to 193.48.255.141. Terms of Use: http://spiedl.org/terms

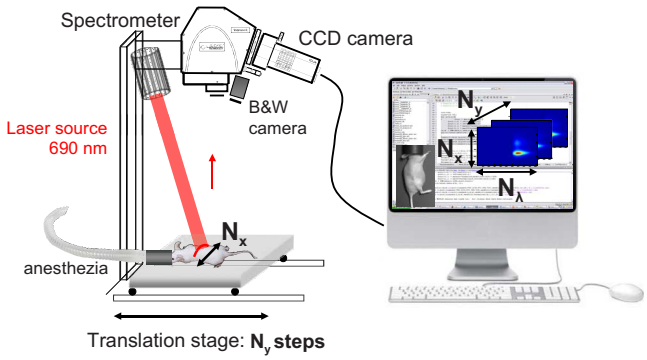

Fig. 9 Experimental setup for acquisition of the animal and data processing.

We prove Theorem 2 in the appendix, by using an auxiliary function as in Refs. 11 and 12, and conserving the maximum of the mathematical notations of Lee and Seung in Ref. 12. To illustrate the efficiency of NMF on spectroscopic data, we propose now an example on *in vivo* experimental data, where up to three different fluorescence sources need to be unmixed.

5 Spectral Unmixing on Experimental Data

For *in vivo* experiments, an autofluorescence signal is necessarily measured. Then several specific markers may be used to simulate marked targets, such as tumors. In this section, we test NMF to unmix three overlapping different fluorescence sources, including the autofluorescence on mice.

To acquire spectrally resolved measurements, the animal was illuminated with a planar laser at 690 nm. The emitted back fluorescence signal was collected along a line of N_x points by a spectrometer coupled with a charge-coupled device camera (Andor Technologies): a $N_x \times N_\lambda$ acquisition was measured (see Fig. 9). For this experiment, N_x was equal to 250 and N_λ to 1024, which corresponds to a wavelength range around 600 to 975 nm. A translation stage, covering N_y steps, was then used to get a scanning of the whole animal: N_y acquisitions were obtained, each of size $N_x \times N_\lambda$. Before to run the NMF algorithm, mixed data of size $(N_y \times N_x \times N_\lambda)$ were reordered as a 2-D array of size $(N_y \times N_x, N_\lambda)$ (N_s is defined in NMF equation is now equal to $N_y \times N_x$). Separately, a black and white picture of the animal may be taken, to superimpose fluorescence signal and environment image if necessary.

For each *in vivo* experiment, a precise protocol is defined. All protocols have in common the following cares. The animal procedure is in compliance with the guidelines of the European Union (regulation No. 86/609), taken in the French law (decree 87/848) regulating animal experimentation. All efforts are made to minimize animal suffering. The animal manipulation is performed with sterile techniques and approved by the Grenoble Animal Care and Use Committee (France) (registration number 20_iRTSV Léti-FNG-02). An adult female nude mouse (Janvier, Le Genest saint-isle, France) are used throughout the experiments. They are housed in approved facilities, at 21 ± 1 °C under diurnal lighting conditions. The mice arrive at the animal facility 2 weeks before the experiments start and had free access to food and water.

5.1 *Multimarker Experiment*

5.1.1 *Presentation*

We present a first feasibility experiment on a mouse. Two glass capillary tubes respectively filled with 5 μl of indocyanine green loaded into lipid nanoparticules[41] (ICG-LNP) at 0.35 μmol/l and 5 μl of Alexa 750 at 0.1 μmol/l were inserted subcutaneously to simulate marked targets [see Fig. 10(a)]. Three distinct fluorescent sources—autofluorescence, ICG-LNP, and Alexa 750—whose emission spectra are overlapping had to be unmixed. To draw a parallel between acquisitions with or without specific fluorescence, a first acquisition of the animal was run as a reference, without any specific markers inserted.

In this precise example, we chose $\alpha = 10^{10}$ to constraint the autofluorescence spectrum to remain close to initialization

