A SAEM ALGORITHM FOR THE ESTIMATION OF TEMPLATE AND DEFORMATION PARAMETERS IN MEDICAL IMAGE SEQUENCES

FRÉDÉRIC J.P. RICHARD, ADELINE M.M. SAMSON, AND CHARLES A. CUÉNOD.

CORRESPONDING AUTHOR: Frédéric RICHARD.

AUTHOR INFORMATION.

Frédéric RICHARD and Adeline SAMSON
University Paris Descartes,
MAP5, CNRS UMR 8145,
Department of Mathematics,
Tel. : +331-44-553541,
Fax. : +331-42-864144,
(richard,samson)@math-info.univ-paris5.fr

and,

Charles CUÉNOD
University Paris Descartes,
LRI-EA4062
Hospital Georges Pompidou,
Service of Radiology.
Tel.: +331-56-093841,
Fax.: +331-56-093850
ca@cuenod.net



Date: November 15, 2007.

Abstract. This paper is about object deformations observed throughout a sequence of images. We present a statistical framework in which the observed images are defined as noisy realizations of a randomly deformed template image. In this framework, we focus on the problem of the estimation of parameters related to the template and deformations. Our main motivation is the construction of estimation framework and algorithm which can be applied to short sequences of complex and highly-dimensional images. The originality of our approach lies in the representations of the template and deformations, which are defined on a common triangulated domain, adapted to the geometry of the observed images. In this way, we have joint representations of the template and deformations which are compact and parsimonious. Using such representations, we are able to drastically reduce the number of parameters in the model. Besides, we adapt to our framework the Stochastic Approximation EM algorithm combined with a Monte Carlo Markov Chain procedure which was proposed in 2004 by Kuhn and Lavielle. Our implementation of this algorithm takes advantage of some properties which are specific to our framework. More precisely, we use the Markovian properties of deformations to build an efficient simulation strategy based on a Metropolis-Hasting-Within-Gibbs sampler. Finally, we present some experiments on sequences of medical images and synthetic data.

Keywords: SAEM algorithm, medical image sequence, template estimation, image registration, image denoising.

1. Introduction

One of the major new trend in medical imaging is the development of dynamic imaging modalities such as Dynamic Contrast-Enhanced Computed Tomography (DCE-CT) [24] and Magnetic Resonance Imaging (DCE-MRI) [25]. Such modalities usually produce time series of 2D or 3D images of human organs, which are relevant for the diagnosis and the management of cancers [9, 21, 26, 33], ischemic diseases such as stroke [31] and myocardial ischemia [20] and inflammation. The clinical analysis of these image series is however difficult and complex due to several factors: the size and complexity of images, their poor signal-to-noise ratio, and movements of patients during the acquisition.

Such difficulties motivate the development of computer tools which can facilitate the image interpretation by either improving the image quality, correcting the movement of patients, or localizing the different anatomical structures in images. For the construction of such tools, it is relevant to take into account the temporal aspect peculiar to medical image sequences. For that, it is necessary to analyze deformations of organs observed throughout the sequence.

In this paper, we propose a statistical framework for the analysis of deformations in the sequence. In this framework, observed images are viewed as noisy realizations of a randomly deformed template image, deformations being governed by a probability law. Such an approach takes inspiration in the pionneer work of Grenander on pattern theory [17] and is akin to the one used in [18] for characterizing the geometric variability of the brain, or the one used in [1] for the analysis of handwritten

digit images. Such approaches raise some open and challenging issues: how to construct a template which is adapted to data and how to set a deformation probability distribution which suits observations? In this paper, we propose to address these problems by estimating the template and deformation parameters of our statistical model.

In [16], Glasbey and Mardia already set up a statistical framework in which the template is estimated by averaging some images registered onto a common selected image [16]. However, in their approach, the deformation distribution is fixed and not adapted to observations. In [1], Allassonnière *et al* recently defined a hierarchical Bayesian framework related to the one in [3]. In this framework, the template and deformations are both represented as a weighted sum of kernel functions. The two vectors of weights are then assumed randomly Gaussian distributed, the individual deformations being interpreted as non-observed variables. In their paper, these authors proved that the posterior maximum estimator of the parameters (means and covariance matrices) of the two Gaussian distributions exists and is consistent.

They applied their estimation algorithm to the analysis of handwritten digit images. These images are small and binary. Compared to medical images we deal with, they have a very low dimension and a low geometrical complexity. As presented in [1], the model cannot easily be applied to highly dimensional and complex images. The number of parameters associated to the template and deformation representations is too large and makes the estimation problem intractable in practice.

In this paper, our main motivation is to build a statistical framework in which the parameter estimation is feasible on sequences of complex and highly dimensional medical images. One of our main contributions is the construction and the use of some parsimonious representations of the template and deformations. These representations, which are both based on a common finite-element decomposition of the template domain, are adapted to the geometry of the observed images. This adaptation enables to drastically decrease the number of parameters in the model and to reduce the problem complexity.

Although mandatory and critical, the parameter reduction strategy is still not sufficient for dealing with medical sequences. Indeed, in some applications, the number of available observed images may be very low. Hence, the estimation of template and deformations is a difficult statistical problem.

Our model belongs to the class of statistical model with non-observed random variables. In this context, the well-known Expectation-Maximization (EM) algorithm is a powerful tool to estimate parameters by likelihood maximum [12]. However, due to the complexity of our model, the EM algorithm can not be directly applied. In the framework of [1], Allassoniere *et al* first proposed a deterministic version of the EM algorithm, in which the E-step is done using a simple approximation by the mode of the posterior distribution. Due to this approximation, this deterministic EM algorithm is not convergent. However, the use of stochastic versions of the EM algorithm could circumvent this convergence problem. For example, authors used

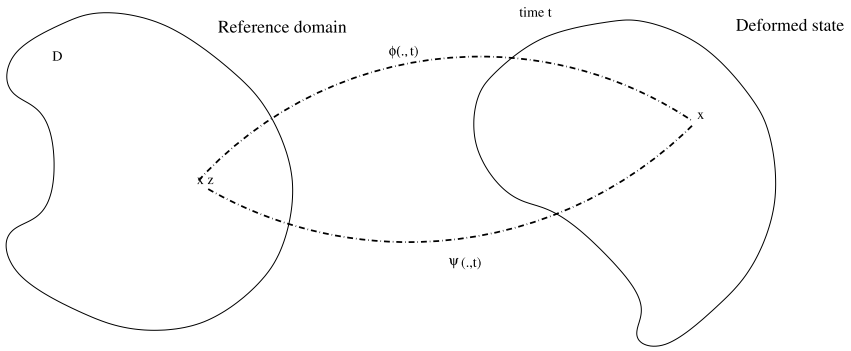


FIGURE 1. The object deformations.

a version of the EM algorithm based on a Stochastic Approximation (SAEM) and combined with a Monte Carlo Markov Chain procedure [2]. The convergence of this algorithm was demonstrated in [22].

Due to the good theoretical and computational properties of SAEM, we also adopt this strategy. Hence, we present a complete construction of a SAEM algorithm which is suitable for our statistical framework. Due to template and deformation representations we use, we take advantage of some Markovian properties of deformations for defining the simulation step of the SAEM algorithm.

The paper is organized as follows. In Section 2, we present the observation statistical model. The estimation stochastic algorithm is described in Section 3 with convergence results. In Section 4, we detail some applications to a medical sequence and to a synthetic sequence simulating real cases with a ground truth.

2. Registration Model

2.1. **A scene model.** A non-rigid object which has undergone some deformations in a time interval $[0, T]$ is observed through a finite series of images. In our application, the observed object is a part of the human body. For describing deformations of the object, let $\mathcal{D}$ denote a common reference domain which is a connected subset of $\mathbb{R}^3$. Then, let us define a mapping ϕ of $\mathcal{D} \times [0, T]$ onto $\mathbb{R}^3$ which gives the positions $\phi(z, t)$ at any time t of any material point located at z in the reference domain (see Figure 1 for illustration). Each map $\phi(\cdot, t)$ is one-to-one and has an inverse denoted by $\psi(\cdot, t)$; the value $\psi(x, t)$ gives the position in the reference domain of a material point which is located in x at time t. In continuum media mechanics, coordinates z of the reference domain are called Lagrangian whereas coordinates x in the deformed domain are called Eulerian. In our context, physical deformations of the object are unknown.

