

# Bayesian Approach for Data and Image Fusion

Ali Mohammad-Djafari

*Laboratoire des Signaux et Systèmes,  
Unité mixte de recherche 8506 (CNRS-Supélec-UPS)  
Supélec, Plateau de Moulon, 91192 Gif-sur-Yvette, France*

**Abstract.** This paper is a tutorial on Bayesian estimation approach to multi-sensor data and image fusion. First a few examples of simple image fusion problems are presented. Then, the simple case of registered image fusion problem is considered to show the basics of the Bayesian estimation approach and its link to classical data fusion methods such as simple mean or median values, Principal Component Analysis (PCA), Factor Analysis (FA) and Independent Component Analysis (ICA). Then, the case of simultaneous registration and fusion of images is considered. Finally, the problem of fusion of really heterogeneous data such as X-ray radiographic and ultrasound echo-graphic data for computed tomography image reconstruction of 2D or 3D objects are considered. For each of the mentioned data fusion problems, a basic method is presented and illustrated through some simulation results.

## INTRODUCTION

To introduce the basics of the Bayesian approach for data fusion, let start by the simplest problem of data fusion: We have observed a few images (data  $D_i, i = 1, \dots, D_m$ ) of the same unknown object (unknown  $X$ ) and we want to create an image  $\hat{X}$  which represents the fusion of those images.

To apply the Bayesian approach we need first to give a mathematical model relating in some way the data  $D_i$  to the unknowns  $X$  (*Forward model*). This step is crucial for any real application. This mathematical model must be as simple as possible. But, often, the real word problems are too complex to be able to write, with simple mathematical equations, the exact relation between  $D_i$  and  $X$  in a deterministic way. We must also be able to account for the uncertainty associated to this model and the variability of the data measurement system. This is the classical probabilistic modeling of what is called the *likelihood*  $P(D_i|X)$  of parameter  $X$  when the data  $D_i$  is observed. Assigning  $P(D_i|X)$  needs a deterministic mathematical relation between  $D_i$  and  $X$  (Forward model) accounting for physical process of data acquisition and a probabilistic modeling accounting for model uncertainty and what is commonly called the *noise*. Very often, a very simple linear relation plus additive noise give enough satisfaction. From each individual  $P(D_i|X)$ , we can define  $P(D_1, \dots, D_m|X)$  if we can assume that those data have been gathered independently and if there is not any correlation between the different sensors.

The next step is to translate our prior knowledge about  $X$  by assigning to it a prior probability law  $P(X)$ . This step is also crucial particularly when the likelihood model is not too informative (when the likelihood function is not very sharp or when it is not

unimodal.

Assigning the appropriate likelihood function  $P(D_1, \dots, D_m | X)$  and appropriate prior probability law  $P(X)$  is not, in general, an easy task. In this paper, we are not going to discuss this point. We give only simple cases of such models. But, when done, the next step which is to combine these two probability laws through the Bayes rule to obtain the posterior law  $P(X | D_1, \dots, D_m)$  is straightforward. There is only one way to combine  $P(D_1, \dots, D_m | X)$  and  $P(X)$  to obtain  $P(X | D_1, \dots, D_m)$ , *i.e.*, the Bayes rule:

$$P(X | D_1, \dots, D_m) = P(D_1, \dots, D_m | X) P(X) / P(D_1, \dots, D_m) \quad (1)$$

where the denominator is a normalizing factor.

The next step is how to use this posterior law to answer the questions about  $X$ . In fact, from this posterior law we can infer any knowledge on  $X$ . When  $X$  is a scalar variable, we can easily answer to the following questions: What is the value of  $X$  which has highest probability? What is the probability that  $X$  lies between two values  $\underline{X}$  and  $\bar{X}$ ? What is its expected value? What is its variance, its median?, etc. When  $X$  is a vector, not only we can answer all the previous questions about any components  $X_i$  by computing the posterior marginals  $P(X_i | D_1, \dots, D_m)$ , but we can also define the joint conditional laws  $P(X_i, X_j | D_1, \dots, D_m)$  and answer about any question about the *a posteriori* relation between  $X_i$  and other  $X_j$ .