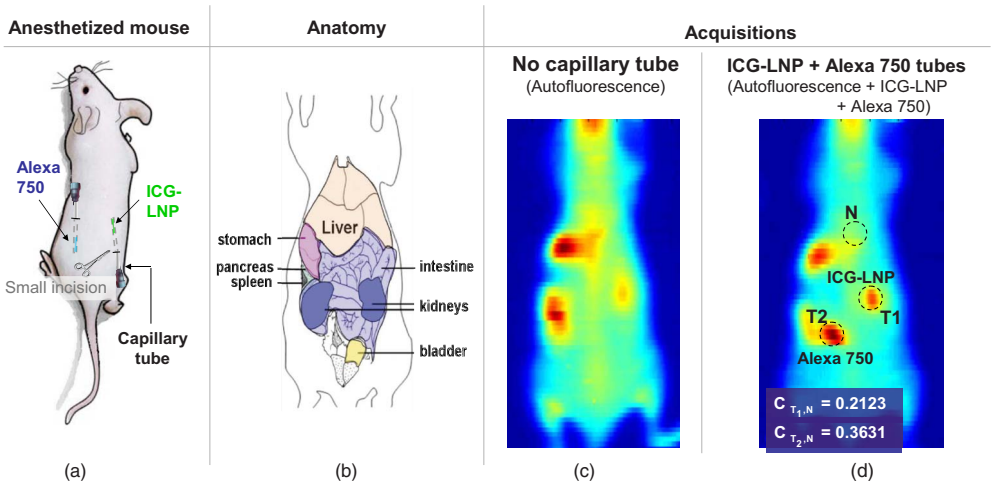

Fig. 10 (a) Schematic of the mouse with two capillary tubes filled with ICG-LNP and Alexa 750; (b) anatomy scheme of the mouse, (c) scanning result without any capillary (autofluorescence signal only), and (d) scanning result with the two capillary tubes inserted (autofluorescence + ICG-LNP + Alexa 750).

Downloaded from SPIE Digital Library on 04 Apr 2011 to 193.48.255.141. Terms of Use: http://spiedl.org/terms

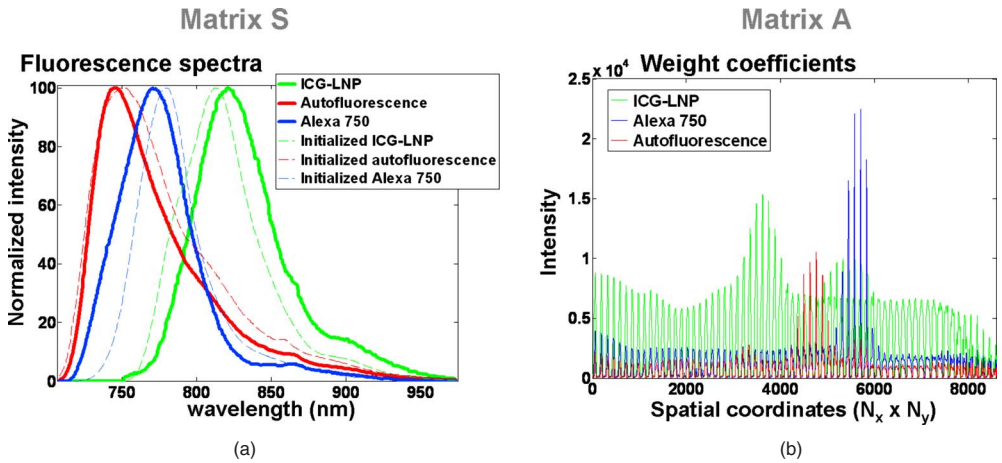

Fig. 11 Results after NMF: (a) fluorescence spectra obtained (matrix **S**): initialization (dotted lines), and results of the NMF decomposition (continuous lines), and (b) weight coefficients (matrix **A**).

and have a plausible profile, and $\alpha=0$ for ICG-LNP, and Alexa 750 spectra. The regularization term helped to smooth over aberrations due to the two components crosstalk on resulting spectra. We chose initialization from empirical results, but could not use multistart initialization: regularization on initialization asks for coherent initialization spectra, while multistart initialization may use translated spectra unconnected to real expected fluorescence spectra. The choice has to be made between different regularization methods, depending on application.

5.1.2 *Results*

Figure 10(c) shows the reference acquisition of the mouse without any capillary: the autofluorescence is the only fluorescence signal measured. Autofluorescent areas may be attributed to some specific organs, known to emit natural fluorescence, such as the stomach, the liver, the intestine, or the kidneys of the animal, if we compare the autofluorescence scanning to the anatomy scheme of a mouse [Fig. 10(b)]. In the same figure, we present the resulting scanning of the mouse with the two capillary tubes of ICG-LNP and Alexa 750 [Fig. 10(d)], on which the regularized NMF algorithm is run.

Figure 11 presents the overlapping spectra that NMF successfully unmixed. The three normalized spectra of matrix **S** presented in Fig. 11(a) are weighted by coefficients of matrix **A** [Fig. 11(b)]. The three resulting contributions of fluorescence are presented in Fig. 12: the autofluorescence [Fig. 12(a)], the ICG-LNP [Fig. 12(b)], and the Alexa 750 [Fig. 12(c)]. Three sources unmixing allowed improving detection for each marker, Alexa 750 and ICG LNP. Indeed, initial contrast values (measured on mixed data, cf. Fig. 10 between markers area and normal tissue area were equal to 0.2123 for ICG-LNP and 0.3631 for Alexa 750). After unmixing, contrast values respectively reached 0.6366 (ICG-LNP) and 0.6763 (Alexa 750) (see Figs. 12(b) and 12(c)].