In this paper, we only present a 2D description of our method. Hence, the observations exclusively concern a part of the object which is located in an horizontal plane of $\mathbb{R}^3$. We assume that while it deforms, the part of the object in the plane

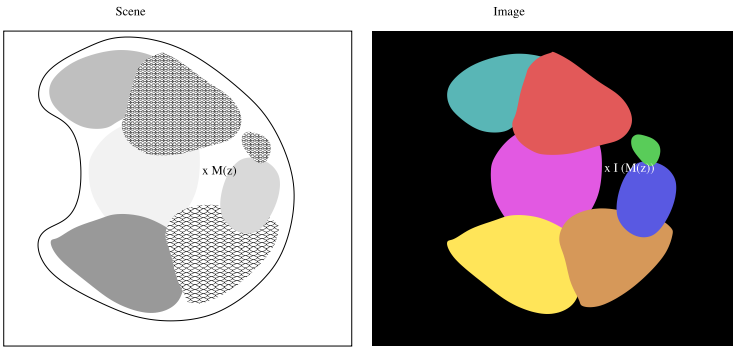


FIGURE 2. The scene representation and the image acquisition.

remains within the plane. In other words, at points of the plane, there is no displacement of the objects in the vertical direction. In the remaining of the text, $\mathcal{D}$ refers to a connected set in the plane of interest. Points of $\mathcal{D}$ are located by their two coordinates on the plane. The application ϕ is identified with its restriction to the plane and has values in $\mathbb{R}^2$.

As illustrated on Figure 2, the object observed throughout the images sequence is made of several tissues which have different material properties. The object is thus non-homogeneous. We define an application $\mathcal{M}$ on $\mathcal{D}$ giving a description of the material properties of the object in the reference configuration. This application is unknown. It is introduced here for the sake of the presentation, but as it will appear later, it will not be used as such.

2.2. **Observation model.** We assume that gray-level values observed in images are related to the physical properties of the object, as illustrated on Figure 2 where each kind of tissues has a specific gray-level value. Due to acquisition noise, the image is however a noisy observation of the physical properties of the object. Let f denote a mapping from $\mathcal{D}$ to $\mathbb{R}$ describing the noise-free image of the object as positioned in the reference domain. Due to the above assumption, f is of the form $f = \mathcal{I}(\mathcal{M})$,where $\mathcal{I}$ represents the action of the image acquisition system which transforms the object properties into a signal. Here, a physical acquisition model could be used for defining precisely the action $\mathcal{I}$. However, this paper focuses exclusively on the image f and not on the physical properties $\mathcal{M}$ the images are related to. So a precise definition of $\mathcal{I}$ is not needed. In the followings, f will be called the template.

The images are acquired on a finite set of n_t times $\{t_1, \cdots, t_{n_t}\}$. Throughout this image sequence, the object is observed as it deforms from a fixed view point. At any fixed observation point x and for any time $t \in \{t_1, \cdots, t_{n_t}\}$, we observe a signal g

$$g(x,t) = \mathcal{I}(\mathcal{M} \circ \psi(\cdot,t))(x) + \varepsilon(x,t), \tag{1}$$

where ε is an additive noise and $\mathcal{I}(\mathcal{M} \circ \psi(\cdot, t))(\cdot)$ is the noise-free image of the deformed object at time t; $\mathcal{I}(\mathcal{M} \circ \psi(\cdot, t))(x)$ reflects the properties of the material point which passes through the position x at time t. We further assume that the noise-free image of the deformed object is equal to the deformed image of the object observed in the reference domain, leading to the expression

$$\mathcal{I}(\mathcal{M} \circ \psi(\cdot, t))(x) = \mathcal{I}(\mathcal{M})(\psi(x, t)) = f(\psi(x, t)). \tag{2}$$

We then have the following observation model

$$\forall\, x \in \mathcal{D}, \forall\, t \in \{t_1, \cdots, t_{n_t}\}, g(x, t) = f(\psi(x, t)) + \varepsilon(x, t). \tag{3}$$

For each t, this observation model is defined on a continuous space of $\mathbb{R}^3$. In practice, images are however observed on a fixed and finite set $\mathcal{G}$ of points of $\mathbb{R}^3$ of dimension n_g (called the image grid),

$$\mathcal{G} = \{x_n, n = 1, \cdots, n_g\}. \tag{4}$$

Let us denote by y_{in} the gray level value $g(x_n, t_i)$ which is observed at position x_n and at time t_i and by ε_{in} the random noise at same position and time. The discrete observation model is then defined as

$$\begin{aligned} y_{in} &= f(\psi(x_n, t_i)) + \varepsilon_{in}, \forall i \in \{1, \cdots, n_t\}, \forall\, n \in \{1, \cdots, n_g\}, \\ \varepsilon_{in} &\sim \mathcal{N}(0, \sigma^2) \end{aligned} \tag{5}$$

where the acquisition noise ε_{in} is assumed to be normally distributed with mean 0 and noise variance σ^2. The template model is detailed in the next section.

2.3. **Template model.** We assume that the domain $\mathcal{D}$ is partitioned into n_e regions where the material properties of the object are homogeneous, *i.e.* where $\mathcal{M}$ is constant. Let p_e denote the regions of the partition and by $\dot{p}_e$ the interiors of these regions. We have

$$\mathcal{D} = \cup_{e=1}^{n_e} p_e \quad \text{and} \quad \forall e_1 \neq e_2, \dot{p}_{e_1} \cap \dot{p}_{e_2} = \emptyset.$$

Due to the above assumptions, the template f is constant on each element of the partition. Let ρ_e be the gray-level value of f on the region p_e. The template can then be written as

$$\forall\, z \in \mathcal{D}, f(z) = \sum_{e=1}^{n_e} \rho_e \mathbb{1}_{z \in p_e}. \tag{6}$$

In practice, a partition of $\mathcal{D}$ is defined by triangulation. The regions p_e are triangles of the plane which are called finite elements and defined by three vertices. Let $\mathcal{N}$ denote the set of all the vertices of the triangulation, n_v the number of vertices, and $\mathcal{N}_e$ the set of vertices defining the finite element p_e. The coordinates of the j^{th} vertex in $\mathcal{D}$ are denoted $\nu_j = (\nu_{j1}, \nu_{j2}) \in \mathbb{R}^2$. Notice that the described template model can be easily extended to 3D by partitioning the domain $\mathcal{D}$ with tridimensional simplices.

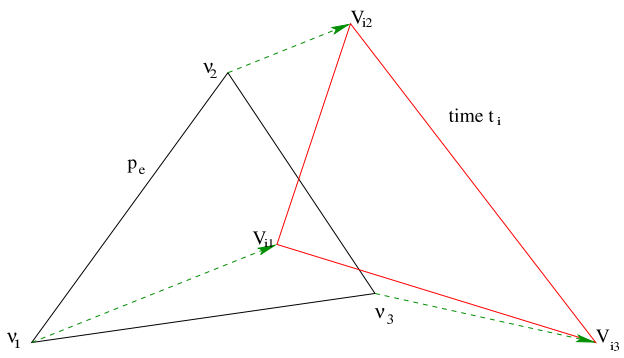


FIGURE 3. Displacements of finite-elements at time t_j.

The template model f is defined in the reference domain whereas the image is observed in the deformed domain. The model of deformations linking these two models is detailed in the next section.

2.4. **Deformation model.** Deformations $\phi(\cdot, t_i)$ are described with piecewise-affine functions: let $\phi(\cdot, t_i) = \mathrm{id} + u(\cdot, t_i)$, where id is the identity map and u is a displacement map. We assume that for all time t_i, displacements $u(\cdot, t_i)$ are affine on each finite elements. Deformations $\phi(\cdot, t_i)$ are thus continuous. In practice, we will also ensure that $\phi(\cdot, t_i)$ are invertible on $\mathcal{D}$ (see Section 3.3.4). Furthermore, the image $\phi(p_e, t_i)$ of a finite element p_e by $\phi(\cdot, t_i)$ remains a triangle (denoted by p_{ei}). As shown on Figure 3, the vertices of the triangle p_{ei} are the images of the vertices of p_e. Let v_{ij} denote the position at time t_i of the jth vertex ν_j of the triangulation. The triangle p_{ei} is entirely defined by the set of vertex positions $\{v_{ij}, j \in \mathcal{N}_e\}$, which is written formally as $p_{ei} = \mathcal{P}(v_{ij}, j \in \mathcal{N}_e)$, where $\mathcal{P}(\cdot)$ denotes the triangle defined by these vertex positions. Moreover, as deformation $\phi(\cdot, t_i)$ at time t_i is piecewise-affine, it is completely characterized by the set of vertex positions $\{v_{ij}, j = 1, \cdots, n_v\}$.

