OPTIMIZATION TRANSFER APPROACH TO JOINT REGISTRATION / RECONSTRUCTION FOR MOTION-COMPENSATED IMAGE RECONSTRUCTION

Jeffrey A. Fessler

Department of Electrical Engineering and Computer Science
The University of Michigan
1301 Beal Avenue, Ann Arbor, MI 48109-2122

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 alway 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].

This work was supported in part by NIH/NCI grant 1P01 CA87634.

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" $\{\boldsymbol{y}_1, \dots, \boldsymbol{y}_M\}$, where $\boldsymbol{y}_m$ denotes the data associated with the mth "scan," *i.e.*, the mth frame in a dynamic study or the mth gate in a gated study. Let $\boldsymbol{x}_m$ denote the (unknown) object corresponding to the mth scan, for $m = 1, \dots, M$. (In the absence of motion we would have $\boldsymbol{x}_1 = \cdots = \boldsymbol{x}_M$.) We assume that the measurements are related to the object linearly as follows:

$$\boldsymbol{y}_m = \boldsymbol{A}_m \boldsymbol{x}_m + \boldsymbol{\varepsilon}_m, \tag{1}$$

where $\boldsymbol{A}_m$ denotes the system model for the mth frame and $\boldsymbol{\varepsilon}_m$ denotes noise. We assume that for each m the object $\boldsymbol{x}_m$ and measurement $\boldsymbol{y}_m$ are motion-free, *i.e.*, the object does not move during the mth 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 $\boldsymbol{c}$ with frame-dependent warp:

$$\boldsymbol{x}_m = \boldsymbol{T}(\boldsymbol{\alpha}_m)\, \boldsymbol{c}, \tag{2}$$

where $\boldsymbol{T}(\cdot)$ denotes an operator that represents a nonrigid warp with (unknown) motion parameters $\boldsymbol{\alpha}_m$ associated with the mth frame. The elements of $\boldsymbol{T}$ 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:

$$\boldsymbol{y}_m = \boldsymbol{A}_m \boldsymbol{T}(\boldsymbol{\alpha}_m)\, \boldsymbol{c} + \boldsymbol{\varepsilon}_m, \quad m = 1, \dots, M. \tag{3}$$

This model has been used widely in MCIR methods. For the model (3), the goal is to jointly estimate the image coefficient vector $\boldsymbol{c}$ and the motion parameters $\boldsymbol{\alpha} \triangleq (\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_M)$ from the overall measurement vector $\boldsymbol{y} \triangleq (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_M)$. Stacking up these models yields the overall model

$$\boldsymbol{y} = \boldsymbol{A}\,\boldsymbol{T}(\boldsymbol{\alpha})\,\boldsymbol{c} + \boldsymbol{\varepsilon}, \tag{4}$$

where $\boldsymbol{A} \triangleq \mathsf{diag}\{\boldsymbol{A}_1, \ldots, \boldsymbol{A}_M\}$, $\boldsymbol{\varepsilon} \triangleq (\boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{\varepsilon}_M)$. and with a slight reuse of notation:

$$\boldsymbol{T}(\boldsymbol{\alpha}) \triangleq \begin{bmatrix} \boldsymbol{T}(\boldsymbol{\alpha}_1) \\ \vdots \\ \boldsymbol{T}(\boldsymbol{\alpha}_M) \end{bmatrix}.$$

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

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

$$\Psi(\boldsymbol{c}, \boldsymbol{\alpha}) = \|\boldsymbol{y} - \boldsymbol{A}\,\boldsymbol{T}(\boldsymbol{\alpha})\,\boldsymbol{c}\|_{\boldsymbol{W}}^2 + R_1(\boldsymbol{c}) + R_2(\boldsymbol{\alpha}), \tag{5}$$

where $\boldsymbol{W}$ is a weight matrix that approximates the inverse of the covariance of $\boldsymbol{y}$, $R_1(\boldsymbol{c})$ is a spatial regularization term, and $R_2(\boldsymbol{\alpha})$ 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 $\boldsymbol{\alpha}$ 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 $\boldsymbol{\alpha}$ holding $\boldsymbol{c}$ fixed and vice versa, *i.e.*:

$$\boldsymbol{\alpha}^{n+1} = \arg\min_{\boldsymbol{\alpha}} \Psi(\boldsymbol{\alpha}, \boldsymbol{c}^n) \tag{6}$$

$$\boldsymbol{c}^{n+1} = \arg\min_{\boldsymbol{c}} \Psi\big(\boldsymbol{\alpha}^{n+1}, \boldsymbol{c}\big), \tag{7}$$