To conclude concerning this feasibility experiment, a last comparison is made, between the original autofluorescence scanning (without any capillary tube inserted to the animal) and the separated autofluorescence results obtained for both past experiments (see Fig. 13). First, a consistent intensity level is obtained after the NMF decomposition. Indeed, the intensity obtained after NMF decomposition for the autofluorescence part [Fig. 13(b)] is close to the intensity observed on the reference autofluorescence acquisition [Fig. 13(a)]. Finally, the principal autofluorescent area observed on the autofluorescence scanning [Fig. 13(a)] are also present on autofluorescence contribution obtained after the NMF decomposition [Fig. 13(b)].

5.2 *Multidepth Experiment*

In a second experiment, we proposed to test the NMF algorithm robustness on a precise case of the same fluorescent marker at two different depths in mouse tissues. In such a case, the deeper the markers are embedded in tissues, the more their fluorescence spectra are distorted and their intensity is exponentially decreased.

Two glass capillary tubes both filled with 5 μl of ICG-LNP (Ref. 41) at 5 μmol/l were inserted to simulate marked targets. The first tube was inserted subcutaneously (around 1 mm deep), while the second one was placed in the rectum of the animal (around 6 mm deep from the surface), as depicted in Fig. 14(a). The same experimental setup was used to scan the mouse (see Fig. 9), and the obtained data are presented Fig. 14(b). The 6-mm-deep marker signal was mixed with the autofluorescence signal and masked by the 1-mm-deep intense marker signal, and the 5 mm-deep-difference was responsible for light distortion (emission peak translation) and loss of intensity between both emitted spectra, measured in the rectum or subcutaneously [see Fig. 14(b)].

5.2.1 *Results*

As for the previous experiment, we ran the NMF algorithm on the mixed data to obtain separated fluorescence contributions of autofluorescence and the specific ICG-LNP fluorescence. The results are presented Fig. 15, with a unique average spectrum obtained in matrix **S** after NMF decomposition for the ICG-LNP despite the slightly different spectra emitted from the subcutaneous and rectum capillaries (see Fig. 15, ICG-LNP dotted line), and accurate unmixing results were ob-

Downloaded from SPIE Digital Library on 04 Apr 2011 to 193.48.255.141. Terms of Use: http://spiedl.org/terms

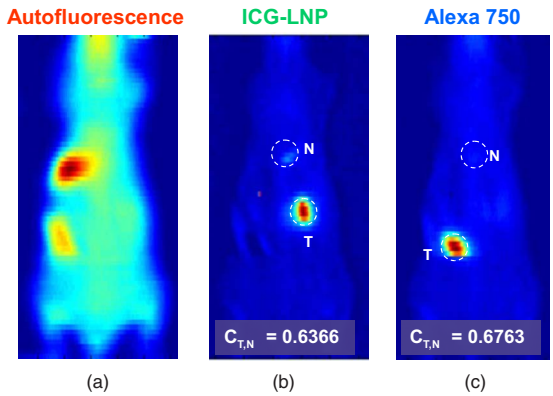

Fig. 12 Results after NMF: (a) autofluorescence intensity contribution, (b) ICG-LNP intensity contribution, and (c) Alexa 750 intensity contribution.

tained [see Figs. 15(a) and 15(b)]. Deep markers that were originally mixed with autofluorescence signal are now easily detectable. Indeed, after autofluorescence removal, detection of 6-mm-deep markers was considerably improved. The initial contrast value between the ICG-LNP rectum area and the normal tissue area was equal to 0.36 before unmixing [see Fig. 14(b)], and reached 0.77 once the fluorescence contributions were separated [see Fig. 15(b)]. Moreover, in the obtained ICG-LNP contribution, all the autofluorescence critical zones (the stomach or intestine, for example) were quenched.