We further introduce a stochastic representation of deformations. At time t_i, the vector $\mathbf{v}_i = (v_{ij})_{j=1}^{n_v}$ of vertex positions is assumed to be a random vector. Vectors $(\mathbf{v}_i)_{i=1}^{n_t}$ represent variations around positions $\nu = (\nu_j)_{j=1}^{n_v}$ of the vertices in the reference domain at different times. The definition of the probability distribution governing these vectors strongly depends on the application at hand. In our application, we assume that vectors $\mathbf{v}_i$ are time-independent and identically distributed, with a multidimensional Gaussian distribution of mean ν and covariance matrix Σ. Note that in some applications, the assumption that vectors $\mathbf{v}_i$ are time-independent may not be realistic. In such a situation, we can extend our approach by describing distributions of $\mathbf{v}_i$ given $\mathbf{v}_{i-1}, \ldots, \mathbf{v}_1$.

In order to reduce the algorithmic complexity, we made some additional assumptions on vertex displacement dependencies. Let us define a neighborhood structure on the vertex set $\mathcal{N}$: two vertices are declared neighbors if and only if they are distinct and if they belong to a same finite element (see Figure 4 for an illustration).

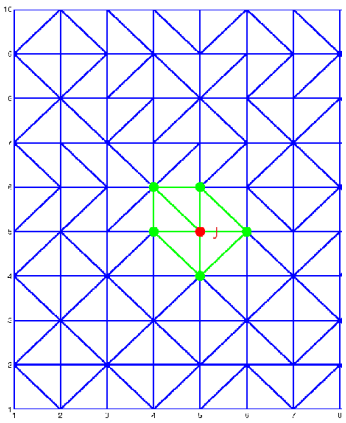


FIGURE 4. The neighborhood structure on the finite-element representation.

Supplied with such a neighborhood structure, the vertex set $\mathcal{N}$ is a graph. On the graph $\mathcal{N}$, we assume that vertex positions are Markovian, *i.e.* the conditional distribution of any vertex position v_{ij} given the other positions only depends on positions in the neighborhood of j:

$$p\left(v_{ij}|v_{ik}, k \neq j, \mathbf{y}; \theta\right) = p\left(v_{ij}|v_{ik}, k \in \mathcal{N}_j, \mathbf{y}; \theta\right). \tag{7}$$

The interest of this assumption will appear in Section 3.2 when describing the estimation algorithm.

2.5. **The complete statistical model.** Due to the template and deformation representations introduced above, the expression of the term $f(\psi(x_n, t_i))$ in Equation (5) can be simplified. Replacing f by its expression given by Equation (6), we first obtain

$$f(\psi(x_n, t_i)) = \sum_{e=1}^{n_e} \rho_e \mathbb{1}_{\psi(x_n,t_i)\in p_e}.$$

But, as the point $\psi(x_n, t_i)$ is in p_e if and only if the point x_n is in $p_{ei} = \phi(p_e, t_i) = \mathcal{P}(v_{ij}, j \in \mathcal{N}_e)$, we can write

$$f(\psi(x_n, t_i)) = \sum_{e=1}^{n_e} \rho_e \mathbb{1}_{x_n\in\mathcal{P}(v_{ij},j\in\mathcal{N}_e)}.$$

In this expression, the inverse deformation ψ is not present anymore and deformations are described in a Lagrangian way using only variables v_{ij}. Such a property is important from a practical point of view, because the computation of the inverse of ϕ is not required.

We can now define the final statistical model:

$$
\begin{aligned}
y_{in} &= \sum_{e=1}^{n_e} \rho_e \mathbb{1}_{x_n \in \mathcal{P}(v_{ij}, j \in \mathcal{N}_e)} + \varepsilon_{in}, \forall i \in \{1, \cdots, n_t\}, \forall\, n \in \{1, \cdots, n_g\}, \\
\varepsilon_{in} &\sim \mathcal{N}(0, \sigma^2), \\
\mathbf{v}_i &\sim \mathcal{N}(\nu, \Sigma),
\end{aligned} \tag{8}
$$

where $\mathbf{y}_i = (y_{in})_{n=1}^{n_g}$ is the vector of observed image gray-level values at time t_i, $\rho = (\rho_e)_{e=1}^{n_e}$ is the vector of the template gray-level values, $\mathbf{v}_i = (v_{ij})_{j=1}^{n_v}$ are the unobserved positions of the vertex at time t_i, randomly distributed in a multidimensional Gaussian distribution of mean ν and variance Σ, ε_{in} represents the image acquisition noise at position x_n at time t_i, with a noise variance σ^2.

Now, our objective is to propose a maximum likelihood estimation of the parameters vector θ from the observations $(y_i)_{i=1}^{n_t}$, where $\theta = (\rho, \sigma^2, \nu, \Sigma)$. This statistical model is a non-linear mixed effect model [27]: unobserved displacements $\mathbf{v}_i$ associated to each image can be interpreted as random effects, and positions ν as fixed effects. Due to the definition of the template, the regression of variables y_{in} on descriptive variables x_n is non-linear.

For the ease of presentation, we now write the model with matrix notations. Let $\mathbf{y} = (y_{in})_{in}$ denote the vector of gray-level values of the n_t images, $\varepsilon = (\varepsilon_{in})_{in}$ the residual errors, and $\mathbf{v} = (v_{ij})_{ij}$ the vector of vertex positions at any time. Let us define $\mathbf{x} = (x_n)_{n=1}^{n_g}$. Let us define for all $e = 1, \cdots, n_e$ the random vertical vector $W_e(\mathbf{v}, \mathbf{x})$ of size $n_t n_g$ which has the random variable $\mathbb{1}_{x_n \in \mathcal{P}(v_{ij}, j \in \mathcal{N}_e)}$ on the $(in)^{th}$ element. The vector $W_e(\mathbf{v}, \mathbf{x})$ is thus a non-linear function of $\mathbf{v}$. Let us further introduced the random matrix

$$
W(\mathbf{v}, \mathbf{x}) = (W_1(\mathbf{v}, \mathbf{x}) | \cdots | W_{n_e}(\mathbf{v}, \mathbf{x}))
$$

defined by horizontal concatenation of vectors $W_e(\mathbf{v}, \mathbf{x})$. Using these notations, the model becomes

$$
\begin{aligned}
\mathbf{y} &= W(\mathbf{v}, \mathbf{x})\rho + \varepsilon, \\
\varepsilon &\sim \mathcal{N}(0, \sigma^2 I), \\
\mathbf{v}_i &\sim \mathcal{N}(\nu, \Sigma), \quad i = 1, \ldots, n_t.
\end{aligned} \tag{9}
$$

Remark: The model is entirely based on the finite-element representation defining both the template and the deformations. The structure of this representation is characterized by a set of vertices and finite-elements. The number of vertices and the definition of finite elements can also be viewed as model parameters. In this paper, these structure parameters are however estimated before computations and set once for all. The neighborhood structure on the vertex set defined in Section 2.4 could also be viewed as a model parameter and be set differently. In this paper, this structure is however fixed.

2.6. **Markovian properties.** In this section, we focus on properties of the random field $\mathbf{v}_i|\mathbf{y}_i$ which at time t_i describes vertex positions given observations. We show that this random field is Markovian for the graph structure $\mathcal{N}$.

By the Hammersley-Clifford theorem [14], a random field is Markovian on a given graph if and only if its distribution has a Gibbsian form on this graph. But, using a Bayes formula, the distribution of $\mathbf{v}_i|\mathbf{y}_i$ can be expressed as:

$$p(\mathbf{v}_i|\mathbf{y}_i;\theta) = \frac{p(\mathbf{y}_i|\mathbf{v}_i;\theta)p(\mathbf{v}_i;\theta)}{\int p(\mathbf{y}_i|\mathbf{v}_i';\theta)p(\mathbf{v}_i';\theta)d\mathbf{v}_i'}, \ \forall\, i = 1, \cdots, n_t.$$

Hence, the random field $\mathbf{v}_i|\mathbf{y}_i$ is Markovian if the conditional distribution and the prior distribution are both Gibbsian for the graph $\mathcal{N}$. The prior distribution is Gibbsian due to the Markovian assumption stated on displacements in Section 2.4. So it suffices to show that the conditional distribution is Gibbsian.