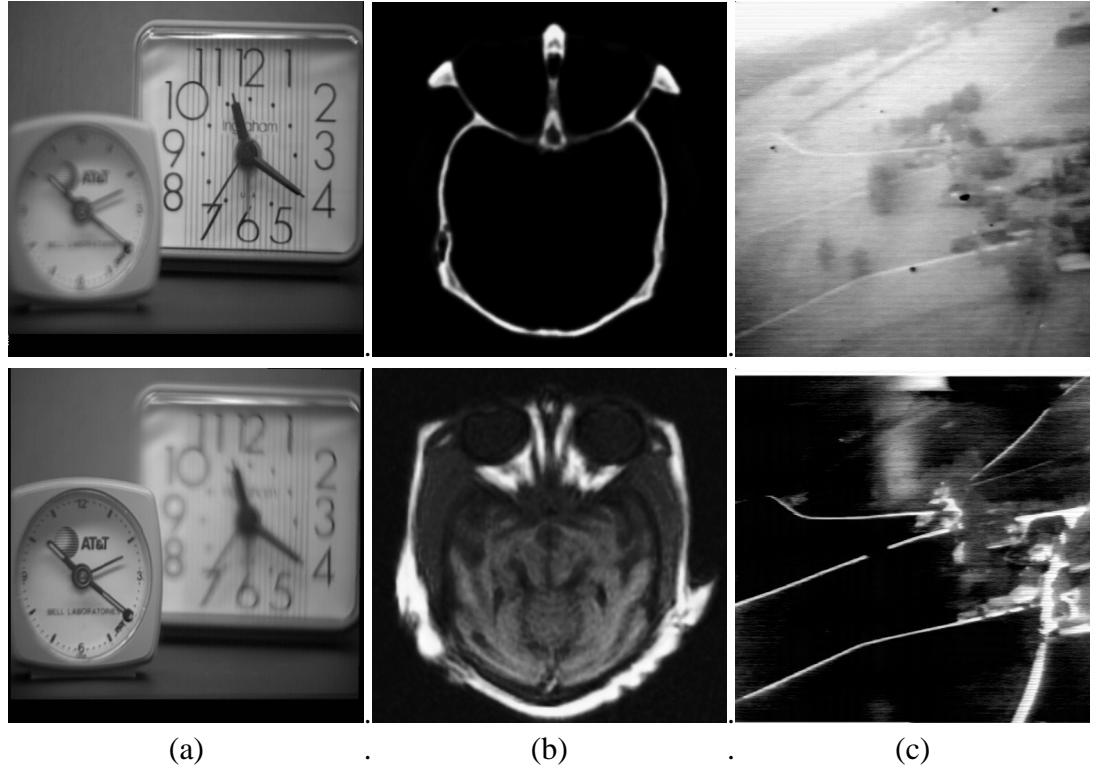
We may also want to use this posterior law to make a decision: Choose one value  $\hat{X}$ , “the best” in some sense. For example, we may define a cost function  $C(X, \hat{X})$  measuring the cost of making an error, *i.e.*, choosing the solution  $\hat{X}$  in place of the true one  $X$  and then want to compute the the solution with lowest posterior mean cost value  $\bar{C}(\hat{X}) = E[C(X, \hat{X})]$ . It is interesting to know that for some particular and natural choices of the cost function  $C(X, \hat{X})$  (those which are increasing functions of the error  $\tilde{X} = (X - \hat{X})$ ), we find classical mode, mean and median values of the posterior law as the best estimators. For example, when  $C(\tilde{X}) = 1 - \delta(\tilde{X})$  we obtain the mode  $\hat{X} = \arg \max_X \{P(X | D_1, \dots, D_m)\}$  and when  $C(\tilde{X}) = \tilde{X}^2$  we obtain the mean  $\hat{X} = E_{P(X | D_1, \dots, D_m)}[X]$  and when  $C(\tilde{X}) = |\tilde{X}|$  we obtain the median  $\hat{X} : P(X > \hat{X} | D_1, \dots, D_m) = P(X < \hat{X} | D_1, \dots, D_m)$  and when  $C(\tilde{X}) = |\tilde{X}|^\beta$  we obtain the  $\beta$ -median.

When  $X$  is a vector, still we can use the decision theory to define an estimator. For example, when  $C(\tilde{X}) = 1 - \delta(\tilde{X})$  we again obtain the mode  $\hat{X} = \arg \max_X \{P(X | D_1, \dots, D_m)\}$  and when  $C(\tilde{X}) = \prod_j (1 - \delta(\tilde{X}_j))$  we obtain the marginal mode  $\hat{X}_j = \arg \max_{X_j} \{P(X_j | D_1, \dots, D_m)\}$  and when  $C(\tilde{X}) = \|\tilde{X}\|^2$  we again obtain the posterior mean  $\hat{X} = E_{P(X | D_1, \dots, D_m)}[X]$ .

In the following, very often, we choose the mode where the corresponding estimator is called the maximum *a posteriori* (MAP). The main objective of this paper is, through simple examples, to show how different practical image fusion problems can be handled easily through the Bayesian approach.

## The simplest model

To give the basics of the Bayesian approach for data fusion, let start by the simplest problem of data fusion: We have observed, in a same geometrical and illumination configuration, a few images  $g_i(x, y)$  of the same unknown object  $f(x, y)$  and we want to create an image  $\hat{f}(x, y)$  which represents the fusion of those images.



**FIGURE 1.** Fusion problem of registered images: a) Two photographic images, b) MRI and PET images in medical imaging, c) Two different radar images

The simplest model for this image fusion problem (when the images have already been registered) is the following:

$$g_i(x, y) = f(x, y) + \epsilon_i(x, y), \quad i = 1, \dots, M, \quad (2)$$

where  $g_i$  are the observed images,  $f$  the original image and  $\epsilon_i$  are the errors or degradations associated to each acquisition. In what follows, we assume all images to be white process meaning that we can work pixel by pixel independently, thus omitting the pixel position  $(x, y)$  from the equations. We consider first this simple model to present the Bayesian approach. Then we extend this simple model to more realistic models accounting for spatial correlation, registration and heterogeneous data.

The Bayesian approach starts with some assumptions on  $\epsilon_i$  which can be translated to probability laws for them, from which we can deduce the conditional probability laws  $p(g_i|f)$ . For example, when we only know the first two moments of  $\epsilon_i$ , the ME principle or any other logical sense leads us to choose Gaussian laws for them. So, assuming  $\epsilon_i$  to

be centered with fixed variances  $\sigma_i^2$  we obtain

$$p(g_i|f) = p_{\epsilon_i}(g_i - f) = \mathcal{N}(0, \sigma_i^2) \propto \exp\left\{-\frac{1}{\sigma_i^2}(g_i - f)^2\right\}. \quad (3)$$

The next step is modeling  $f$  through a prior probability law  $p(f)$ . The third step is to compute the posterior law  $p(f|g_1, \dots, g_M)$  and finally, defining an estimator for  $f$ , for example the Maximum a posteriori (MAP) estimate which is defined as

$$\begin{aligned} \hat{f} &= \arg \max_f \{p(f|g_1, \dots, g_M)\} \\ &= \arg \min_f \{J(f) = -\ln p(f|g_1, \dots, g_M)\}. \end{aligned} \quad (4)$$

