34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

A New Method to Estimate Abundances of Multiple Components using Multi-Spectral Fluorescence Lifetime Imaging Microscopy

O. Gutierrez-Navarro[1], E. R. Arce-Santana[1], Daniel U. Campos-Delgado[1], M. O. Mendez[1] and Javier A. Jo[2].

Abstract— **Multi-Spectral Fluorescent Lifetime Imaging Microscopy (m-FLIM) is a technique that aims to perform non-invasive in situ clinical diagnosis of several diseases. It measures the endogenous fluorescence of molecules, recording their lifetime decay in different wavelength bands. This signal is a mixed response of multiple fluorescent components present in a tissue sample. The goal is to decompose the mixture and estimate the proportional contributions of its constituents. Estimation of such quantitative description will help to characterize the molecular constitution of a given sample.**
This paper presents a new method to estimate the abundances of multiple components present in a mixture measured using m-FLIM data. It provides a closed-form solution under the fully constrained linear unmixing model and assuming the number of components as well as their ideal lifetime decays are known. Its performance is tested using synthetic samples with three components, where performance can be measured accurately and the percentage error is around 6%. The algorithm was also validated performing unmixing of ex vivo data samples from atherosclerotic human tissue containing collagen, elastin and low-density lipoproteins. These experiments were validated against ground-truth maps, which only give a quantitative description, and the estimated accuracy was around 88%.

I. INTRODUCTION

Devices capable of non-invasive in vivo clinical diagnosis are of great interest in biomedical research, mainly to reduce the need for biopsies and other invasive techniques that are time consuming and potentially harmful for the patients. Fluorescence microscopy imaging is widely employed for chemometrical studies and a recent trend in the area is precisely to take advantage of endogenous fluorophores, molecules with a natural fluorescence emission, present in living systems.

The problem is how to identify each component and quantify their abundance given the mixed measurement. We assume that lifetime measurements correspond to a linear mixture, this case is known in literature as linear unmixing [1]. The main difficulty is that multiple solutions exist and usually restrictions are imposed in order to find a best match. Linear Unmixing has been approached using different techniques like optimization and pattern recognition methods. Several algorithms are restricted to solve mixtures of a limited number of components [2]. Only a few proposals deal with mixtures of multiple fluorophores [3], [4], [5]. Most of these methods are not tolerant to highly overlapping spectra components like those emitted by autofluorescence molecules.

We make use of multi-spectral Fluorescence Lifetime Imaging Microscopy (m-FLIM) data, also known as hyper-spectral FLIM [6]. It allows us to better differentiate overlapping spectra like endogenous fluorescent molecules, thus avoiding the use of invasive markers. The m-FLIM data measures the lifetime, the decay rate of the fluorescent luminous response, at each pixel of the sampled area in multiple wavelength bands, see Figure 1. However, accurate and quantitative characterization of in-vivo tissue using m-FLIM hast not been demonstrated yet [7].

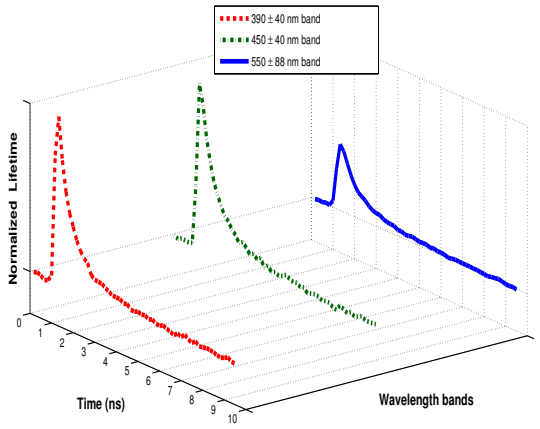

Fig. 1. Multi-spectral Fluorescence Lifetime Imaging Microscopy Data

The Linear Mixture model [1], [8] is employed to state the contribution of each fluorophore in a mixture $y_\lambda(k)$, on every wavelength $\lambda = 1, \cdots, P$ over a window of M measurements as

$$y_\lambda(k) = \sum_{n=1}^{N} \alpha_n p_{n,\lambda}(k) \qquad k = 1, \ldots, M \tag{1}$$

subject to

$$\alpha_n \geq 0 \quad \forall\, n = 1, \ldots, N \tag{2}$$

$$\sum_{n=1}^{N} \alpha_n = 1. \tag{3}$$

[1]Facultad de Ciencias, Universidad Autonoma de San Luis Potosi, SLP, Mexico.
[2] Javier A. Jo is with the Department of Biomedical Engineering, Texas A&M University, College Station, TX, USA.