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

$$\boldsymbol{\alpha}^{n+1} = \arg\min_{\boldsymbol{\alpha}} \|\boldsymbol{y} - \boldsymbol{A}\,\boldsymbol{T}(\boldsymbol{\alpha})\,\boldsymbol{c}^n\|_{\boldsymbol{W}}^2 + R_2(\boldsymbol{\alpha}). \tag{8}$$

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

2.4. Optimization transfer approach

Consider the simple WLS cost function

$$L(\boldsymbol{x}) = \frac{1}{2}\,\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_{\boldsymbol{W}}^2 = \sum_i \frac{w_i}{2}(y_i - [\boldsymbol{A}\boldsymbol{x}]_i)^2.$$

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

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

where

$$\boldsymbol{g}^n \triangleq -\nabla L(\boldsymbol{x}^n) = \boldsymbol{A}'\boldsymbol{W}\,(\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}^n)$$

and

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

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

$$Q(\boldsymbol{x}, \boldsymbol{x}^n) \equiv \frac{1}{2}\left\|\boldsymbol{D}^{1/2}\left[\boldsymbol{x} - \left(\boldsymbol{x}^n + \boldsymbol{D}^{-1}\boldsymbol{g}^n\right)\right]\right\|^2, \tag{9}$$

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

$$\phi(\boldsymbol{\alpha}, \boldsymbol{\alpha}^n) \triangleq Q(\boldsymbol{T}(\boldsymbol{\alpha})\,\boldsymbol{c}^n, \boldsymbol{T}(\boldsymbol{\alpha}^n)\,\boldsymbol{c}^n) + R_2(\boldsymbol{\alpha}), \tag{10}$$

where

$$\boldsymbol{g}^n \triangleq -\nabla L(\boldsymbol{x}^n) = \boldsymbol{A}'\boldsymbol{W}\,(\boldsymbol{y} - \boldsymbol{A}\,\boldsymbol{T}(\boldsymbol{\alpha}^n)\,\boldsymbol{c}^n)\,.$$

One can show that this surrogate satisfies the majorization conditions:

$$\begin{aligned} \phi(\boldsymbol{\alpha}^n, \boldsymbol{\alpha}^n) &= \Psi(\boldsymbol{c}^n, \boldsymbol{\alpha}^n) \\ \phi(\boldsymbol{\alpha}, \boldsymbol{\alpha}^n) &\geq \Psi(\boldsymbol{c}^n, \boldsymbol{\alpha}), \quad \forall \boldsymbol{\alpha}. \end{aligned}$$

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

$$\boldsymbol{\alpha}^{n+1} = \arg\min_{\boldsymbol{\alpha}} \phi(\boldsymbol{\alpha}, \boldsymbol{\alpha}^n). \tag{11}$$

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 $\boldsymbol{D}$ 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

[1] F. Qiao, T. Pan, J. W. Clark, and O. R. Mawlawi. A motion-incorporated reconstruction method for gated PET studies. *Phys. Med. Biol.*, 51(15):3769–84, August 2006.

[2] T. Li, B. Thorndyke, E. Schreibmann, Y. Yang, and L. Xing. Model-based image reconstruction for four-dimensional PET. *Med. Phys.*, 33(5):1288–98, May 2006.

[3] F. Lamare, M. J. Ledesma Carbayo, T. Cresson, G. Kontaxakis, A. Santos, C. Cheze LeRest, A. J. Reader, and D. Visvikis. List-mode-based reconstruction for respiratory motion correction in PET using non-rigid body transformations. *Phys. Med. Biol.*, 52(17):5187–204, September 2007.

[4] M. W. Jacobson and J. A. Fessler. Joint estimation of image and deformation parameters in motion-corrected PET. In *Proc. IEEE Nuc. Sci. Symp. Med. Im. Conf.*, volume 5, pages 3290–4, 2003.

[5] K. Taguchi, M. Zhang, E. C. Frey, J. Xu, W. Paul Segars, and B. M. W. Tsui. Image-domain material decomposition using photon-counting CT. In *Proc. SPIE 6510, Medical Imaging 2007: Phys. Med. Im.*, page 651008, 2007.