Furthermore assuming mutual independence of  $\epsilon_i$ , and accounting for all the pixels, we obtain

$$p(g_1, \dots, g_M|f) = \prod_i p(g_i|f) \propto \exp\left\{-\sum_i \frac{1}{\sigma_i^2}|g_i - f|^2\right\}. \quad (5)$$

Now assuming a generalized Gaussian prior law for  $f$ :

$$p(f|\alpha, \beta, f_0) = \mathcal{GE}(\alpha, \beta, f_0) = \frac{\beta}{2\alpha\Gamma(1/\beta)} \exp\left\{-|(f - f_0)/\alpha|^\beta\right\}, \quad (6)$$

the MAP estimate becomes the optimum of the following criterion:

$$J(f) = \sum_i \frac{1}{\sigma_i^2}|g_i - f|^2 + \frac{\beta}{2\alpha\Gamma(1/\beta)}|(f - f_0)/\alpha|^\beta. \quad (7)$$

It is interesting to see that for  $\beta = 2$  we have  $p(f|\alpha, \beta = 2, f_0) = \mathcal{N}(f_0, \sigma_f^2 = \alpha)$ , then  $J(f)$  becomes a quadratic function of  $f$  and we have an explicit expression for the MAP solution which is obtained by putting  $\frac{dJ}{df} = 0$ :

$$\hat{f} = \frac{1}{\sum_i \lambda_i} \left( \sum_i \lambda_i g_i + \lambda_0 f_0 \right)$$

with  $\lambda_i = \frac{1}{\sigma_i^2}$ ,  $i = 1, \dots, M$  and  $\lambda_0 = \frac{1}{\sigma_f^2}$ . For the particular case of  $\lambda_0 = 0$  (meaning no a priori) and  $\lambda_1 = \dots = \lambda_M = 1$  meaning equal confidence for all the data we obtain

$$\hat{f} = \frac{1}{M} \sum_i g_i$$

which gives the simplest data fusion algorithm: *mean value*.

If, in place of a Gaussian law for the noise, we choose a GE one, we obtain

$$J(f) = \sum_i \lambda_i |g_i - f|^{\beta_i} + \lambda_0 |(f - f_0)|^{\beta_0}. \quad (8)$$

When  $1 < \beta_i < 2$ , the criterion  $J(f)$  is still convex and one can show that the solution is obtained by weighted summation of what is called *soft thresholding* of the inputs  $g_i$  and the *a priori*  $f_0$ .

## A more realistic model

Here we account for the diversity of the scales of the observations

$$g_i(x, y) = h_i f(x, y) + \epsilon_i(x, y), \quad i = 1, \dots, M \quad (9)$$

where  $h_i$  are unknown scalar factors representing the scale of the observations. The problem here is to estimate both  $\mathbf{h} = [h_1, \dots, h_M]$  and  $f$ . Here again, we can use the Bayesian approach to define the posterior law  $p(f, \mathbf{h} | g_1, \dots, g_M)$  and then we have the following options:

- Estimate simultaneously both unknowns  $f$  and  $\mathbf{h}$

$$(\hat{f}, \hat{\mathbf{h}}) = \arg \max_{(f, \mathbf{h})} \{p(f, \mathbf{h} | \mathbf{g})\} \quad (10)$$

One way to do this is to try the following iterative algorithm:

$$\begin{cases} \hat{\mathbf{h}} = \arg \max_{\mathbf{h}} \{p(\hat{f}, \mathbf{h} | \mathbf{g})\} \\ \hat{f} = \arg \max_f \{p(f, \hat{\mathbf{h}} | \mathbf{g})\} \end{cases} \quad (11)$$

but it may or may not converge to the right solution.

- Estimate first  $\mathbf{h}$  through

$$\hat{\mathbf{h}} = \arg \max_{\mathbf{h}} \{p(\mathbf{h} | \mathbf{g})\} \quad (12)$$

and then  $f$  through  $p(f | \hat{\mathbf{h}}, \mathbf{g})$ :

$$\hat{f} = \arg \max_f \{p(f | \hat{\mathbf{h}}, \mathbf{g})\}$$

but this needs

$$p(\mathbf{h} | \mathbf{g}) = \int p(\mathbf{h}, f | \mathbf{g}) df$$

which may not be always easy to do except the Gaussian case. There is however, an iterative algorithm called EM (Expectation-Maximization) which can be used to achieve this estimate without the need for this integration.

- EM algorithm can be summarized as follows:

$$\begin{cases} Q(\mathbf{h}, \hat{\mathbf{h}}) = E \ln p(f, \mathbf{h} | \mathbf{g}) = \int p(f | \mathbf{g}, \hat{\mathbf{h}}) \ln p(f, \mathbf{h} | \mathbf{g}) df \\ \hat{\mathbf{h}} = \arg \max_{\mathbf{h}} \{Q(\mathbf{h}, \hat{\mathbf{h}})\} \end{cases}$$

The main difficulty however stills the computation of  $Q(\mathbf{h}, \hat{\mathbf{h}})$ , but there has been a lot of works to find approximations for it.