where k and n denotes the time and component index respectively. The lifetime decay profiles are described by $p_{n,\lambda}(k)$ and α_n stands for their proportional contributions we want to estimate.

II. Constrained Optimization for Spectral Unmixing with a Closed Form Solution

We propose a new closed-form solution for (1) under the sum-to-one (3) and non-negativity (2) constraints. It can be estimated using the ideal lifetime decay profiles $p_{n,\lambda}(k)$ and assuming the number of components is known. To perform multi-spectral Fluorescence Lifetime Imaging Microscopy (m-FLIM) data unmixing the following cost function is employed

$$\min_{\alpha_1,\ldots,\alpha_N} \sum_{\lambda=1}^{P} \|\mathbf{y}_\lambda - \hat{\mathbf{y}}_\lambda\|_2^2 \tag{4}$$

such that

$$\sum_{n=1}^{N} \alpha_n = 1 \qquad \alpha_n \geq 0 \; \forall n \tag{5}$$

where $\|\cdot\|_2$ denotes the Euclidean norm, the estimated response at wavelength λ is given by $\hat{\mathbf{y}}_\lambda = \sum_{n=1}^{N} \alpha_n \mathbf{p}_{n,\lambda}$, for and $\lambda = 1, \ldots, P$. The vector $\mathbf{p}_{n,\lambda} = [p_{n,\lambda}(1) \ldots p_{n,\lambda}(M)]$ contains the lifetime decay samples.

In fact, the solution to (4) can be briefly described as a two step procedure. First the approximation problem is solved incorporating the equality constraint $\sum_n \alpha_n = 1$ using a Lagrange multiplier. Second, if the resulting optimal solution has negative entries those elements are set to zero by introducing new equality constraints and the optimal solution is recalculated according to the Karush-Kuhn-Tucker conditions [9]. The proposed methodology is detailed in Algorithm 1.

The solution obtained is still optimal considering that there is random and uncorrelated uncertainty in the estimation of the measurements vector $\mathbf{y}_\lambda$ at each wavelength λ, i.e.

$$\hat{\mathbf{y}}_\lambda = \sum_{n=1}^{N} \alpha_n \mathbf{p}_{n,\lambda} + \boldsymbol{\xi}_\lambda \tag{9}$$

where $\boldsymbol{\xi}_\lambda \in \mathbb{R}^N$ is a noise vector of uncorrelated elements with zero-mean ($E\{\boldsymbol{\xi}_\lambda\} = \mathbf{0}$) and finite variance $\left(E\{\boldsymbol{\xi}_\lambda^T \boldsymbol{\xi}_\lambda\} < \infty\right)$.

III. Evaluation

The proposed algorithm was tested using two sets of experiments. First, the performance was evaluated using synthetic m-FLIM data stacks, which allows to accurately measure the error on the estimation of the abundances. Later, the algorithm was validated using ex-vivo m-FLIM samples of atherosclerotic human tissue from [11], which were measured in three different excitation wavelength bands (i.e. $P = 3$), namely 390 ± 40 nm, 450 ± 40 nm, and 550 ± 88 nm.