6 Conclusions

We presented a blind positive source separation method applied to *in vivo* spectrally resolved data. Beyond the specific fluorescence signal of specific markers used in optical imaging, the autofluorescence of biological tissues was also detected in the wavelength range we used, and must be removed to achieve accurate detection results. This property increased with the depth of biological tissues explored, since the specific signal decreases exponentially while the autofluorescence signal remains constant. To remove autofluorescence and unmix different fluorescent markers whose fluorescence

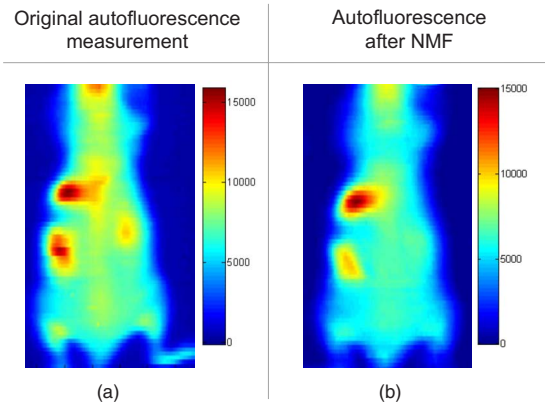

Fig. 13 (a) Autofluorescence measurement, without any specific fluorescence (no capillary tubes) added and (b) autofluorescence contribution obtained after NMF decomposition (two capillary tubes of ICG-LNP and Alexa 750 experiment).

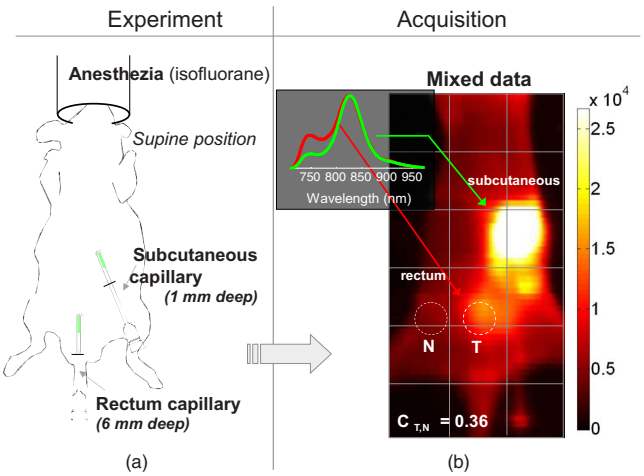

Fig. 14 (a) Two capillary tubes of ICG-LNP are placed at two different depths and (b) intensity data obtained and average fluorescence spectra measured in the capillary tubes.

spectra are not perfectly known, a blind source separation method, the NMF, was chosen. By notably taking into account the nonnegativity of the data, it is particularly suitable for fluorescence imaging data. We briefly presented this classical method, from the choice of the cost function and the optimization step that minimizes it and leads to a nonunique NMF decomposition. Indeed, when using blind source separation, we estimated fluorescence sources without knowing the mixing process; without some *a priori* knowledge, it is not possible to uniquely estimate the sources. We presented studies of simulated breast data that underline interest of *a priori* information for initialization choice and additional constraints that may be applied to cost functions. An original regularized NMF algorithm that takes into account some prior knowledge about the fluorescence spectra was proposed. We did not discuss a method to chose the value of the regularization parameter, which is currently empirically chosen. Finally, our theory was successfully validated with *in vivo* experiments on mice. The aim of those experiments was to remove the autofluorescence contribution from the experimental data, and to unmix upto two specific markers, or the same fluorescent marker at different depths in tissues. NMF computed satisfactory results and enabled us to localize specific marker contributions initially lost in the autofluorescence signal.

As optical imaging tries to detect deeper and deeper embedded targets, NMF is a useful preprocessing step to remove unwanted autofluorescence and unmix different spectra of interest. By returning separated fluorescence contribution data, the method presents the possibility to perform accurate tomographic reconstructions and thus to confirm the 3-D position of marked tumors.

Appendix: Theorem 1: Proofs of Convergence

In this section, we propose a proof of convergence of Theorem 2. Note that a proof for convergence of the update rule for $\mathbf{A}$ was already given by Lee and Seung.[12] We introduce the following definition:

Downloaded from SPIE Digital Library on 04 Apr 2011 to 193.48.255.141. Terms of Use: http://spiedl.org/terms