- Link with FA, PCA and ICA:

It is interesting to see that, when the prior laws are assumed to be white and Gaussian, the

estimation of  $\mathbf{h}$  becomes equivalent to Factor analysis (FA) and equivalent to Principal Component Analysis (PCA) when neglecting the noises  $\epsilon_i$ . In fact, when  $f$  and  $\epsilon_i$  are assumed Gaussian:

$$f \sim \mathcal{N}(0, \sigma_f^2), \text{ and } \epsilon \sim \mathcal{N}(0, \Sigma_\epsilon)$$

then  $\mathbf{g}$  is Gaussian too and we have:

$$\mathbf{g}|\mathbf{h} \sim \mathcal{N}(\mathbf{0}, \Sigma_g) \text{ with } \Sigma_g = \sigma_f^2 \mathbf{h} \mathbf{h}^t + \Sigma_\epsilon$$

$$p(\mathbf{h}|\mathbf{g}) \propto p(\mathbf{g}|\mathbf{h}) p(\mathbf{h})$$

When  $p(\mathbf{h})$  assumed uniform,  $\sigma_f^2 = 1$  and  $\Sigma_\epsilon = \sigma_\epsilon^2 \mathbf{I}$  then  $\Sigma_g = \mathbf{h} \mathbf{h}^t + \sigma_\epsilon^2 \mathbf{I}$  and the optimization problem  $\hat{\mathbf{h}} = \arg \max_{\mathbf{h}} \{p(\mathbf{h}|\mathbf{g})\}$  reduces to the computation of the eigenvalues of  $\Sigma_g = \mathbf{h} \mathbf{h}^t + \sigma_\epsilon^2 \mathbf{I}$ . Thus the algorithm becomes:

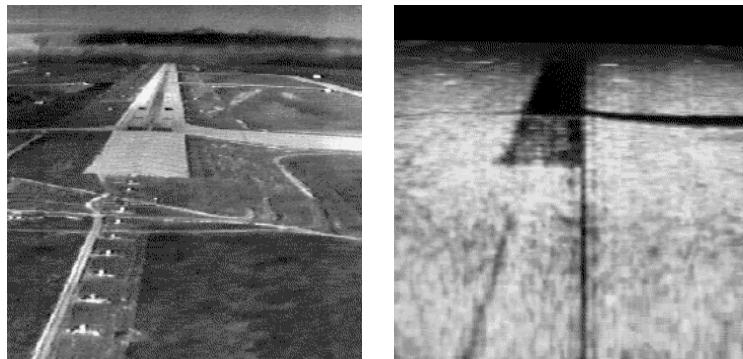
- Estimation of the covariance matrix  $\Sigma_g$  with  $(\Sigma_g)_{i,j} = \langle g_i, g_j \rangle$  from the data and
- Computation of the eigenvalues  $\lambda_i$  of  $\Sigma_g$  and

$$h_i = \begin{cases} \lambda_i - \sigma_\epsilon^2 & \lambda_i > \sigma_\epsilon^2 \\ 0 & \lambda_i < \sigma_\epsilon^2 \end{cases}$$

When the noises  $\epsilon_i$  are neglected but a non Gaussian prior law is chosen for  $f$ , one finds the Independent component analysis (ICA) algorithms [1, 2, 3].

## Registration and fusion of images

In practical image fusion problems, each observed image may have been obtained with different geometrical projection axes and we have to account for this axis transformation. The problem then becomes registration and fusion.



**FIGURE 2.** Fusion problem for unregistered images.

There are two main axis transformation models: global or local.

## A global axis transformation model

When the observed images are not registered, the simplest model is

$$g_i(x, y) = h_i f(x + u(x, y, \boldsymbol{\theta}_i), y + v(x, y, \boldsymbol{\theta}_i)) + \epsilon_i(x, y), \quad i = 1, \dots, M \quad (13)$$

where there are at least two models for the relation between  $(u(x, y, \boldsymbol{\theta}), v(x, y, \boldsymbol{\theta}))$  and  $(x, y)$ :

- A simple affine transformation:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} dx \\ dy \end{bmatrix} + \begin{bmatrix} \cos \theta & \sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} \alpha x \\ \beta y \end{bmatrix}$$

with  $\boldsymbol{\theta} = (dx, dy, \theta, \alpha, \beta)$ .

- A projective model of a 3D scene on a plane:

$$\begin{cases} u = \theta_1 + \theta_3 x + \theta_5 y + \theta_7 x^2 + \theta_8 xy \\ v = \theta_2 + \theta_4 x + \theta_6 y + \theta_7 xy + \theta_8 y^2 \end{cases}$$

which becomes an affine transformation when  $\theta_7 = \theta_8 = 0$ .

In both cases, the only complication, at least theoretically, is the estimation of the registration parameters  $\boldsymbol{\theta}_i$  for each image  $g_i$ <sup>1</sup>.

One can again either estimate all the unknowns jointly by

$$(\hat{f}, \hat{\mathbf{h}}, \hat{\boldsymbol{\theta}}) = \arg \max_{(f, \mathbf{h}, \boldsymbol{\theta})} \{p(f, \mathbf{h}, \boldsymbol{\theta} | \mathbf{g})\}$$

or hierarchically by

$$\begin{cases} \hat{\boldsymbol{\theta}} = \arg \max_{\boldsymbol{\theta}} \{p(\boldsymbol{\theta} | \mathbf{g})\}, \\ \hat{\mathbf{h}} = \arg \max_{\mathbf{h}} \{p(\mathbf{h} | \mathbf{g}, \hat{\boldsymbol{\theta}})\}, \\ \hat{f} = \arg \max_f \{p(f | \hat{\mathbf{h}}, \hat{\boldsymbol{\theta}}, \mathbf{g})\} \end{cases} \quad (14)$$

or still by the following iterative algorithm:

$$\begin{cases} \hat{\boldsymbol{\theta}} = \arg \max_{\boldsymbol{\theta}} \{p(\boldsymbol{\theta} | \mathbf{g}, \hat{\mathbf{h}}, \hat{f})\}, \\ \hat{\mathbf{h}} = \arg \max_{\mathbf{h}} \{p(\mathbf{h} | \mathbf{g}, \hat{\boldsymbol{\theta}}, \hat{f})\}, \\ \hat{f} = \arg \max_f \{p(f | \hat{\mathbf{h}}, \hat{\boldsymbol{\theta}}, \mathbf{g})\} \end{cases} \quad (15)$$

In both case the main difficulty is the estimation of the parameters  $\boldsymbol{\theta}$  due to the nonlinear relation of them to the data. When the noises  $\epsilon_i$  are assumed Gaussian, this estimation

<sup>1</sup> In case where we have only two data sets, we only need to estimate one  $\boldsymbol{\theta}$ , because one of the images can be taken as the reference image. In the following we consider this case.

step in the last algorithm becomes equivalent to least squares (LS). However, the LS criterion is still non quadratic in  $\theta$  and its optimization needs global techniques. There have been many publications on the subject. We mention here the stochastic methods such as MCMC where the optimizations are replaced by sampling and other deterministic approaches such as multi-grid, multi-resolution and pyramidal optimization techniques.