Algorithm 1 Fully Constrained Method for Estimation of Abundances under the Linear Unmixing Model

1) Given the m-FLIM data measurements $\mathbf{y}_\lambda$ and the component profiles $\mathbf{p}_{n,\lambda}$ $n = 1, \ldots, N$ the following elements are calculated:

$$\mathbf{A} = \sum_{\lambda=1}^{P} \begin{bmatrix} \langle \mathbf{p}_{1,\lambda}, \mathbf{p}_{1,\lambda} \rangle & \cdots & \langle \mathbf{p}_{1,\lambda}, \mathbf{p}_{N,\lambda} \rangle \\ \vdots & \ddots & \vdots \\ \langle \mathbf{p}_{N,\lambda}, \mathbf{p}_{1,\lambda} \rangle & \cdots & \langle \mathbf{p}_{N,\lambda}, \mathbf{p}_{N,\lambda} \rangle \end{bmatrix}$$

$$\mathbf{b} = -2 \sum_{\lambda=1}^{P} [\langle \mathbf{y}_\lambda, \mathbf{p}_{1,\lambda} \rangle \ldots \langle \mathbf{y}_\lambda, \mathbf{p}_{N,\lambda} \rangle]^T$$

$$e = \textstyle\sum_{\lambda=1}^{P} \langle \mathbf{y}_\lambda, \mathbf{y}_\lambda \rangle$$

$$\mathbf{c} = [1 \ \ldots \ 1]_{1\times N}^T$$

$$d = 1$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product between vectors. The Gram matrix $\mathbf{A}$ [10] is assumed non-singular, i.e. the set of profiles $(\mathbf{p}_{\lambda,1}, \ldots, \mathbf{p}_{\lambda,N})$ are linearly independent for each wavelength.

2) The solution to the optimal approximation problem subject to constraint (3) is

$$\mathbf{x} = -\frac{1}{2}\mathbf{A}^{-1}\mathbf{b} + \mathbf{A}^{-1}\mathbf{c}\left[\frac{\frac{1}{2}\mathbf{c}^T\mathbf{A}^{-1}\mathbf{b} + d}{\mathbf{c}^T\mathbf{A}^{-1}\mathbf{c}}\right]$$

3) Define the set $\Omega = \{1, \ldots, N\}$, and compute the set of active inequality constraints $\mathcal{I}_\mathbf{x} \subset \Omega$ such that

$$\mathcal{I}_\mathbf{x} = \{i \in \Omega | x_i < 0\}. \tag{6}$$

Let $L = \text{card}(\mathcal{I}_\mathbf{x})$ define the total number of negative terms in the first step of the optimization process. If $\mathcal{I}_\mathbf{x} = \emptyset$, then $\mathbf{x}$ satisfies all the inequality constraints and $\mathbf{x}$ is the solution.

4) If $\mathcal{I}_\mathbf{x} \neq \emptyset$, the equality constraints are augmented by

$$\mathbf{H}_i \mathbf{x} = 0 \quad \forall i \in \mathcal{I}_\mathbf{x} \tag{7}$$

where $\mathbf{H}_i = [0 \ldots 0 \ 1 \ 0 \ldots 0] \in \mathbb{R}^{1\times N}$ is a row vector of zeros with one in the i-th position. The Lagrange multipliers related to the new equality constraints are

$$\delta_i > 0 \ \forall i \in \mathcal{I}_\mathbf{x}$$

In this way, the next optimality sufficient conditions [9], [10] define a set of $N + 1 + L$ linear equations (non-singular by construction) and the same number of unknowns $(\mathbf{x}, \mu, \delta_1, \ldots, \delta_L)$ to solve

$$\begin{aligned} 2\mathbf{A}\mathbf{x} + \mathbf{b} + \mu\mathbf{c} + \sum_{i\in\mathcal{I}_\mathbf{x}} \delta_i \mathbf{H}_i^T &= 0 \\ \mathbf{c}^T\mathbf{x} &= d \\ \mathbf{H}_i\mathbf{x} &= 0 \quad \forall i \in \mathcal{I}_\mathbf{x} \end{aligned} \tag{8}$$

where $\mu > 0$ is the Lagrange multiplier associated to the constraint in (3).

5) If $\mathcal{I}_\mathbf{x} = \Omega$ there is no feasible solution for the optimization problem.