Using the definition of the observation model in Equation (9), we have

$$p(\mathbf{y}_i|\mathbf{v}_i;\theta) \propto \exp\left(-\frac{1}{2\sigma^2}\sum_{n=1}^{n_g}(y_{in} - \sum_{e=1}^{n_e} W_e(\mathbf{v},\mathbf{x})_{in}\rho_e)^2\right).$$

This equation can also be written as

$$p(\mathbf{y}_i|\mathbf{v}_i;\theta) \propto \exp\left(-\frac{1}{2\sigma^2}\sum_{e=1}^{n_e}\sum_{n=1}^{n_g} W_e(\mathbf{v},\mathbf{x})_{in}(y_{in} - \rho_e)^2\right).$$

But, by definition, $W_e(\mathbf{v},\mathbf{x})_{in}$ only depends on v_{ij} for $j \in \mathcal{N}_e$, hence, there exist functions $\tilde{q}$ such that

$$p(\mathbf{y}_i|\mathbf{v}_i;\theta) \propto \exp\left(-\frac{1}{2\sigma^2}\sum_{e=1}^{n_e}\tilde{q}(v_{ij}, j \in \mathcal{N}_e, \mathbf{y}_i;\rho_e)\right).$$

But, the three vertices of a finite element p_e (indexed in the set $\mathcal{N}_e$) are pairwise-neighbors by definition of the graph structure. Hence, potentials $\tilde{q}_j$ are defined on cliques of order 3 and the conditional distribution $p(\mathbf{y}_i|\mathbf{v}_i;\theta)$ has a Gibbsian form. Finally the random field $\mathbf{v}_i|\mathbf{y}_i$ is Markovian on the graph $\mathcal{N}$.

Besides, using Hammersley-Clifford theorem, this conditional probability is proportional to:

$$\begin{aligned}(10)\qquad & p\left(v_{ij}|v_{ik}, k \neq j, \mathbf{y};\theta\right) \propto \\ & \exp\left(-\sum_{k\in\mathcal{N}_j}(v_{ij}-\nu_j)^t\Gamma_{j,k}(v_{ik}-\nu_k) - \frac{1}{2}(v_{ij}-\nu_j)^t\Gamma_{j,j}(v_{ij}-\nu_j)\right) \\ & \times \exp\left(-\frac{1}{2\sigma^2}\sum_{e\in\mathcal{P}_j}\sum_{n=1}^{n_g} W_e(\mathbf{v},\mathbf{x})_{in}(y_{in}-\rho_e)^2\right),\end{aligned}$$

where $\mathcal{P}_j$ denotes the set of finite elements which the vertex j belongs to. This Markovian property is used in the estimation algorithm presented in the next section.

3. Estimation by maximum likelihood

In this section, we present the construction of a SAEM algorithm for the estimation of parameters in the model defined by Equation (8).

3.1. **The statistical problem.** The model (8) is an incomplete data model, $\mathbf{y}$ being the observed data, $\mathbf{v}$ the missing data (or unobserved data) and $(\mathbf{y}, \mathbf{v})$ the complete data. Due to the non-linearity of the regression function in the unobserved data, the likelihood has no closed form. The estimation of θ is therefore a difficult challenge. We consider a stochastic version of the EM algorithm introduced in [12] and dedicated to the parameter estimation of incomplete data model.

3.2. **SAEM algorithm.** The classical EM algorithm is based on an iterative procedure which aims at the maximization of the conditional expectation of the log likelihood $p(\mathbf{y}, \mathbf{v}; \theta)$ of the complete data with respect to parameters θ. At each iteration k, the EM is decomposed into two successive steps called the expectation step (E-step) and the maximization step (M-step):

E-step: the conditional expectation of the complete log-likelihood, given the observations $\mathbf{y}$ and a current approximation $\theta^{(k-1)}$ of parameters is evaluated

$$\forall\, \theta,\ Q^{(k)}(\theta) = \mathbb{E}(\log p(\mathbf{y}, \mathbf{v}; \theta)|\mathbf{y}; \theta^{(k-1)}).$$

M-step: the parameters are updated by

$$\theta^{(k)} = \arg\max_{\theta} Q^{(k)}(\theta).$$

The initial values of Q and θ are arbitrary chosen.

The EM algorithm cannot be directly applied whenever the conditional expectation $Q^{(k)}$ cannot be explicitly computed in the E-step. For dealing with such a situation, Delyon, Lavielle and Moulines propose to use a stochastic approximation of $Q^{(k)}$ based on a sampling of the posterior distribution [11]. In their algorithm (called SAEM), the traditional E-step of the EM algorithm is replaced by two new steps, a simulation step (S-step) and a stochastic approximation step (SA-step):

S-step: the non-observed data $\mathbf{v}^{(k)}$ are sampled from the posterior distribution $p(\mathbf{v}|\mathbf{y}; \theta^{(k-1)})$.

SA-step: a stochastic approximation is done on the complete log-likelihood using the simulated value of the non-observed data

$$Q^{(k)}(\theta) = (1 - \gamma_k)Q^{(k-1)}(\theta) + \gamma_k \log(p(\mathbf{y}, \mathbf{v}^{(k)}; \theta^{(k-1)})),$$

where $(\gamma_k)_k$ is a decreasing sequence of positive step-sizes.

In [11], the convergence of the SAEM algorithm is proved under general conditions when the likelihood $p(\mathbf{y}, \mathbf{v}; \theta)$ belongs to a regular curved exponential family

$$\log p(\mathbf{y}, \mathbf{v}; \theta) = -\Psi(\theta) + \langle S(\mathbf{y}, \mathbf{v}), \Phi(\theta)\rangle, \tag{11}$$

where $\langle ., . \rangle$ is the scalar product and $S(\mathbf{y}, \mathbf{v})$ is the minimal sufficient statistic of the model. For such a model, we have

$$Q^{(k)}(\theta) = -\Psi(\theta) + \langle s^{(k)}, \Phi(\theta) \rangle,$$

where $s^{(k)}$ denotes the conditional expectation $\mathbb{E}(S(\mathbf{y}, \mathbf{v})|\mathbf{y}; \theta^{(k-1)})$ at the k^{th} iteration. Hence, the SA-step of SAEM reduces to the approximation of $s^{(k)}$ and can be simplified as follows.

SA-step (curved exponential model): $s^{(k)}$ is updated by the stochastic approximation scheme

$$s^{(k)} = (1 - \gamma_k) s^{(k-1)} + \gamma_k S(\mathbf{y}, \mathbf{v}^{(k)}). \tag{12}$$

For the same kind of models, the M-step is

M-step (curved exponential model): $\theta^{(k)}$ is updated by

$$\theta^{(k)} = \arg\max \left(-\Psi(\theta) + \langle s^{(k)}, \Phi(\theta) \rangle \right). \tag{13}$$

The S-step of SAEM requires a sampling from the posterior distribution. In many situations, this distribution cannot however be sampled directly. For dealing with such a situation, Kuhn and Lavielle propose to couple SAEM with a Monte Carlo Markov Chain scheme (MCMC) [22]. In their algorithm, the S-step at the k^{th} iteration is based on the construction of a non-homogeneous Markov chain with $p(\mathbf{v}|\mathbf{y}; \theta^{(k-1)})$ as unique stationary distribution. Kuhn and Lavielle prove the convergence of the combined SAEM/MCMC algorithm under general assumptions. The more restrictive assumption is that the non observed variables live in a compact set. Allassonnière *et al* generalize the convergence proof to a non compact setting, which is the case when the prior of the non observed data is a Gaussian distribution [2]. Finally, the main assumptions to be checked to ensure the convergence of the SAEM algorithm coupled with MCMC are the two followings:

A1: For any $\theta \in \Theta$, the MCMC algorithm generates a uniformly ergodic chain for which the invariant probability is $p(\mathbf{v}|\mathbf{y}; \theta)$.
A2: For all k in $\mathbb{N}$, $\gamma_k \in [0, 1]$, $\sum_{k=1}^{\infty} \gamma_k = \infty$ and $\sum_{k=1}^{\infty} \gamma_k^2 < \infty$.

The assumption A2 can be trivially checked by choosing a careful step size sequence. The assumption A1 is more delicate and requires a careful implementation of MCMC as detailed in next section.

3.3. **Application of SAEM to image sequence.** This section details why the version of SAEM adapted to exponential model can be applied to model (8) and the different steps of the algorithm.