### *A local axis transformation model*

When the 3D scene is a compact body, the previous model can be valid (one value for  $\theta$  for the whole image), but for more realistic 3D scenes, we have to consider a local variation for the parameters, *i.e.*,  $\theta(x, y)$ . The estimation of  $\theta(x, y)$  is then a difficult problem. Many investigations have been performed on this subject: optical flow techniques [] where  $\theta(x, y)$  is assumed to be a very smooth function and estimated via a variational method or the motion estimation techniques where  $\theta(x, y)$  is assumed to be constant inside a window around the point  $(x, y)$ . More detailed investigation of these methods is not within the focus of this paper.

## **3D image recovery from a set of 2D data**

One example is a 3D shape recovery of a compact object from a set of its shadows [4, 5] or from a set of its pictures [6, 7, 8], or still from a set of its X-ray radiographic data at different view angles.

An object is called compact if it can be represented by a function  $f(\vec{r})$  which is equal to one inside a region and zero outside that region:

$$f(\vec{r}) = \begin{cases} 1 & \text{if } \vec{r} \in \mathcal{C} \\ 0 & \text{elsewhere} \end{cases} \quad (16)$$

where  $\vec{r}$  is a point in the coordinate space of the object  $(\xi, \eta, \zeta)$ .

In all these examples we can write:

$$g_i(x, y) = [\mathcal{H}_i f(\vec{r})](x, y) + \epsilon_i(x, y), \quad i = 1, \dots, M \quad (17)$$

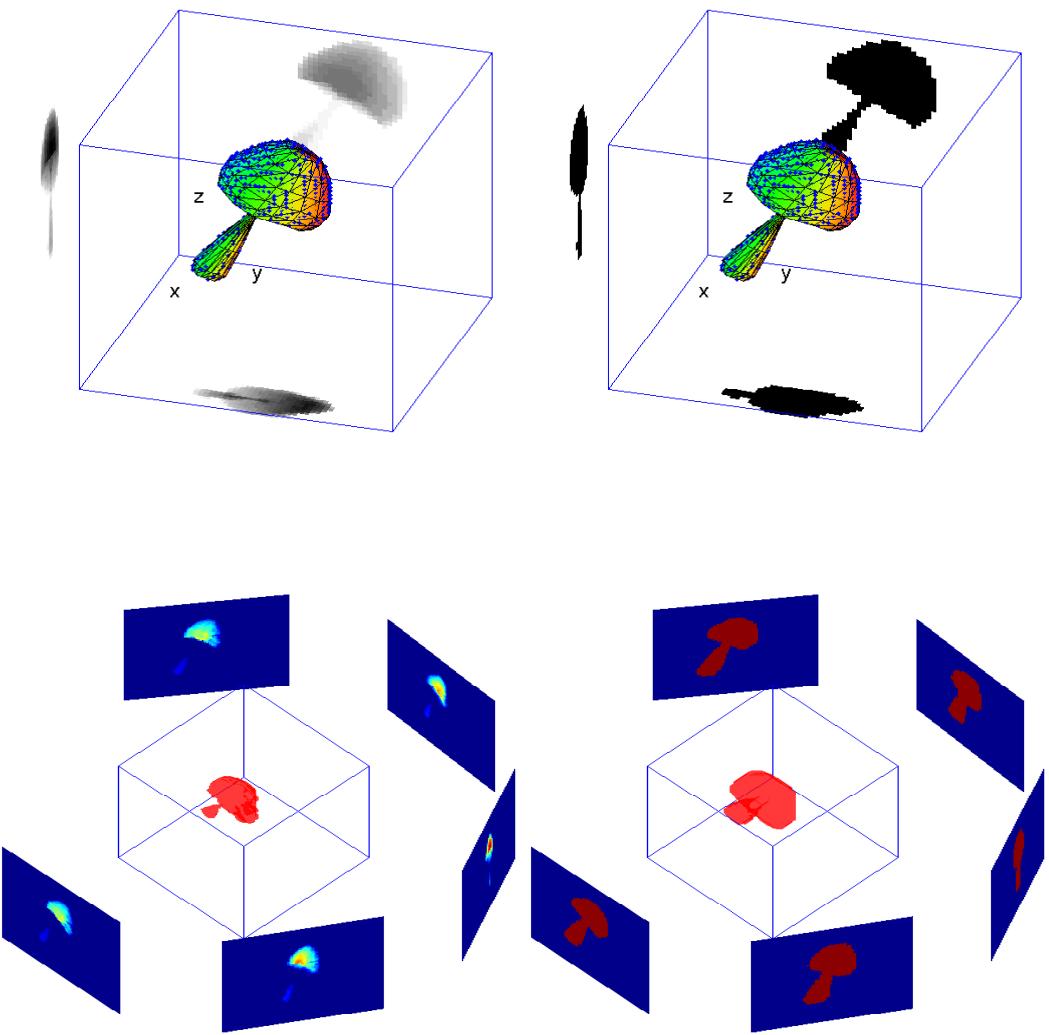
where  $\mathcal{H}_i$  are operators corresponding to the mathematical relation between the data set  $g_i$  and a particular feature of the object  $f(\vec{r})$ .