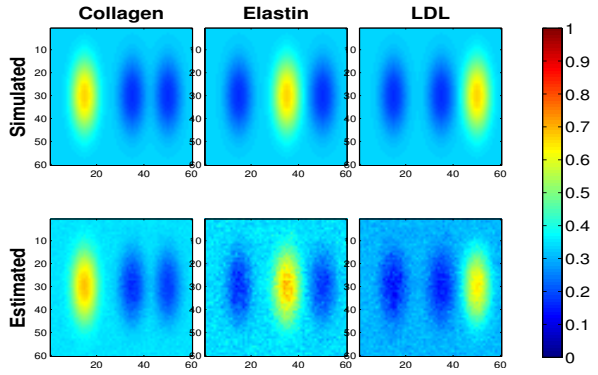

Fig. 2. Simulated and estimated abundances from a synthetic m-FLIM data sample which depicts three fluorophores without co-localized high concentrations.

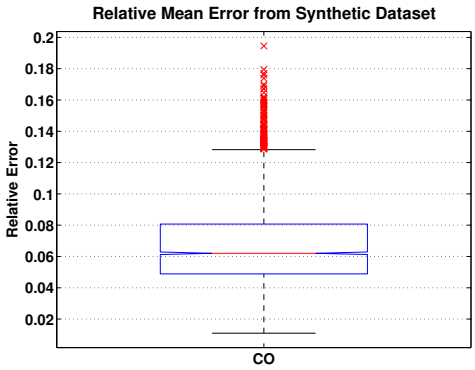

Fig. 3. Boxplot for the relative error estimated on the synthetic m-FLIM sample from Figure 2. The estimated median error was 6%.

Synthetic data

Different synthetic m-FLIM data stacks were generated using a parametric model [12] to synthesize lifetime decay profiles. The experiments simulated mixtures of three autofluorescence molecules: collagen, elastin, and low density lipoproteins (LDL), i.e. $N = 3$. Normal noise with mean 0 and standard deviation $\sigma = 0.02$ was added to the parameters of the synthetic profiles in order to simulate uncertain data. The spatial distributions for molecule concentrations were generated using normal distributions, the pattern simulated without noise can be seen in Figure 2, along with the abundances estimated by the proposed method. The algorithm estimated the component percentages by using the profiles of each component in laboratory samples in concentration of 100%. Please note that the real profiles could be different from the ideal ones due to the protein arrangement in tissue. These profiles $\mathbf{p}_{n,\lambda}$ are composed of samples with dimensions $60 \times 60 \times 510$.

The error on the predictions was measured using the relative error (RE) given by

$$RE = \frac{\|\hat{\mathbf{x}} - \mathbf{x}\|_2}{\|\mathbf{x}\|_2} \tag{10}$$

where $\hat{\mathbf{x}}$ is the estimation of the exact fractional concentrations vector $\mathbf{x}$. A boxplot of the estimated RE is shown in Figure 3, where the median error was 0.062 (6%)

Ex vivo data

To validate our algorithm on ex-vivo data, we used the methodology proposed in [13], where Ground-Truth maps are obtained using a multi-class Fisher's linear discriminant analysis applied to 6 fluorescence features (intensity and average lifetime per wavelength band). Every pixel of the Ground-Truth maps are classified as either "High-Collagen" (Green), "High-Elastin"(Blue) or "High-LDL"(Red). This method provides a good qualitative description of the sample's composition and it has already been validated against histology images [13]. To test the performance of the unmixing algorithm, the estimated concentrations at each position where classified according to the component with the highest value. The classification is simply made as

$$\text{Class} = \arg \max_{i=1,2,3} \alpha_i. \tag{11}$$

where the indexes $i = 1, 2, 3$ represent Collagen, Elastin and LDL, respectively.