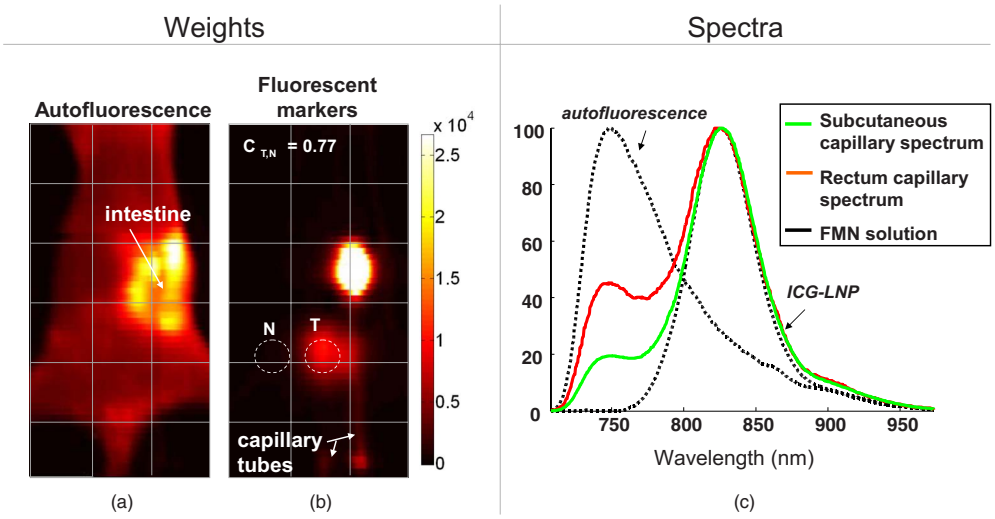

Fig. 15 Unmixing results: (a) and (b) weights (matrix **A**) of respectively autofluorescence and ICG-LNP fluorescence, and (c) unmixed spectra (matrix **S**) obtained at convergence, and a comparison with unmixed spectra measured in capillary tubes on the mouse, subcutaneously and in the rectum.

Definition 1. $G(s,s^i)$ is an auxiliary function for $F_2(s)$ if the following conditions are satisfied:

$$G(s,s^i) \geqslant F_2(s) \quad \text{and} \quad G(s,s) = F_2(s). \tag{16}$$

The auxiliary function definition is useful for the following lemma:[12]

Lemma 1. If G is an auxiliary function for F_2, then F_2 is nonincreasing under the update:

$$s^{i+1} = \operatorname{argmin}_s G(s,s^i) \tag{17}$$

Lemma 1 is illustrated in Fig. 16.

By defining an appropriate auxiliary function G for F_2, as defined earlier, the update rules presented in Theorem 1 simply follow from Lemma 1 [Eq. (17)]. The auxiliary function G is presented in the following lemma.

Lemma 2. If $\mathbf{K}(s^i)$ is the diagonal matrix

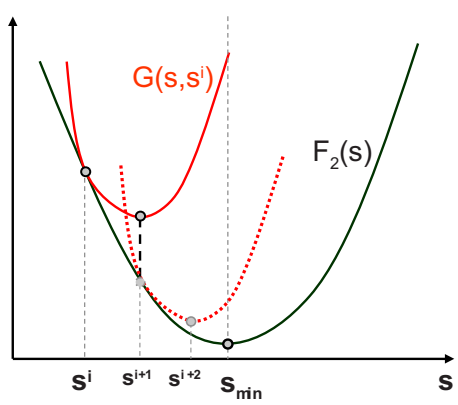

Fig. 16 Minimizing the auxiliary function $G(s,s^i) \geqslant F_2(s)$ ensures that $F_2(s^{i+1}) \leqslant F_2(s^i)$ for $s^{i+1} = \operatorname{argmin}_s G(s,s^i)$.

$$\mathbf{K}_{ab}(s^i) = \frac{\delta_{ab}(\mathbf{A}^t\mathbf{A}s^i + \boldsymbol{\alpha} \times s^i)_a}{s_a^i}, \tag{18}$$

then, setting $\Delta s = (s - s^i)$,

$$G(s,s^i) = F_2(s^i) + \Delta s^t \nabla F_2(s^i) + \frac{1}{2}\Delta s^t \mathbf{K}(s^i)\Delta s, \tag{19}$$

is an auxiliary function for

$$F_2(s) = \frac{1}{2}\sum_i \left(v_i - \sum_a \mathbf{A}_{ia}s_a\right)^2 + \frac{\alpha}{2}\sum_a (s_a - s_{0_a})^2. \tag{20}$$

Proof of lemma 2. $G(s,s^i)$ is an auxiliary function for $F_2(s)$ if the conditions defined in Eq. (16) are verified. If $G(s,s) = F_2(s)$ is obvious, the second condition $G(s,s^i) \geqslant F_2(s)$ must be demonstrated.

Let ∇F_2 be the gradient of F_2 and $\mathbf{Hess}(F_2)$ the Hessian matrix. We obtain

$$F_2(s) = F_2(s^i) + \Delta s^t \nabla F_2(s^i) + \frac{1}{2}\Delta s^t \mathrm{Hess}(F_2)\Delta s, \tag{21}$$

with

$$F_2(s^i) = \frac{1}{2}\|m - \mathbf{A}s^i\|^2 + \frac{\alpha}{2}\|s^i - s_0^i\|^2 = \frac{1}{2}(m - \mathbf{A}s^i)^t(m - \mathbf{A}s^i) + \frac{\alpha}{2}(s^i - s_0^i)^t(s^i - s_0^i), \tag{22}$$

and finally the gradient is

$$\nabla F_2(s^i) = \mathbf{A}^t(\mathbf{A}s^i - m) + \alpha(s^i - s_0^i), \tag{23}$$

while the Hessian matrix gives

Downloaded from SPIE Digital Library on 04 Apr 2011 to 193.48.255.141. Terms of Use: http://spiedl.org/terms

$$\mathrm{Hess}(F_2) = \mathbf{A}^t\mathbf{A} + a\mathbf{I}_p, \tag{24}$$

where $\mathbf{I}$ is the $p \times p$ identity matrix.