The example of computed tomography (CT) is interesting, because in this application  $f(\vec{r})$  is related to the volumetric material density of the body and  $\mathcal{H}_i$  are linear operators (related to the Radon transform). The discretized version of the above relation then becomes

$$\mathbf{g}_i = \mathbf{H}_i \mathbf{f} + \boldsymbol{\epsilon}_i, \quad i = 1, \dots, M \quad (18)$$

where  $f_j$  is the voxel  $j$  of the volume and  $\mathbf{H}_i$  are matrices corresponding to the line integrals. It is then easy to estimate  $f$  from the line integration data  $\mathbf{g}_i$ .

The shape from shadow problem is very similar. In fact, a shadow can be considered as the support of a X-ray radiographic data for a compact object. However, the operators



**FIGURE 3.** Shape recovery from X-ray radiographic data in computed tomography (left) or from shadows (right). The upper row correspond to a situation where we have three orthogonal views and the lower row corresponds to a situation where we have view around the object.

$\mathcal{H}_i$  are no more linear. One have to model the contour (external surface) of the body by a parametric function and estimate these parameters from the data.

The shape from pictures problem is also very similar, but here the pictures are related not only to the volume or the surface of the body, but also to the optical properties of the body surface.

## DATA FUSION IN COMPUTED TOMOGRAPHY

A more general relation when facing heterogeneous data is

$$g_i = [\mathcal{H}_i f_i] + \epsilon_i, \quad i = 1, \dots, M \quad (19)$$

where  $\mathcal{H}_i$  are operators corresponding to the mathematical relation between the data set  $g_i$  and particular features  $f_i$  of the object. When discretized, we have:

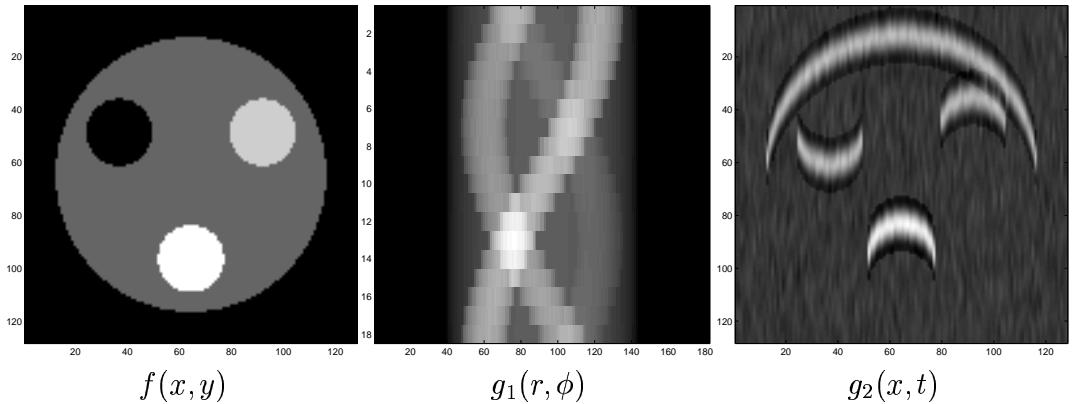
$$\mathbf{g}_i = \mathbf{H}_i \mathbf{f}_i + \boldsymbol{\epsilon}_i, \quad i = 1, \dots, M \quad (20)$$

where  $\mathbf{H}_i$  are matrices corresponding to discretized versions of the operators  $\mathcal{H}_i$ .

When we have homogeneous data, *i.e.*, when all data are related to the same feature  $f$  of the body, we find the problems of shape recovery from X-rays, shadows or pictures of the last section. But, when the data are of different nature, for example radiographic data  $g_1$  and ultrasound echo-graphic data  $g_2$ , then we have

$$\begin{cases} g_1(r, \phi) = \int_{L(r, \phi)} f(x, y) \, dl + \epsilon_1(r, \phi) \\ g_2(x, t) = \int q(x, y) h(y - t) \, dy + \epsilon_2(x, t) \end{cases} \quad (21)$$

where  $f(x, y)$  is volume material density of the body and  $q(x, y)$  is the reflection coefficient distribution of the body which is more related to the variation of the density  $f(x, y)$ .



**FIGURE 4.** X-ray radiographic and ultrasound data fusion in computed tomography

In this case then we have:

$$\begin{cases} g_1 = \mathbf{H}_1 \mathbf{f} + \boldsymbol{\epsilon}_1 \\ g_2 = \mathbf{H}_2 \mathbf{q} + \boldsymbol{\epsilon}_2 \end{cases} \quad (22)$$

The main difficulty here is modeling the relation between  $\mathbf{f}$  and  $\mathbf{q}$ . When this is done, we can estimate the unknowns  $(\mathbf{f}, \mathbf{q})$  either through a MAP criterion:

$$(\hat{\mathbf{f}}, \hat{\mathbf{q}}) = \arg \max_{(\mathbf{f}, \mathbf{q})} \{p(\mathbf{f}, \mathbf{q} | \mathbf{g}_1, \mathbf{g}_2)\} \quad (23)$$

or through the Generalized maximum likelihood (GML) criterion:

$$\begin{cases} \hat{\mathbf{q}} = \arg \max_{\mathbf{q}} \{p(\mathbf{q} | \mathbf{g}_1, \mathbf{g}_2)\}, \\ \hat{\mathbf{f}} = \arg \max_{\mathbf{f}} \{p(\mathbf{f} | \hat{\mathbf{q}}, \mathbf{g}_1, \mathbf{g}_2)\} \end{cases} \quad (24)$$

where

$$p(\mathbf{f}, \mathbf{q} | \mathbf{g}_1, \mathbf{g}_2) \propto p(\mathbf{g}_1 | \mathbf{f}) p(\mathbf{g}_2 | \mathbf{q}) p(\mathbf{f}, \mathbf{q})$$

and

$$p(\mathbf{q} | \mathbf{g}_1, \mathbf{g}_2) = \int p(\mathbf{f}, \mathbf{q} | \mathbf{g}_1, \mathbf{g}_2) d\mathbf{f}.$$