IV. RESULTS

The proposed method estimates the abundances of the constituents in mixture samples of m-FLIM data. The performance of the algorithm has been tested using synthetic data. However validation with living tissue is needed. The concentration maps obtained from one ex-vivo sample is shown in Figure 4, here the localized abundances of each component are displayed. The Ground-truth map employed for validation is shown in Figure 5, this map was obtained following the methodology in [13]. The data set depicts a region with high concentrations of LDL and elastin, as can be confirmed with the concentration maps in Figure 4. The percentage of successful classifications was calculated for each possible case using (11), and represented by a confusion matrix in Table I. The numbers in the diagonal represent the percentage of each true positive case. Around 98% of the pixels with high concentrations of Collagen were correctly classified, while 87% of correct classification was made for LDL rich zones. The misclassification of elastin rich zones was low, less than 40%, however these pixels where only detected in the borders by our method, Figure 4, as well as in the Ground-truth map as blue regions in Figure 5. The later are qualitative results, and not exactly accurate. They only provide a good reference since it is impractical to know the real co-localized abundances from living tissue. The total percentage of misclassified pixels ($100\times$ misclassified pixels/total number of pixels by m-FLIM stack) is 12.08% (435 out of 3600).

V. CONCLUSIONS

In this study, we addressed the fully constrained Linear Unmixing problem. The method assumes that the ideal profiles of the components are known. The performance was tested solving noisy synthetic samples, and the proposed

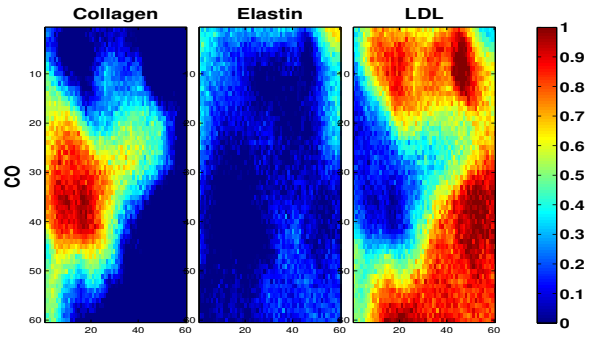

Fig. 4. Estimated Abundances for Ex-Vivo Data Set.

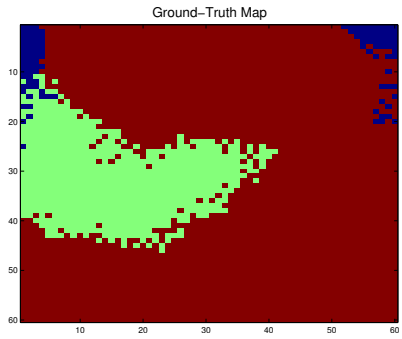

Fig. 5. Ground-Truth map for Ex-Vivo Sample. The sample depicts regions of high concentrations of LDL in red, Collagen and Elastin rich zones are represented by green and blue pixels, respectively.

TABLE I
CONFUSION MATRIX FOR EX-VIVO DATA SET

Real \ Predicted	High Elastin	High Collagen	High LDL
High Elastin	0.39815	0.21296	0.38889
High Collagen	0	0.98317	0.01683
High LDL	0	0.12882	0.87118

method was validated by ground-truth maps from ex-vivo samples. Simulation and experimental results suggest that the proposed methodology provides a good quantitative description of fluorescent components found naturally within tissue. The relative error measurements estimated a percentage error around 6% when performing unmixing on noisy m-FLIM data.

The closed-form solution implies a low computational cost; the abundances of $60 \times 60 \times 510$ m-FLIM data stacks of three components are estimated in less than 2 seconds. Furthermore, the proposed methodology could solve mixtures of multiple fluorophores, however it was only validated in experiments using three components.

There exist plenty of approaches based on constrained optimization to solve spectral unmixing/end-member extraction problems in similar applications, like [14] and [15]. However, to the author's knowledge, there are no similar proposals with a closed-form solution subject to constraints (2) and (3).

The algorithm still needs further validation, and its performance should be evaluated against other methodologies in the state of the art. Since the method is based on a linear mixture model, it could be applied to several different applications, from chemometric studies to remote sensing procedures. An extension is being developed to perform blind end-member extraction from m-FLIM samples, i.e. solving the linear unmixing model without the need for prior information and estimating the lifetime decay profiles of each component, as well as, the abundances.