3.3.1. *Exponential model family.* The log-likelihood of the complete data $(\mathbf{y}, \mathbf{v})$ in the model described by Equation (9) can be expressed as

$$\begin{aligned} \log p(\mathbf{y}, \mathbf{v}; \theta) &= \log p(\mathbf{y}|\mathbf{v}; \theta) + \log p(\mathbf{v}; \theta), \\ &= -\frac{1}{2\sigma^2}(\mathbf{y} - W(\mathbf{v}, \mathbf{x})\rho)^t(\mathbf{y} - W(\mathbf{v}, \mathbf{x})\rho) \\ &\quad -\frac{1}{2}\sum_{i=1}^{n_t}(\mathbf{v}_i - \nu)^t\Sigma^{-1}(\mathbf{v}_i - \nu) - K(\theta), \end{aligned} \tag{14}$$

where $K(\theta)$ is the logarithm of the normalization constant of the probability distribution $p(\mathbf{y}, \mathbf{v}; \theta)$. This equation can further be written as

$$\log p(\mathbf{y}, \mathbf{v}; \theta) = \sum_{i=1}^{5} S_i(\mathbf{y}, \mathbf{v}), \Phi_i(\theta) - \Psi(\theta),$$

with

$$\begin{array}{lll} S_1(\mathbf{y}, \mathbf{v}) = W(\mathbf{v}, \mathbf{x})^t\mathbf{y}, & S_2(\mathbf{y}, \mathbf{v}) = W(\mathbf{v}, \mathbf{x})^tW(\mathbf{v}, \mathbf{x}), & S_3(\mathbf{y}, \mathbf{v}) = \sum_{i=1}^{n_t}\mathbf{v}_i, \\ S_4(\mathbf{y}, \mathbf{v}) = \sum_{i=1}^{n_t}\mathbf{v}_i\mathbf{v}_i^t, & S_5(\mathbf{y}, \mathbf{v}) = \mathbf{y}^t\mathbf{y}. & \end{array}$$

and

$$\begin{array}{lll} \Phi_1(\theta) = -\sigma^{-2}\rho^t, & \Phi_2(\theta) = \frac{1}{2}\sigma^{-2}\rho^t\rho & \Phi_3(\theta) = -\nu^t\Sigma^{-1}, \\ \Phi_4(\theta) = \frac{1}{2}\Sigma^{-1}, & \Phi_5(\theta) = \frac{1}{2}\sigma^{-2}. & \end{array}$$

Therefore, this model belongs to the curved exponential family defined by Equation (11). Moreover, the sufficient statistics associated to the model and to be updated in the SAEM algorithm are $S_1(\mathbf{y}, \mathbf{v})$, $S_2(\mathbf{y}, \mathbf{v})$, $S_3(\mathbf{y}, \mathbf{v})$ and $S_4(\mathbf{y}, \mathbf{v})$.

3.3.2. *SA-step.* Consequently, when applying SAEM to our model, the SA-step at the k^{th} iteration simply consists of computing

$$s_1^{(k)} = (1 - \gamma_k)s_1^{(k-1)} + \gamma_k W(\mathbf{v}^{(k)t}, \mathbf{x})\mathbf{y}, \tag{15}$$

$$s_2^{(k)} = (1 - \gamma_k)s_2^{(k-1)} + \gamma_k W(\mathbf{v}^{(k)t}, \mathbf{x})W(\mathbf{v}^{(k)}, \mathbf{x}), \tag{16}$$

$$s_3^{(k)} = (1 - \gamma_k)s_3^{(k-1)} + \gamma_k \sum_{i=1}^{n_t}\mathbf{v}_i^{(k)}, \tag{17}$$

$$s_4^{(k)} = (1 - \gamma_k)s_4^{(k-1)} + \gamma_k \sum_{i=1}^{n_t}\mathbf{v}_i^{(k)t}\mathbf{v}_i^{(k)}. \tag{18}$$

3.3.3. *M-step.* In the M-step, parameter θ is estimated as the maximum of $-\Psi(\theta) + \langle s^{(k)}, \Phi(\theta)\rangle$. This maximization is explicit in the case of model (14). A simple

differentiation implies the following expressions for the M-step:

$$\rho^{(k)} = (s_2^{(k)})^{-1} s_1^{(k)}, \tag{19}$$

$$\sigma^{2(k)} = \frac{1}{n_g}(\mathbf{y}^t\mathbf{y} - 2\rho^{(k)t} s_1^{(k)} + \rho^{(k)t} s_2^{(k)} \rho^{(k)}), \tag{20}$$

$$\nu^{(k)} = \frac{1}{n_t} s_3^{(k)}, \tag{21}$$

$$\Sigma^{(k)} = \frac{1}{n_t} s_4^{(k)} - \nu^{(k)t}\nu^{(k)}. \tag{22}$$

Furthermore, the inverse matrix Γ of the matrix Σ is required in the S-step of SAEM. But, the matrix $\Sigma^{(k)}$ is ill-conditioned and the computation of its inverse is numerically unstable. More accurately, if a vertex position v_{ij} is almost fixed throughout the sequence, the diagonal element of Σ corresponding to this vertex is null or very small and Σ is not invertible. Therefore, we decide to sidestep this conditioning problem by substituting $\Sigma^{(k)}$ with $\Sigma^{(k)}+\lambda I_{2n_g}$, where I_{2n_g} is the identity matrix of size $2n_g \times 2n_g$ and λ is a fixed positive value. The small diagonal elements of Σ are therefore artificially replaced by a greater value, avoiding the problem of inverse computation. This is equivalent to the assumption that the distribution of each vertex position v_{ij} has a variance greater or equal to λ. This can also be interpreted as an additional prior distribution on Σ in the complete statistical model, which is a deterministic distribution.

Finally, we obtain an estimator of Γ at iteration k of SAEM by computing

$$\Gamma^{(k)} = (\Sigma^{(k)} + \lambda I_{2n_g})^{-1}. \tag{23}$$

3.3.4. *S-step.* When combining a MCMC algorithm to the SAEM algorithm in the S-step, the choice of the MCMC is a critical issue. This section details the implemented MCMC scheme.

At iteration k of the S-step, vectors $\mathbf{v}_i$ are updated separately, since vertex positions are assumed to be time-independent. The dimension of vector $\mathbf{v}_i$ being very large for standard images, the direct simulation of the whole vector $\mathbf{v}_i^{(k)}$ in the posterior distribution $p(\mathbf{v}_i|\mathbf{y}_i;\theta^{(k-1)})$ is inefficient. Therefore a Gibbs algorithm is absolutely essential to simulate $\mathbf{v}_i^{(k)}$ coordinates by coordinates. For each time t_i, all vertices are therefore successively scanned and for each vertex j, we update the vertex position v_{ij} by simulating in the target distribution

$$\pi_{ij}^{(k)}(v_{ij}) = p\left(v_{ij}|\{v_{il}^{(k-1)}, l<j\}, \{v_{il}^{(k)}, j<l\}, \mathbf{y}_i; \theta^{(k-1)}\right), \tag{24}$$

this expression being detailed in Equation (10). This uni-dimensional conditional distribution being known up to a constant, a Metropolis-Hastings scheme (MH) is coupled with the Gibbs algorithm, in a so-called Metropolis-Hastings-Within-Gibbs (MHWG) algorithm.

In Section 2.6, we have shown that the random field $\mathbf{v}|\mathbf{y}$ is Markovian. Hence, the target distribution depends only on positions of vertices in the neighborhood of

j. This means that the simulation of each v_{ij} involves only a few vertices located in the neighborhood of the vertex j. At iteration k of SAEM, to update the current vector $v_{ij}^{(k-1)}$ during the MHWG algorithm, the MH scheme consists of three steps:

MH1: A candidate position v'_{ij} is sampled from a proposal distribution $q_{ij}^{(k)}$, as described next.

MH2: We compute the so-called acceptance probability

$$\alpha = \min\left(\frac{\pi_{ij}^{(k)}(v'_{ij})}{\pi_{ij}^{(k)}(v_{ij}^{(k-1)})}\frac{q_{ij}^{(k)}(v_{ij}^{(k-1)})}{q_{ij}^{(k)}(v'_{ij})}, 1\right)$$

using Equations (10) and (24).

MH3: We simulate u with an uniform distribution $\mathcal{U}[0,1]$. If $u < \alpha$ then $v_{ij}^{(k)} = v'_{ij}$ else $v_{ij}^{(k)} = v_{ij}^{(k-1)}$.