The main difficulty here, as we mentioned before, is modeling the relation between  $\mathbf{f}$  and  $\mathbf{q}$  through  $p(\mathbf{f}, \mathbf{q}) = p(\mathbf{f} | \mathbf{q}) p(\mathbf{q}) = p(\mathbf{q} | \mathbf{f}) p(\mathbf{f})$ . The exact physical relation between  $\mathbf{f}$  and  $\mathbf{q}$  is not easy to establish from the physics of the problem. The only qualitative relation is that  $\mathbf{f}$  is related to the distribution of the material density inside the body, while,  $\mathbf{q}$  representing the ultrasound reflection coefficient distribution is related to the change of material density, *i.e.*,  $q_j$  is almost equal to zero if the pixel  $j$  is inside a homogeneous area and its value is an increasing function of the amount of that change. Indeed, we have only reflection when going from a homogeneous region to another homogeneous region with higher material density. This qualitative relation between  $\mathbf{q}$  and  $\mathbf{x}$  can be modeled through  $p(\mathbf{f}, \mathbf{q})$  or  $p(\mathbf{f} | \mathbf{q})$  or  $p(\mathbf{q} | \mathbf{f})$  using compound Markov modeling. For example, we can choose

$$p(\mathbf{q}) = \exp\{-\lambda_1 \|\mathbf{q}\|^{\beta_1}\} = \exp\{-\lambda_1 \sum_j |q_j|^{\beta_1}\} \quad (25)$$

to account for the fact that the distribution of  $\mathbf{q}$  is very concentrated around the zero, and

$$p(\mathbf{f} | \mathbf{q}) = \exp\{-\lambda_1 \sum_j (1 - q_j) \phi(f_j - \bar{f}_j)\} \quad (26)$$

where  $\bar{f}_j$  is the mean value of pixels in the neighborhood of the pixel  $j$  and  $\phi(t) = \{|t|^\beta, \sqrt{1+|t|^2}\}$  or any other convex function, to account for the fact that  $q_j = 0$  means we are in a homogeneous region and  $q_j \neq 0$  means that we are on the borders of the two regions.

We can also do in reverse, trying to model

$$p(\mathbf{q} | \mathbf{f}) = \exp\{-\lambda_1 \sum_j |q_j - \phi(f_j - \bar{f}_j)|^\beta\} \quad (27)$$

where  $\phi$  is a monotonic increasing function and  $1 < \beta < 2$  meaning that, when  $(f_j - \bar{f}_j)$  is high the probability of having  $q_j = \phi(f_j - \bar{f}_j)$  is very high and when  $(f_j - \bar{f}_j)$  is low the probability of having  $q_j = 0$  is high.

We have recently developed methods based on this approach for two application areas: medical imaging [9] and non destructive testing (NDT) imaging [10, 11, 12, 13]. In both cases, the main idea has been to estimate  $\mathbf{q}$  from the ultrasound echo-graphic data using

$$\hat{\mathbf{q}} = \arg \min_{\mathbf{q}} \{J(\mathbf{q}) = -\ln p(\mathbf{q} | \mathbf{g}_1)\} \quad (28)$$

with

$$J(\mathbf{q}) = \|\mathbf{g}_1 - \mathbf{H}_1 \mathbf{q}\|^2 + \lambda_1 \|\mathbf{q}\|^{\beta_1} \quad (29)$$

and use it in tomography reconstruction from the X-ray radiographic data using

$$\hat{\mathbf{f}} = \arg \min_{\mathbf{f}} \{J(\mathbf{f}) = -\ln p(\mathbf{f} | \hat{\mathbf{q}}, \mathbf{g}_2)\} \quad (30)$$

