ABSTRACT

Motion artifacts in image reconstruction problems can be reduced by performing image motion estimation and image reconstruction jointly using a penalized-likelihood cost function. However, updating the motion parameters by conventional gradient-based iterations can be computationally demanding due to the system model required in inverse problems. This paper describes an optimization transfer approach that leads to minimization steps for the motion parameters that have comparable complexity to those needed in image registration problems. This approach can simplify the implementation of motion-compensated image reconstruction (MCIR) methods when the motion parameters are estimated jointly with the reconstructed image.

Keywords: motion-compensated image reconstruction, image registration, tomography

1. INTRODUCTION

Most of the work on motion-compensated image reconstruction has considered the case where the motion parameters are determined separately from the reconstructed image. For example, in PET-CT systems one can estimate the motion from gated CT scans and then apply those parameters to the PET reconstruction process [1–3]. Separate motion information is not always available, so in some applications one must estimate the motion from the same data used for reconstructing the image(s) [4–7]. This paper focuses on such methods for jointly estimating the motion parameters and the reconstructed image from the same data. We describe an optimization transfer approach [8] that simplifies the update of the motion parameters. The principles are applicable to all imaging modalities.

The methods described here are also applicable to super-resolution problems [9–16].

2. THEORY

We first review the joint registration/reconstruction approach described in [4]. We then describe how optimization transfer [8, 17], also known as majorize-minimize methods [18], can simplify the optimization problem.

2.1. Measurement model

Consider an imaging scenario where the data consists of M “scans” \{y_1, \ldots, y_M\}, where \(y_m\) denotes the data associated with the \(m\)th “scan,” i.e., the \(m\)th frame in a dynamic study or the \(m\)th gate in a gated study. Let \(x_m\) denote the (unknown) object corresponding to the \(m\)th scan, for \(m = 1, \ldots, M\). (In the absence of motion we would have \(x_1 = \cdots = x_M\).) We assume that the measurements are related to the object linearly as follows:

\[ y_m = A_m x_m + \varepsilon_m, \quad m = 1, \ldots, M. \]  

where \(A_m\) denotes the system model for the \(m\)th frame and \(\varepsilon_m\) denotes noise. We assume that for each \(m\) the object \(x_m\) and measurement \(y_m\) are motion-free, i.e., the object does not move during the \(m\)th scan (gate or frame).

2.2. Object model

We assume that the object state at each of the \(M\) frames can be written in terms of a common underlying image coefficient vector \(c\) with frame-dependent warp:

\[ x_m = T(\alpha_m) c, \]  

where \(T(\cdot)\) denotes an operator that represents a nonrigid warp with (unknown) motion parameters \(\alpha_m\) associated with the \(m\)th frame. The elements of \(T\) depend on the motion model and type of image interpolator used [7, 19].

2.3. Joint registration/reconstruction

Substituting (2) into (1) yields the following measurement model:

\[ y_m = A_m T(\alpha_m) c + \varepsilon_m, \quad m = 1, \ldots, M. \]
This model has been used widely in MCIR methods. For the model (3), the goal is to jointly estimate the image coefficient $c \triangleq (c_1, \ldots, c_M)$ from the overall measurement vector $y \triangleq (y_1, \ldots, y_M)$. Stacking up these models yields the overall model

$$y = AT(\alpha)c + \epsilon,$$

where $A \triangleq \text{diag}\{A_1, \ldots, A_M\}$, $\epsilon \triangleq (\epsilon_1, \ldots, \epsilon_M)$, and with a slight reuse of notation:

$$T(\alpha) \triangleq \begin{bmatrix} T(\alpha_1) \\ \vdots \\ T(\alpha_M) \end{bmatrix}.$$

One can apply many estimation methods to the model (4). For simplicity, we focus here on penalized weighted least squares (PWLS) estimation [20]:

$$(\hat{c}, \hat{\alpha}) = \arg\min_{c, \alpha} \Psi(c, \alpha)$$

where $W$ is a weight matrix that approximates the inverse of the covariance of $y$, $R_1(c)$ is a spatial regularization term, and $R_2(\alpha)$ is an optional regularization term for the motion parameters [21]. The methods described below are easily generalized from PWLS to penalized-likelihood (PL) estimation by using quadratic surrogate functions for the marginal log-likelihood functions for each $i$ [22].