ACKNOWLEDGEMENT

The authors thank the support of PROMEP through grant F-PROMEP-39/REV-03, SEP-23-005. Omar Gutierrez Navarro also acknowledges the support of CONACYT through grant number 266831.

REFERENCES

[1] J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*. Springer, 2006.

[2] J. G. Gert-Jan Kremers, Erik B. van Munster and J. Theodorus W. J. Gadella, "Quantitative lifetime unmixing of multiexponentially decaying fluorophores using single-frequency fluorescence lifetime imaging microscopy," *Biophys J.*, vol. 95(1), no. Published online 2008 March 21. doi: 10.1529/biophysj.107.125229, 2008 July 1.

[3] A. E. G. K. S. G. M. S. Schlachter, S. Schwedler and C. Kaminski, "A method to unmix multiple fluorophores in microscopy images with minimal a priori information," *Optics Express*, vol. 17, 2009.

[4] H. Xu and . C. D. A. C. . Brad W. Rice. Caliper Life Sciences, Inc., "In-vivo fluorescence imaging with a multivariate curve resolution spectral unmixing technique," *J. Biomed. Opt.*, vol. 14, 064011, 2009.

[5] F. J. Theis, R. Neher, and A. Zeug, "Blind decomposition of spectral imaging microscopy: A study on artificial and real test data," in *Proceedings of the 8th International Conference on Independent Component Analysis and Signal Separation*, ser. ICA '09. Berlin, Heidelberg: Springer-Verlag, 2009, pp. 548–556.

[6] J. McGinty, C. Dunsby, E. Auksorius, R. K. P. Benninger, P. De Beule, D. S. Elson, N. Galletly, D. Grant, O. Hofmann, G. Kennedy, and et al., "Chapter 4 multidimensional fluorescence imaging," *Laboratory Techniques in Biochemistry and Molecular Biology*, vol. 33, p. 133, 2009.

[7] J. Phipps, Y. Sun, N. Hatami, M. C. Fishbein, A. Rajaram, R. Saroufeem, and L. Marcu, "Endoscopic fluorescence lifetime imaging microscopy (flim) images of aortic plaque: an automated classification method," vol. 7548, no. 1. SPIE, 2010, p. 754839.

[8] T. Zimmermann, "Spectral imaging and linear unmixing in light microscopy," *Molecular Biology*, vol. 95, pp. 245– 265, 2005.

[9] J. Nocedal and S. J. Wright, *Numerical Optimization*. Springer, Aug. 2000. [Online]. Available: http://www.worldcat.org/isbn/0387987932

[10] D. Luenberger, *Optimization by vector space methods*, ser. Series in decision and control. Wiley, 1997.

[11] S. Shrestha, B. E. Applegate, J. Park, X. Xiao, P. Pande, and J. A. Jo, "High-speed multispectral fluorescence lifetime imaging implementation for in vivo applications," *Opt. Lett.*, vol. 35, no. 15, pp. 2558–2560, Aug 2010.

[12] B. K. C. Lee, J. Siegel, S. E. D. Webb, L. S. Fort, M. J. Cole, R. Jones, K. Dowling, M. J. Lever, and P. M. W. French, "Application of the Stretched Exponential Function to Fluorescence Lifetime Imaging," *Biophys. J.*, vol. 81, no. 3, pp. 1265–1274, Sep. 2001.

[13] J. Park, P. Pande, S. Shrestha, F. Clubb, B. E. Applegate, and J. A. Jo, "Biochemical characterization of atherosclerotic plaques by endogenous multispectral fluorescence lifetime imaging microscopy," *Atherosclerosis*, no. 0, pp. –, 2011.

[14] D. C. Heinz, S. Member, C. i Chang, and S. Member, "Fully constrained least squares linear mixture analysis for material quantification in hyperspectral imagery," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 39, pp. 529–545, 2001.

[15] D. M. Rogge, B. Rivard, J. Zhang, A. Sanchez, J. Harris, and J. Feng, "Integration of spatialspectral information for the improved extraction of endmembers," *Remote Sensing of Environment*, vol. 110, no. 3, pp. 287–303, 2007.