Thus,

$$\begin{aligned} G(s,s^i) - F_2(s) &= F_2(s^i) + \Delta s^t \nabla F_2(s^i) + \frac{1}{2}\Delta s^t \mathbf{K}(s^i)\Delta s - F_2(s^i) \\ &\quad - \Delta s^t \nabla F_2(s^i) - \frac{1}{2}\Delta s^t \mathrm{Hess}(F_2)\Delta s \\ &= \frac{1}{2}\Delta s^t[\mathbf{K}(s^i) - \mathrm{Hess}(F_2)]\Delta s = \frac{1}{2}\Delta s^t(\mathbf{K}(s^i) \\ &\quad - \mathbf{A}^t\mathbf{A} - \alpha\mathbf{I})\Delta s. \end{aligned} \tag{25}$$

Hence, the following equivalence is given

$$G(s,s^i) \geqslant F_2(s) \Leftrightarrow \Delta s^t[\mathbf{K}(s^i) - \mathbf{A}^t\mathbf{A} - \alpha\mathbf{I}]\Delta s \geqslant 0. \tag{26}$$

We can now demonstrate that $\Delta s^t[\mathbf{K}(s^i) - \mathbf{A}^t\mathbf{A} - \alpha\mathbf{I}]\Delta s$ is positive:

$$\begin{aligned} &\Delta s^t[\mathbf{K}(s^i) - \mathbf{A}^t\mathbf{A} - \alpha\mathbf{I}]\Delta s \\ &= \sum_{ab} \Delta s_a[\mathbf{K}(s) - \mathbf{A}^t\mathbf{A} - \alpha\mathbf{I}]_{ab}\Delta s_b \\ &= \sum_{ab} \Delta s_a\left[\frac{\delta_{ab}(\mathbf{A}^t\mathbf{A}s^i + \alpha s^i)_a}{s_a^i}\right]\Delta s_b - \sum_{ab} \Delta s_a[(\mathbf{A}^t\mathbf{A} \\ &+ \alpha\mathbf{I})_{ab}]\Delta s_b \\ &= \sum_{ab} \Delta s_a\left\{\frac{\delta_{ab}\left[\sum_c\sum_d (\mathbf{A}_{da}\mathbf{A}_{dc}s_c^i) + \alpha s_a^i\right]}{s_a^i}\right\}\Delta s_b \\ &- \sum_{ab} \Delta s_a\left(\sum_d \mathbf{A}_{da}\mathbf{A}_{db} - \alpha\delta_{ab}\right)\Delta s_b \\ &= \sum_{a,c,d} \mathbf{A}_{da}\mathbf{A}_{dc}\left(\Delta s_a \Delta s_a \frac{s_c^i}{s_a^i} - \Delta s_a \Delta s_c\right) \\ &+ \sum_a \Delta s_a\left(\frac{\alpha \times s_a^i}{s_a^i} - \alpha\right)\Delta s_a = \frac{1}{2}\sum_{a,c,d} \mathbf{A}_{da}\mathbf{A}_{dc}\left[\Delta s_a\left(\frac{s_c}{s_a}\right)^{1/2} \right. \\ &\left. - \Delta s_c\left(\frac{s_a}{s_c}\right)^{1/2}\right]^2 \geqslant 0. \end{aligned} \tag{27}$$

Following is a demonstration of Theorem 1.

Proof of theorem 1. Lemma 1 gives $s^{i+1} = \mathrm{argmin}_s G(s,s^i)$.

Since

$$\begin{aligned} \forall a, \frac{\delta G(s,s^i)}{\delta s_a^i} = [\nabla F_2(s)]_a + [\mathbf{K}(s)\Delta s]_a = 0 \Rightarrow \Delta s = \\ - \mathbf{K}(s)^{-1} \nabla F_2(s), \end{aligned} \tag{28}$$

we obtain

$$\forall a, s_a^{i+1} - s_a^i = -\sum_b \mathbf{K}_{ab}(s^i)^{-1} \times [\mathbf{A}^t(\mathbf{A}s^i - m) + \alpha(s^i - s_0^i)]_b. \tag{29}$$

Thus,

$$s_a^{i+1} = s_a^i \frac{(\mathbf{A}^t v + \alpha s_0^i)_a}{(\mathbf{A}^t\mathbf{A}s^i + \alpha s^i)_a}. \tag{30}$$