The definition of the proposal distribution q_{ij} in the step **MH1** is crucial for the speed of convergence of the whole SAEM. We define this distribution carefully so as to ensure that simulated candidates are reasonable. We define the set $c_{ij}^{(k)}$ of points in the interior of the polygon formed by positions $v_{il}^{(k)}$ of neighbor vertices. The new candidate position v'_{ij} is chosen in the set $c_{ij}^{(k)}$, excluding points which would generate mesh singularities (crossings of triangle edges). This choice of the new candidate guarantees that deformations remain one-to-one. The selection of the position v'_{ij} is done randomly using the uniform distribution $q_{ij}^{(k)}$ on non-singular points of the set $c_{ij}^{(k)}$. Notice that this distribution does not depend on the current vertex position $v_{ij}^{(k-1)}$ and that $q_{ij}^{(k)}(v'_{ij}) = q_{ij}^{(k)}(v_{ij}^{(k-1)})$. Hence the acceptance probability reduces to

$$\alpha = \min\left(\frac{\pi_{ij}^{(k)}(v'_{ij})}{\pi_{ij}^{(k)}(v_{ij}^{(k-1)})}, 1\right).$$

Due to the Markovian properties shown in Section 2.6, the computation of the acceptance probability for the jth vertex position only involves the values of vertex positions in the neighborhood of j. This reduces drastically the complexity of the MH algorithm.

Besides, let us notice that this MH algorithm is a random walk scheme. The convergence and ergodicity of this Random-Walk-Within-Gibbs are therefore ensured by the theorem demonstrated in [23, 28].

4. Results

4.1. **Application to a temporal series of medical images.** We apply the SAEM algorithm to a series of ten consecutives computed tomography (CT) images of the upper abdomen (Figure 5). These images were acquired every seconds at the same level of the abdomen with an in-plane resolution of 512x512 pixels. The sequence was acquired after the injection of an iodinated contrast media which enhances

the tissues according to their vascularity. In order to avoid severe image gray-level variations due to the diffusion of the media throught tissues, the sequence acquisition started at 80 seconds after the tissue enhancement had reached a steady state. For this reason, we can assume that the range of gray-level values of a same tissue remains constant with time. On CT images, gray-level ranges of the different tissues, which are measured as the Hounsfied Units (UH), are directly related to their chemical composition (fat, air, bone, water,...) and their content in contrast media. Such a property is consistent with the assumption we made for defining the image acquisition model in Section 2.2. On the images, the liver, the spleen, the aorta, the stomach, a vertebra and some blood vessels are visible. However, images suffer from a poor signal-to-noise ratio, due to the limited irradiation dose used for the sequential acquisition (acquisition parameters were fixed at 80Kv and 50 mAs). Moreover, the observed tissues were deformed during the sequence, due to the patient breathing, heart pulsations, gastric contractions and some volontary movements. These deformations are clearly visible in Figure 5 on the substracted images. Hence, image restoration and motion correction are required for the analysis of these images.

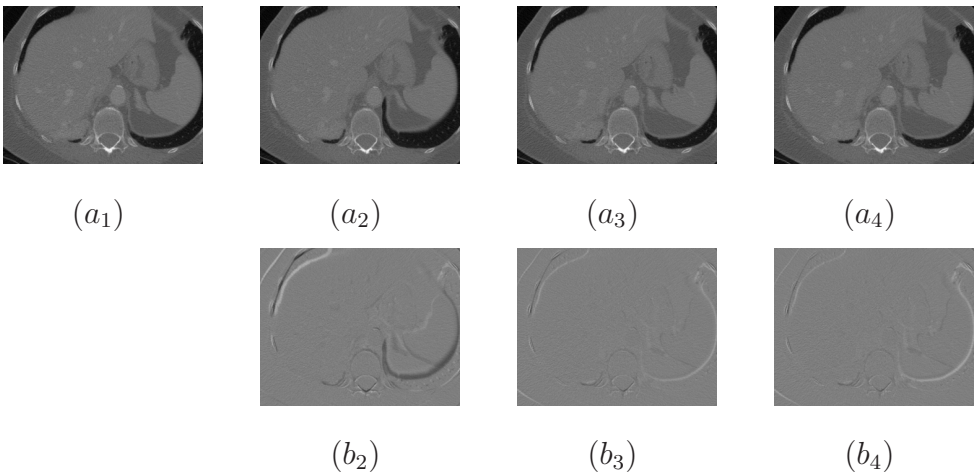


FIGURE 5. Images extracted from a series of CT images of the upper abdomen before estimation. Images (a_1) to (a_4) are taken at different times and the images (b_n) are the differences between (a_n) and (a_1) showing some patient motions.

For the initialization of the template and the finite-element structure, we use an image approximation procedure. We select an image of the sequence and extract borders of its homogeneous regions using a Canny-Deriche edge-detector algorithm [6, 7]. We then select a set of border points which cover all borders and are pairwise distant of at least 2 pixels. We add to this set (1) some equally spaced points located within the homogeneous regions and (2) some points located outside the image which

are used as boundary conditions. Finally, we obtain a triangulation of the image by computing the Delaunay triangulation of the set of all points. On each triangle p_e of this partition, we further initialize the gray-level value ρ_e of the template on p_e by averaging gray-level values of the image on p_e. The initial template and the associated finite-element structure obtained for our sequence are shown on Figures 6 (a) and (b). For this template, the number of finite elements used is around 2400, which is about one third of the image size.

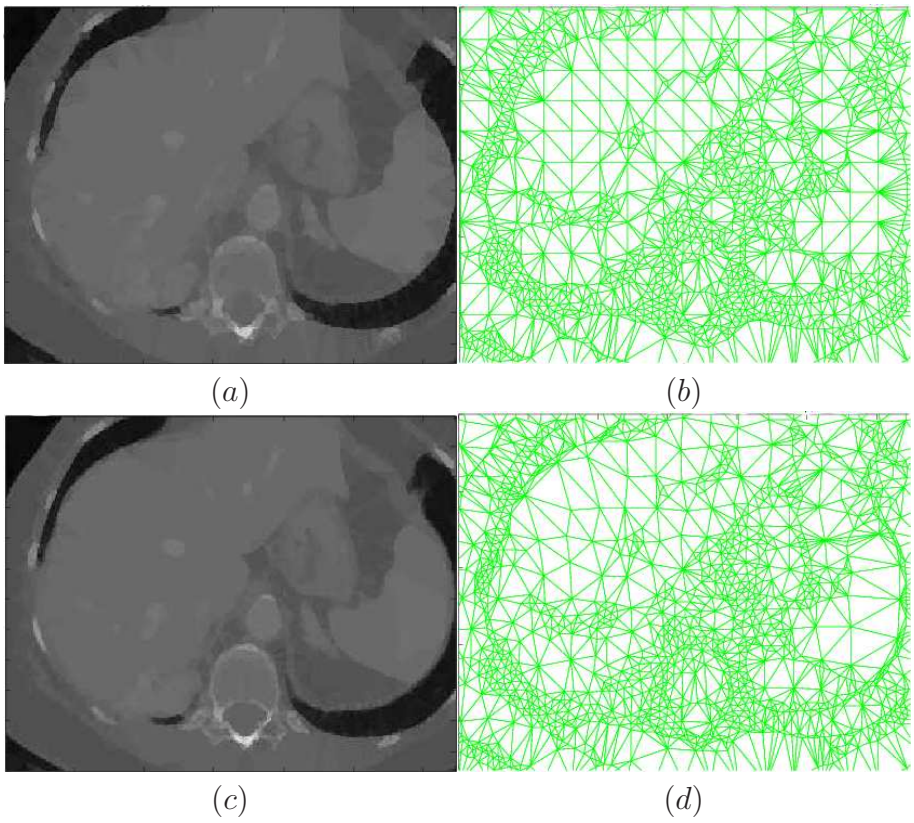
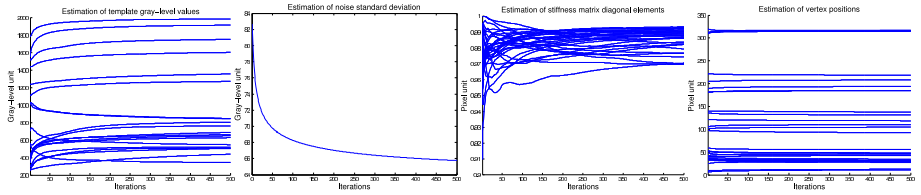


FIGURE 6. Application of the SAEM algorithm to the sequence. (a) initial template gray-level values. (b) initial template mesh. (c) final template gray-level values. (d) final template mesh.