The cost function (5) is a nonconvex function of $\alpha$ and therefore very challenging to minimize. Some methods have used simultaneous gradient descent, e.g., [6]. The optimization problem is simplified by using alternating minimization [4], where we update $\alpha$ holding $c$ fixed and vice versa, i.e.,

$$\alpha^{n+1} = \arg\min_{\alpha} \Psi(\alpha, c^n)$$

$$c^{n+1} = \arg\min_{c} \Psi(\alpha^{n+1}, c),$$

where we initialize with $\alpha^0$ and $c^0$. Minimizing over $c$ is a standard image reconstruction problem with a modified system matrix $AT(\alpha^{n+1})$. There are many well-known methods for performing this step, e.g., [20]. Minimizing over $\alpha$ in (6) is much more challenging. The motion parameter optimization problem at the $n$th iteration is

$$\alpha^{n+1} = \arg\min_{\alpha} \|y - AT(\alpha)c^n\|_W^2 + R_2(\alpha).$$

This is more challenging than a typical image registration problem because of the presence of the system matrix $A$. The next section proposes an approach to simplifying (8), loosely inspired by [23].

### 2.4. Optimization transfer approach

Consider the simple WLS cost function

$$L(x) = \frac{1}{2} \|y - Ax\|_W^2 = \sum_i \frac{w_i}{2} (y_i - [Ax]_i)^2.$$  

One can show that the following function is a valid quadratic surrogate [22]:

$$Q(x, x^n) \triangleq L(x^n) - \langle x - x^n, g^n \rangle + \frac{1}{2} \langle x - x^n, D(x - x^n) \rangle,$$

where

$$g^n \triangleq -\nabla L(x^n) = A^T \{y - Ax^n\}$$

and

$$D = \text{diag}\{d_j\}, \quad d_j \triangleq \sum_i w_i |a_{ij}| \left( \sum_k |a_{ik}| \right).$$

By completing the square, we can rewrite this quadratic surrogate as follows:

$$Q(x, x^n) \equiv \frac{1}{2} \| D^{1/2} [x - (x^n + D^{-1} g^n)] \|_2^2,$$

ignoring constants independent of $x$. All terms in this form of the quadratic surrogate are in the image domain. We now adopt this idea to the motion estimation problem (8) by equating $x = T(\alpha) c^n$ and $x^n = T(\alpha^n) c^n$. This yields the overall surrogate

$$\phi(\alpha, \alpha^n) \triangleq Q(T(\alpha) c^n, T(\alpha^n) c^n) + R_2(\alpha),$$

where

$$g^n \triangleq -\nabla L(x^n) = A^T \{y - AT(\alpha^n) c^n\}.$$

One can show that this surrogate satisfies the majorization conditions:

$$\phi(\alpha^n, \alpha^n) = \Psi(c^n, c^n)$$

$$\phi(\alpha, \alpha^n) \geq \Psi(c^n, \alpha), \quad \forall \alpha.$$

Therefore, minimizing (or decreasing) $\phi(\alpha, \alpha^n)$ is guaranteed to decrease the original cost function $\Psi(\alpha, c^n)$. With this surrogate, the minimization step for $\alpha$ becomes

$$\alpha^{n+1} = \arg\min_{\alpha} \phi(\alpha, \alpha^n).$$

Because (9) involves only image domain terms, unlike the original function (8), the minimization problem (11) is essentially a standard image registration problem with a least-squares similarity metric. The only difference is that there is a diagonal weighting matrix $D$ in that similarity metric, which is easy to incorporate into image registration software. Of
course, intensity-based image registration problems are non-convex in general, so the minimization problem (11) is non-trivial, but its challenges are well known for image registration.

With this proposed surrogate, the problem of joint estimation (5) simplifies to updating the motion parameters using a cost function that is standard for image registration problems (11), alternating with updates of the image using a standard image reconstruction tool (7). This approach may not be the optimal algorithm in terms of convergence rate (that requires numerical investigation) but it is very appealing from the point of view of software modularization and algorithm maintenance, because it allows one to separately tune the implementations of the “registration” step and the “reconstruction” step in a joint estimation problem.

3. SUMMARY

The proposed optimization transfer approach leads to a minimization step for the warp parameters that is an image domain image registration problem. Existing fast methods such as those based on GPUs can be used to solve this part of the optimization problem [24, 25].

Although we focused on a WLS data fit term, the methods generalize easily to non-quadratic data-fit terms by using quadratic surrogates, e.g., [22].

As long as one implements the reconstruction update step (7) and the motion parameter update step (11) properly, the overall method is guaranteed to monotonically decrease the cost function (5) each iteration. The cost function is non-convex in all intensity-based image registration problems, so it will be important to use standard coarse-to-fine strategies [26] to encourage convergence to a desirable local minimizer.

The next step is to compare this approach to a conventional optimization algorithm applied to the PWLS cost function (5) such as preconditioned nonlinear conjugate gradients.

4. REFERENCES