References

1. V. Ntziachristos, "Visualization of antitumor treatment by means of fluorescence molecular tomography with an annexin v-cy5.5 conjugate," *Proc. Natl. Acad. Sci. U.S.A.* **101**(33), 12294–9 (2004).
2. J. Hung, S. Lam, J. Leriche, and B. Palcic, "Autofluorescence of normal and malignant bronchial tissue," *Lasers Surg. Med.* **11**(2), 99–105 (1991).
3. D. Wood, G. Feke, D. Vizard, and R. Papineni, "Refining epifluorescence imaging and analysis with automated multiple-band flat-field correction," *Nat. Methods* **5**, i–ii (2008).
4. C. Vandelest, E. Versteeg, J. Veerkamp, and T. Vankuppevelt, "Elimination of autofluorescence in immunofluorescence microscopy with digital image processing," *J. Histochem. Cytochem.* **43**(7), 727–730 (1995).
5. R. Levenson and P. Cronin, "Spectral imaging of deep tissue," U.S. Patent No. 7,321,791 (2008).
6. C. Hoyt, "Systems and methods for *in vivo* optical imaging and measurement," U.S. Patent No. 7,477,931 (2009).
7. P. Cronin and P. J. Miller, "Spectral imaging," U.S. Patent No. 6,930,773 (2005).
8. L. De Lathauwer, B. De Moor, and J. Vandewalle, "Blind source separation by higher-order singular value decomposition," *Proc. Eur. Signal Processing Conf. (EUSIPCO)*, Vol. **1**, pp. 175–178 (1994).
9. R. C. Henry, "Current factor analysis receptor models are ill-posed," *Atmos. Environ.* **21**(8), 1815–1820 (1987).
10. P. Paatero and U. Tapper, "Positive matrix factorization: a non-negative factor model with optimal utilization of error estimates of data values," *Environmetrics* **5**(2), 111–126 (1994).
11. A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. R. Stat. Soc.* **39**(1), 1–38 (1977).
12. D. Lee and H. Seung, "Algorithms for non-negative matrix factorization," *Adv. Neural Inf. Process. Syst.* **13**, 556–562 (2001).
13. W. Lawton and E. Sylvestre, "Self modeling curve resolution," *Technometrics* **13**(3), 617–633 (1971).
14. A. Pascual-Montano, P. Carmona-Saez, M. Chagoyen, F. Tirado, J. M. Carazo, and R. D. Pascual-Marqui, "bionmf: a versatile tool for non-negative matrix factorization in biology," *BMC Bioinf.* **7**(1), 366–375 (2006).
15. A. Cichocki and A. H. Phan, "Fast local algorithms for large scale nonnegative matrix and tensor factorizations," *IEICE Trans. Fundamentals* **E92A**(3), 708–721 (2009).
16. V. Pauca, F. Shahnaz, M. Berry, and R. Plemmons, "Text mining using nonnegative matrix factorizations," in *Proc. 4th SIAM Int. Conf. on Data Mining*, pp. 452–456 (2004).
17. H. Xu, C. Kuo, and B. W. Rice, "Improved sensitivity by applying spectral unmixing prior to fluorescent tomography," in *Proc. OSA Biomed. Opt.*, paper BMC1 (2008).
18. C. Gobinet, E. Perrin, and R. Huez, "Application of nonnegative matrix factorization to fluorescence spectroscopy," in *Proc. Eur. Signal Processing Conf. (EUSIPCO)*, pp. 6–10 (2004).
19. C. Gobinet, V. Vrabie, A. Tfayli, O. Piot, R. Huez, and M. Manfait, "Preprocessing and source separation methods for Raman spectra analysis of biomedical samples," in *Proc. EMBS 29th Annu. Int. Conf. of the IEEE*, pp. 6207–6210 (2007).
20. H. Xu and B. W. Rice, "*In vivo* fluorescence imaging with a multivariate curve resolution spectral unmixing technique," *J. Biomed. Opt.* **14**(6), 064011 (2009).
21. P. Paatero, "Least squares formulation of robust non-negative factor analysis," *Chemom. Intell. Lab. Syst.* **37**(1), 23–35 (1997).