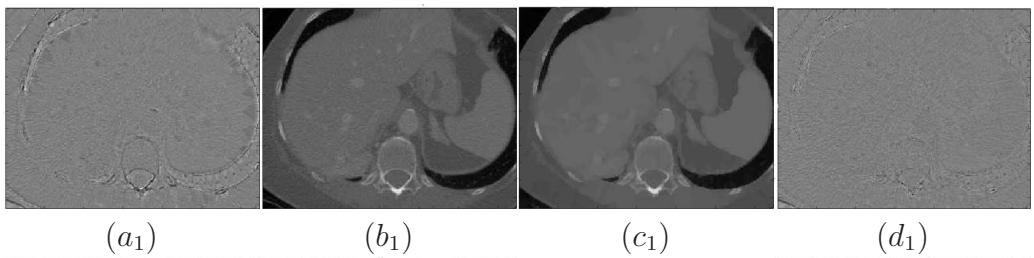
The SAEM algorithm is used to estimate deformation and template parameters from the last ten images of the sequence. Let detail some implementation parameters of the SAEM algorithm. The weight sequence $(\gamma_k)_{k\geq 0}$ used in the SA step has to be chosen to fulfill the condition (A2) (see Section 3.2): we use $\gamma_k = 1$ during the K first iterations $1 \leq k \leq K$, and $\gamma_k = (k-K)^{-1}$ during the last iterations; in our implementation, we set $K = 100$ and have a total of 500 iterations. Five iterations of Metropolis-Hasting algorithm are performed at each S-step of the SAEM algorithm. We implement the SAEM algorithm in Matlab; the code source will be available for

a public use. Figure 7 illustrates the convergence of parameter estimates during the SAEM iterations. SAEM shows a good convergence: most of parameter estimates reach a steady state in about 200 iterations.



(a) (b) (c) (d)

FIGURE 7. Convergence of parameter estimates in the application of SAEM. Parameter estimates are (a) gray-level values of the template, (b) noise standard deviation, (c) stiffness matrix elements, (d) vertex positions.



(a_1) (b_1) (c_1) (d_1)

(a_2) (b_2) (c_2) (d_2)

FIGURE 8. Application of the SAEM algorithm to the sequence. (a_1)-(a_2) difference between the initial template and the observed image. (b_1)-(b_2) observations. (c_1)-(c_2) deformations of the final template at times corresponding to the observed images. (d_1)-(d_2) difference between the deformed template and the observed image. The images (b_1) and (b_2) correspond to images (a_1) and (a_2) of Figure 5, respectively.

On Figures 6 and 8, we illustrate the application of the SAEM algorithm to the sequence. The effect of estimating the template and its deformations can be seen by comparing images of differences between the template and observations before and after estimation (images (a_n) and (d_n)). We observe that the initial differences

are well corrected by the algorithm. After estimation, there only remain some small differences due to either inaccurate approximations of observations by the template (*e.g.* on the spin, middle and top of the image (d_1)) or misregistration of the template and observations. On Figure 8 (a_2)-(d_2), we show a misregistration example: in the part indicated by an arrow, separate regions of tissues have merged, leading to a singular non-injective deformation which is not taken into account in our model. This example is the only significant misregistration we observed.

The quality of the registration can also be quantified by the parameter σ. This noise, which represents the noise of the difference between the template and the images after registration, can not be lower than the image acquisition noise. This acquisition noise can be evaluated by computing the standard deviation of image gray-level values on manually selected homogeneous regions. The obtained mean value is 50 in our images. Before registration, the noise between the initial template and each observation is equal to 104. After registration with the SAEM algorithm, the parameter σ is estimated at the value 66. This means that our method reduces the noise of the difference between two images to a value close to the image acquisition noise. This suggests that using the SAEM algorithm, the template is correctly estimated and accurately registered to observations.

The quality of the template obtained by SAEM is also very good. At initialization, the template has some defaults, mostly located on region borders. After SAEM application, these defaults no longer appear on the template. The noise on the final template is very low. Regions formed by similar tissues have homogeneous gray-level values and regions of different tissues are fairly well-separated. Using HU bounds, we have thresholded the template gray-level values. A classification of different components present in the image is thus obtained. From Figure 9, the physician of our group was able to identify different classes: (a) in brown, the vertebra and ribs, (b) in orange, the blood vessels, (c) in green, the parenchyma (liver and spleen) and muscles (stomach and dorsal muscle), (d) in blue, the fat, and (e) in dark blue, the air within lungs. Notice that some parts of the bones are classified in the blood vessel class. This is due to the normal structure of the bone which has a continuum of intermediate densities varying from the low density of the marrow (in the vertebra center) to the high density of the vertebra wall (in the periphery). Some of these intermediate bone densities have the same gray-level value ranges as the vessels. The classification result we obtained illustrates the quality of the template. Let us also mention that the classification obtained on the template can be propagated to all observed images using the estimated deformations.

4.2. **Evaluation on synthetic data.** For evaluating the precision of the parameter estimation, some experiments are performed on synthetic data. Using the template and parameter values estimated on the medical sequence by SAEM, ten sequences of five observed images are simulated using the following three steps:

(1) simulation of vertex displacements $\mathbf{v}$ with a Gibbs sampler,

(2) application of the simulated displacements to the template for forming a deformed image,
(3) addition of noise to the deformed image.

The SAEM algorithm is then applied to the ten simulated sequences. The relative bias and variance of each parameter estimator are computed and reported on Table 1.

param.	unit	dimension	relative bias (%)	relative std error (%)
ρ	gray-level	1069	0.7	14.9
σ	gray-level	1	7.1	7.1
ν	pixels	1106	2.9	22.2
Σ^{-1} (*)		15180	0	0.03

TABLE 1. Estimation of parameters on synthetic data. (*) For the estimates of Σ^{-1}, bias and variance are in absolute unit.

As we can see, the bias are very low, although the number of samples is low (five images). The relative standard errors are satisfactory. They are of the same order than the ones which are usually obtained using mixed effect models [29].

5. Discussion and conclusion

In Section 2, we present a statistical model describing the observation of deformed objects throughout an image sequence. Similarly to [1, 3, 18], the model defines observed images as noisy realizations of a randomly deformed template. We then construct a SAEM algorithm for the estimation of the template and deformation parameters of the model. The estimation algorithm was finally applied to a medical image sequence and some synthetic data, showing good registration and estimation results. These results encourage us to pursue the application of the algorithm to

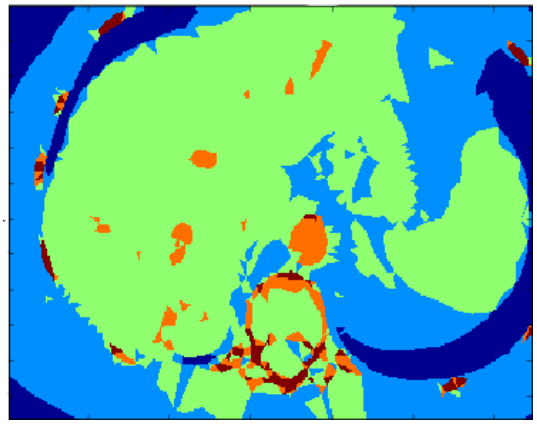


FIGURE 9. Segmentation of the template obtained by a thresholding procedure.

medical image sequences and to extend its implementation in 3D. In collaboration with physicians, we also plan to fully investigate the potential applications of our algorithm: (a) image denoising for the improvement of image quality, (b) image registration for the correction of patient movements, and (c) tissue classification and segmentation of images for the localization of structures of interest.

The model we present departs from the one proposed in [1]. The major difference concerns the way the template and deformations are represented. In [1], the template and deformations are represented as combinations of kernel functions centered around some control points. The sets of control points used for defining the template and the deformations are different, leading to two separate representations. In our model, the template and deformation are defined on a same triangulated domain; the template is defined as a piecewise-constant function on the triangulated domain and deformations as piecewise-affine functions on the same domain. By adapting the domain triangulation to the geometry of objects observed in the sequence, we obtain some joint representations of the template and deformations which are compact and parsimonious. Due to the geometry-adapted representations, the number of template and deformation parameters to be estimated is much lower than the one in [1].