[6] F. Odille, P-A. Vuissoz, P-Y. Marie, and J. Felblinger. Generalized reconstruction by inversion of coupled systems (GRICS) applied to free-breathing MRI. *Mag. Res. Med.*, 60(1):146–57, July 2008.

[7] S. Y. Chun and J. A. Fessler. Joint image reconstruction and nonrigid motion estimation with a simple penalty that encourages local invertibility. In *Proc. SPIE 7258, Medical Imaging 2009: Phys. Med. Im.*, page 72580U, 2009.

[8] M. W. Jacobson and J. A. Fessler. An expanded theoretical treatment of iteration-dependent majorize-minimize algorithms. *IEEE Trans. Im. Proc.*, 16(10):2411–22, October 2007.

[9] M. Irani and S. Peleg. Motion analysis for image enhancement: resolution, occlusion, and transparency. *J. Visual Comm. Im. Rep.*, 4(4):324–35, December 1993.

[10] M. Elad and Y. Hel-Or. A fast super-resolution reconstruction algorithm for pure translational motion and common space-invariant blur. *IEEE Trans. Im. Proc.*, 10(8):1187–93, August 2001.

[11] B. C. Tom and A. K. Katsaggelos. Resolution enhancement of monochrome and color video using motion compensation. *IEEE Trans. Im. Proc.*, 10(2):278–87, February 2001.

[12] E. S. Lee and M. G. Kang. Regularized adaptive high-resolution image reconstruction considering inaccurate sub-pixel registration. *IEEE Trans. Im. Proc.*, 12(7):826–837, July 2003.

[13] S. C. Park, M. K. Park, and M. G. Kang. Super-resolution image reconstruction: a technical overview. *IEEE Sig. Proc. Mag.*, 20(3):21–36, May 2003.

[14] S. Farsiu, M. D. Robinson, M. Elad, and P. Milanfar. Fast and robust multiframe super resolution. *IEEE Trans. Im. Proc.*, 13(10):1327–44, October 2004.

[15] S. Farsiu, M. Elad, and P. Milanfar. Multiframe demosaicing and super-resolution of color images. *IEEE Trans. Im. Proc.*, 15(1):141–59, January 2006.

[16] G. K. Chantas, N. P. Galatsanos, and N. A. Woods. Super-resolution based on fast registration and maximum a posteriori reconstruction. *IEEE Trans. Im. Proc.*, 16(7):1821–30, July 2007.

[17] K. Lange, D. R. Hunter, and I. Yang. Optimization transfer using surrogate objective functions. *J. Computational and Graphical Stat.*, 9(1):1–20, March 2000.

[18] D. R. Hunter and K. Lange. A tutorial on MM algorithms. *American Statistician*, 58(1):30–7, February 2004.

[19] M. Jacobson. *Approaches to motion-corrected PET image reconstruction from respiratory gated projection data*. PhD thesis, Univ. of Michigan, Ann Arbor, MI, 48109-2122, Ann Arbor, MI, 2006.

[20] J. A. Fessler. Penalized weighted least-squares image reconstruction for positron emission tomography. *IEEE Trans. Med. Imag.*, 13(2):290–300, June 1994.

[21] S. Y. Chun and J. A. Fessler. A simple regularizer for B-spline nonrigid image registration that encourages local invertibility. *IEEE J. Sel. Top. Sig. Proc.*, 3(1):159–69, February 2009. Special Issue on Digital Image Processing Techniques for Oncology.

[22] H. Erdoğan and J. A. Fessler. Monotonic algorithms for transmission tomography. *IEEE Trans. Med. Imag.*, 18(9):801–14, September 1999.

[23] G. Wang and J. Qi. Iterative nonlinear least squares algorithms for direct reconstruction of parametric images from dynamic PET. In *Proc. IEEE Intl. Symp. Biomed. Imag.*, pages 1031–4, 2008.

[24] R. Strzodka, M. Droske, and M. Rumpf. Image registration by a regularized gradient flow - a streaming implementation in dx9 graphics hardware. *Computing*, 73(4):373–89, November 2004.

[25] P. Muyan-Ozcelik, J. D. Owens, J. Xia, and S. S. Samant. Fast deformable registration on the GPU: A CUDA implementation of demons. In *Intl. Conf. Computational Sciences and its Appl.*, pages 223–33, 2008.

[26] P. Thevenaz and M. Unser. Optimization of mutual information for multiresolution image registration. *IEEE Trans. Im. Proc.*, 9(12):2083–99, December 2000.