Downloaded from SPIE Digital Library on 04 Apr 2011 to 193.48.255.141. Terms of Use: http://spiedl.org/terms

22. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, Cambridge University Press, New York (2007).
23. S. Moussaoui, D. Brie, and J. Idier, "Non-negative source separation: range of admissible solutions and conditions for the uniqueness of the solution," in *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, Vol. **5**, pp. 289–292 (2005).
24. A. Cichocki, S. Amari, A.-H. Phan, and R. Zdunek, *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multiway Data Analysis and Blind Source Separation*, Wiley-Blackwell, Chichester, UK (2009).
25. Z. Jin, V. Josserand, S. Foillard, D. Boturyn, P. Dumy, M. Favrot, and J. Coll, "*In vivo* optical imaging of integrin alphaV-beta3 in mice using multivalent or monovalent cRGD targeting vectors," *Molec. Cancer* **41**, 1–9 (2007).
26. Y. Urano, D. Asanuma, Y. Hama, Y. Koyama, T. Barrett, M. Kamiya, T. Nagano, T. Watanabe, A. Asegawa, and P. Choyke, "Selective molecular imaging of viable cancer cells with pH-activatable fluorescence probes," *Nat. Med.* **15**, 104–109 (2008).
27. R. Weissleder, C. Tung, U. Mahmood, and A. Bogdanov, "*In vivo* imaging of tumors with protease-activated near-infrared fluorescent probes," *Nat. Biotechnol.* **17**, 375–378 (1999).
28. J. Razkin, V. Josserand, D. Boturyn, Z. Jin, P. Dumy, M. Favrot, J. Coll, and I. Texier, "Activatable fluorescent probes for tumour-targeting imaging in live mice," *ChemMedChem* **1**, 1069–1072 (2006).
29. Z. Jin, J. Razkin, V. Josserand, D. Boturyn, A. Grichine, I. Texier, M. Favrot, P. Dumy, and J. Coll, "*In vivo* noninvasive optical imaging of receptor-mediated RGD internalization using self-quenched Cy5-labeled RAFT-c(-RGDfK-) (4)," *Mol. Imaging* **6**, 43–55 (2007).
30. A. Cerussi, A. Berger, F. Bevilacqua, N. Shah, D. Jakubowski, J. Butler, R. Holcombe, and B. Tromberg, "Sources of absorption and scattering contrast for near-infrared optical mammography," *Acad. Radiol.* **8**, 211–218 (2001).
31. A. Cichocki, R. Zdunek, and S. Amari, "New algorithms for non-negative matrix factorization in applications to blind source separation," in *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, Vol. **5** (2006).
32. M. Berry, M. Browne, A. Langville, V. Pauca, and R. Plemmons, "Algorithms and applications for approximate nonnegative matrix factorization," *Comput. Stat. Data Anal.* **52**(1), 155–173 (2007).
33. H. Kim and H. Park, "Non-negative matrix factorization based on alternating non-negativity constrained least squares and active set method," *SIAM J. Matrix Anal. Appl.* **30**(2), 713–730 (2008).
34. A. Cichocki, R. Zdunek, and S. Amari, "Csiszar's divergences for non-negative matrix factorization: family of new algorithms," *Lect. Notes Comput. Sci.* **3889**, 32–39 (2006).
35. A. Hamza and D. Brady, "Reconstruction of reflectance spectra using robust nonnegative matrix factorization," *IEEE Trans. Signal Process.* **54**(9), 3637–3642 (2006).
36. I. Dhillon and S. Sra, "Generalized nonnegative matrix approximations with Bregman divergences," *Adv. Neural Inf. Process. Syst.* **18**, 283–290 (2006).
37. P. Juzenas, V. Iani, S. Bagdonas, R. Rotomskis, and J. Moan, "Fluorescence spectroscopy of normal mouse skin exposed to 5-aminolaevulinic acid and red light," *J. Photochem. Photobiol., B* **61**(1–2), 78–86 (2001).
38. D. C. G. de Veld, M. Skurichina, M. J. H. Witjes, R. P. W. Duin, D. J. C. M. Sterenborg, W. M. Star, and J. L. N. Roodenburg, "Autofluorescence characteristics of healthy oral mucosa at different anatomical sites," *Lasers Surg. Med.* **32**, 367–376 (2003).
39. G. Weagle, P. E. Paterson, J. Kennedy, and R. Pottier, "The nature of the chromophore responsible for naturally occurring fluorescence in mouse skin," *J. Photochem. Photobiol., B* **2**, 313–320 (1988).
40. K. Onizawa, N. Okamura, H. Saginoya, J. Yusa, T. Yanagawa, and H. Yoshida, "Fluorescence photography as a diagnostic method for oral cancer," *Oral Oncol.* **38**, 343–348 (2002).
41. F. P. Navarro, M. Berger, M. Goutayer, S. Guillermet, V. Josserand, P. Rizo, F. Vinet, and I. Texier, "A novel indocyanine green nanoparticle probe for noninvasive fluorescence imaging *in vivo*," *Progr. Biomed. Opt. Imaging* **10**(30), 1–10 (2009).

Downloaded from SPIE Digital Library on 04 Apr 2011 to 193.48.255.141. Terms of Use: http://spiedl.org/terms