Besides, as mentioned in introduction, we are faced with applications where the number of available observed images is very low and which requires the use of robust estimation techniques. In the literature of Statistics, several algorithms have been proposed for solving maximum likelihood estimation problems associated to hierarchical or mixed models. Those based on the well-known EM algorithm, which takes advantage of the incomplete data model structure of the mixed model [12], are the most adapted to complex data with a low number of images. However, in our context of non-linear mixed models, the application of the EM algorithm is not direct because the E-step has no close form. A solution to this problem is to adopt stochastic versions of the EM algorithm, as proposed by Celeux and Diebolt (Stochastic EM) [8], Wei and Tanner (Monte Carlo EM) [30] and Delyon *et al* (Stochastic Approximation EM) [11]. The SEM and the MCEM algorithms consist of estimating the E-step by the averaging of Monte-Carlo replications, being thus computationally intensive. As an alternative, the SAEM algorithm computes the E-step by a stochastic approximation procedure, requiring only one simulation of the non-observed data. Furthermore, although convergence of the SEM and MCEM are not proved, the almost sure convergence of the SAEM estimates towards a likelihood maximum is proved. Therefore, we only focus on this powerful SAEM algorithm combined with a MCMC scheme to realize the simulation step, as proposed by [22].

In this paper, we propose an efficient SAEM-MCMC algorithm. In such an approach, the crucial issue is the implementation of the MCMC procedure used for the simulation of displacements under the posterior distribution. In our algorithm, the MCMC implementation takes full advantage of a Markovian property of the non-observed displacements, which is specific to our framework. More precisely, we use

a Metropolis-Hastings-Within-Gibbs (MHWG) approach so as to simulate each displacement component successively, the simulation in a one-dimensional distribution being more manageable. The use of the Markovian structure obviously improves the efficacy of this MHWG scheme. In practice, the computation of the acceptance probability of the MH algorithm only requires the positions of the neighboring vertices and the correlations of these neighboring vertex positions. This reduces the complexity and the computational cost of the MCMC scheme.

In our approach, the structure of the domain triangulation used for defining the template and deformations is fixed before the estimation procedure. Hence the precision of the parameter estimation depends on the initial definition of this structure. One way of fixing this limitation would consist of allowing changes of this structure during the estimation procedure. This could be done by adapting the triangulation with some merging and refinement procedures. Such a strategy could be managed by implementing an adaptive version of MCMC algorithms such as the one proposed in [4, 5].

From a last statistical point of view, the issue of parameter identifiability should not be neglected given the complexity of our model. The Markovian assumption on the displacements is therefore crucial to reduce the number parameters used in the estimation algorithm. The risk of data over-fitting obviously exists. However, since our aim is not to build a predictive model but to build a template as adapted as possible to a specific image sequence, this problem is secondary.

References

[1] S. Allassonnière, Y. Amit, and A. Trouvé. Towards a coherent statistical framework for dense deformable template estimation. *J. R. Stat. Soc.: Series B (Statistical Methodology)*, 69(1):3–29, Feb 2007.

[2] S. Allassonnière, E. Kuhn, and A. Trouvé. Bayesian deformable models building via stochastic approximation algorithm: a convergence study. *HAL-CNRS*, 2007.

[3] Y. Amit, U. Grenander, and M. Piccioni. Structural image restoration through deformable templates. *J. Am. Stat. Assoc.*, 86(414):376–387, 1991.

[4] C. Andrieu and E. Moulines. On the ergodicity properties of some adaptive MCMC algorithms. *Ann. Appl. Probab.*, 16:1462–1505, 2006.

[5] Y.F. Atchadé and J.S. Rosenthal. On adaptive Markov Chain Monte Carlo algorithms. *Bernoulli*, 11:815–828, 2005.

[6] J. Canny. A Computational Approach to Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 8(6):679–698, 1986.

[7] R. Deriche. Fast Algorithms for Low-Level Vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 12(1):78–87, 1990.

[8] G. Celeux and J. Diebolt. The SEM algorithm: a probabilistic teacher algorithm derived from the EM algorithm for the mixture problem. *Comput. Stat. Quart.*, 2:73–82, 1985.

[9] C. A. Cuénod, L. Fournier, D. Balvay and J.M. Guinebretiére. Tumor angiogenesis: pathophysiology and implications for contrast-enhanced MRI and CT assessment. *Abdom Imaging*, 31(2):188–193, 2006.

[10] G.E. Christensen, R.D. Rabbitt, and M.I. Miller. Mathematical textbook of deformable neuroanatomies. In *Proc. Natl. Acad. Sci., USA*, volume 90, pages 11944–11948, dec 93.

[11] B. Delyon, M. Lavielle, and E. Moulines. Convergence of a stochastic approximation version of the EM algorithm. *Ann. Statist.*, 27:94–128, 1999.
[12] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. B*, 39:1–38, 1977.
[13] A.A. Ding and H. Wu. Assessing antiviral potency of anti-HIV therapies in vivo by comparing viral decay rates in viral dynamic models. *Biostatistics*, 2(1):13–29, Mar 2001.
[14] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:721–741, 1984.
[15] P. Girard and F. Mentré. A comparison of estimation methods in nonlinear mixed effects models using a blind analysis. *PAGE 14*, Sep 2005.
[16] C.A. Glasbey and K.V. Mardia. A penalized likelihood approach to image warping. *J. R. Stat. Soc. (Series C)*, 63(3):465–514, 2001.
[17] U. Grenander. *General Pattern Theory*. Oxford University Press, 1994.
[18] U. Grenander and M. Miller. Computational anatomy: an emerging discipline. *Quarterly of Applied Mathematics*, 4:617–694, 1998.
[19] J. Guedj, R. Thiébaut, and D. Commenges. Maximum Likelihood Estimation in Dynamical Models of HIV. *Biometrics*, 2007.
[20] R.M. Judd, C.H. Lugo-Olivieri, M. Araj, T. Kondo, *et al.* Physiological basis of myocardial contrast enhancement in fast magnetic resonance images of 2-day-old reperfused canine infarcts. *Circulation*, 92(7):1902–1910, 1995.
[21] C. Kuhl. The current statuts of breast MR imaging. Part I. Choice of technique, image interpretation, diagnostic accuracy, and transfer to clinical practice. *Radiology*, 244(2):356–378, 2007.
[22] E. Kuhn and M. Lavielle. Coupling a stochastic approximation version of EM with a MCMC procedure. *ESAIM Probab. Stat.*, 8:115–131, 2004.
[23] K. L. Mengersen and R. L. Tweedie. Rates of convergence of the Hastings and Metropolis algorithms. *Ann. Statist.*, 24(1):101–121, 1996.
[24] K. A. Miles Perfusion CT for the assessment of tumor vascularity: which protocol? *Br. J. Radiol*, 76(1):36–42, 2003.
[25] J. P. O'Connor, A. Jackson, G.J. Parker and G.C. Jayson. DCE-MRI biomarkers in the clinical evaluation of antiangiogenic and vascular disrupting agents. *Br. J. Cancer*, 96(2):189–195, 2007.
[26] A.R. Padhani. Dynamic contrast-enhanced MRI in clinical oncology: current status and future directions. *J. Magn. Reson. Imaging*, 16(4):407–422, 2002.
[27] J.C. Pinheiro and D.M. Bates. Approximations to the log-likelihood function in the non-linear mixed-effect models. *Journal of Computational Graphics and Statistics*, 4:12–35, 1995.
[28] J.S. Roberts, G.O.and Rosenthal. Harris recurrence of metropolis-within-gibbs and trans-dimensional markov chains. *Ann. Appl. Prob.*, 16:2123–2139, 2006.
[29] A. Samson, M. Lavielle and F. Mentre. The SAEM algorithm for group comparison tests in longitudinal data analysis based on non-linear mixed-effects model. *Statist. Med.*, 26, 2007.
[30] G.C. Wei and M. A. Tanner. Calculating the content and boundary of the highest posterior density region via data augmentation. *Biometrika*, 77:649–652, 1990.
[31] M. Wintermark. Brain perfusion-CT in acute stroke patients. *Eur. Radiol.*, 15(4):D28–31, 2005.
[32] R. Wolfinger. Laplace's approximation for nonlinear mixed models. *Biometrika*, 80(4):791–795, 1993.

[33] M.A. Zahra, K.G. Hollingsworth, E. Sala, D.J. Lomas and L.T. Tan. Dynamic constrast-enhanced MRI as a predictor of tumour response to radiotherapy. *Lancet Oncol.*, 8(1):63–74, 2007.