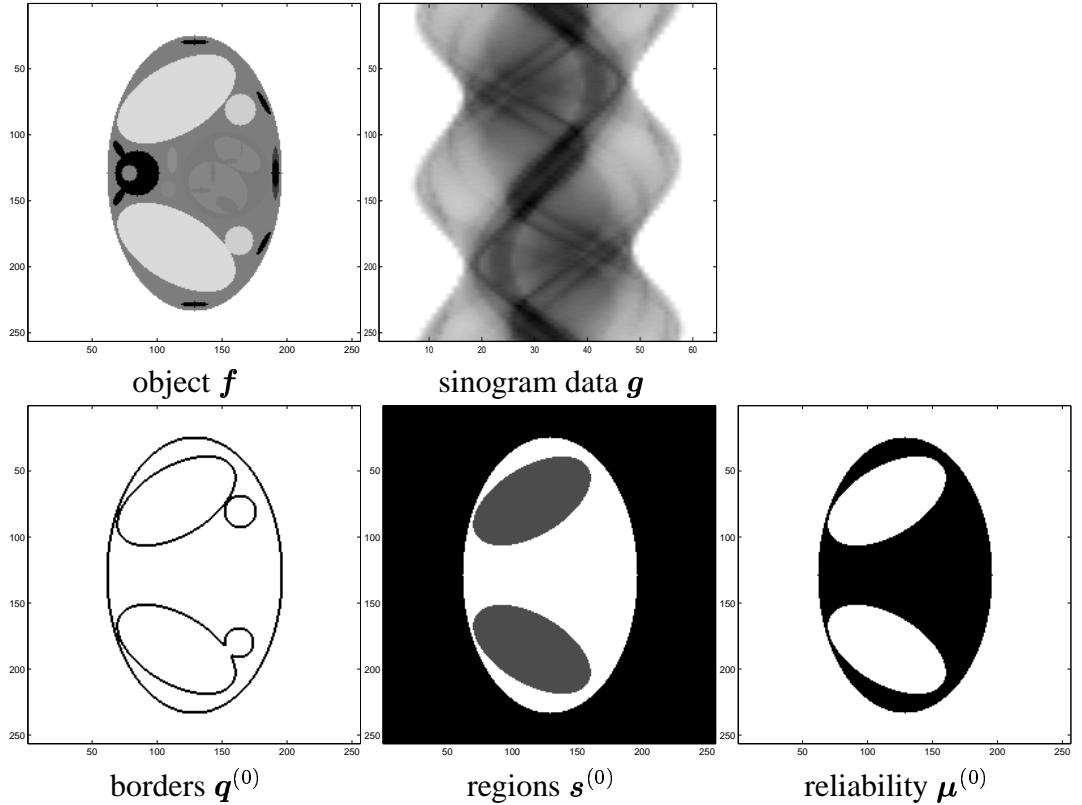
with

$$J(\mathbf{f}) = \|\mathbf{g}_2 - \mathbf{H}_2 \mathbf{f}\|^2 + \lambda_1 \sum_j (1 - \hat{q}_j) \phi(f_j - \bar{f}_j) \quad (31)$$

where  $\bar{f}_j$  is the mean value of pixels in the neighborhood of the pixel  $j$  and  $\phi(t) = \sqrt{1 + |t|^2}$ . The reader can refer to [9, 11, 12, 13] for more details.

## Fusion of radiographic and anatomical data in medical imaging

We applied the same approach in a situation where we have both radiographic data and some partial knowledge about the inside anatomy of the body. For example, assume that we can localize the positions of the borders  $\mathbf{q}^{(0)}$  of some of the regions and also we may know the material density in those or some other regions  $\mathbf{s}^{(0)}$  with a map of its corresponding degree of confidence  $\mu^{(0)}$ .



**FIGURE 5.** Fusion of radiographic and anatomical data in medical imaging

Again here, we can write the following equations:

$$\mathbf{g} = \mathbf{H}_1 \mathbf{f} + \epsilon_1 \longrightarrow p(\mathbf{g} | \mathbf{f}) \propto \exp\{||\mathbf{g} - \mathbf{H}_1 \mathbf{f}||^2\}$$

$$p(\mathbf{f}|\mathbf{q}^{(0)}) \propto \exp\left\{-\lambda_1 \sum_j (1 - q_j^{(0)}) \phi_1(f_j - \bar{f}_j)\right\}$$

$$p(\mathbf{f}|\mathbf{s}^{(0)}, \boldsymbol{\mu}^{(0)}) \propto \exp\left\{-\lambda_2 \sum_j \mu_j^{(0)} \phi_2(f_j - s_j^{(0)})\right\}$$

Then, based on these priors, we can write down the expression of the posterior  $p(\mathbf{f}|\mathbf{g}, \mathbf{q}^{(0)}, \mathbf{s}^{(0)}, \boldsymbol{\mu}^{(0)})$  and compute the MAP estimate:

$$\begin{aligned}\hat{\mathbf{f}} &= \arg \max_{\mathbf{f}} \{p(\mathbf{f}|\mathbf{g}, \mathbf{q}^{(0)}, \mathbf{s}^{(0)}, \boldsymbol{\mu}^{(0)})\} \\ &= \arg \min_{\mathbf{f}} \{J(\mathbf{f}) = -\ln p(\mathbf{f}|\mathbf{g}, \mathbf{q}^{(0)}, \mathbf{s}^{(0)}, \boldsymbol{\mu}^{(0)})\}\end{aligned}$$

$$J(\mathbf{f}|\mathbf{y}, \mathbf{q}, \mathbf{s}, \boldsymbol{\mu}) = \|\mathbf{y} - \mathbf{H}_1 \mathbf{x}\|^2 + \lambda_1 \sum_j (1 - q_j) \phi_1(x_{j+1} - x_j) + \lambda_2 \sum_j \mu_j \phi_2(x_j - s_j) \quad (32)$$

Note that, our knowledge about the borders and regions is partial. We may also want to estimate the unknown parts of borders and regions as well using an iterative algorithm as described in the following.

**Algorithm:**

1. Initialize  $\mathbf{q} = \mathbf{q}^{(0)}$  and  $\mathbf{s} = \mathbf{s}^{(0)}$  (and  $\boldsymbol{\mu} = \boldsymbol{\mu}^{(0)}$ )
2. Compute  $\mathbf{f}$  by optimizing  $J(\mathbf{f}|\mathbf{g}, \mathbf{q}, \mathbf{s}, \boldsymbol{\mu})$
3. Estimate new values for  $\mathbf{q}$  and  $\mathbf{s}$  (and  $\boldsymbol{\mu}$ ) from  $\mathbf{f}$  and  $\mathbf{q}^{(0)}$  and  $\mathbf{s}^{(0)}$  (and  $\boldsymbol{\mu}^{(0)}$ )
4. Return to 2. until convergence.

This algorithm has given satisfaction results as are shown in the next section and in [14, 9, 11]. However, one can do a better job by doing a better joint estimation via the joint posterior

$$p(\mathbf{f}, \mathbf{q}, \mathbf{s}, \boldsymbol{\mu}|\mathbf{g}) \propto p(\mathbf{g}|\mathbf{f}) p(\mathbf{f}|\mathbf{q}) p(\mathbf{q}) p(\mathbf{f}|\mathbf{s}, \boldsymbol{\mu}) p(\mathbf{s}) p(\boldsymbol{\mu})$$

and the following MCMC algorithm:

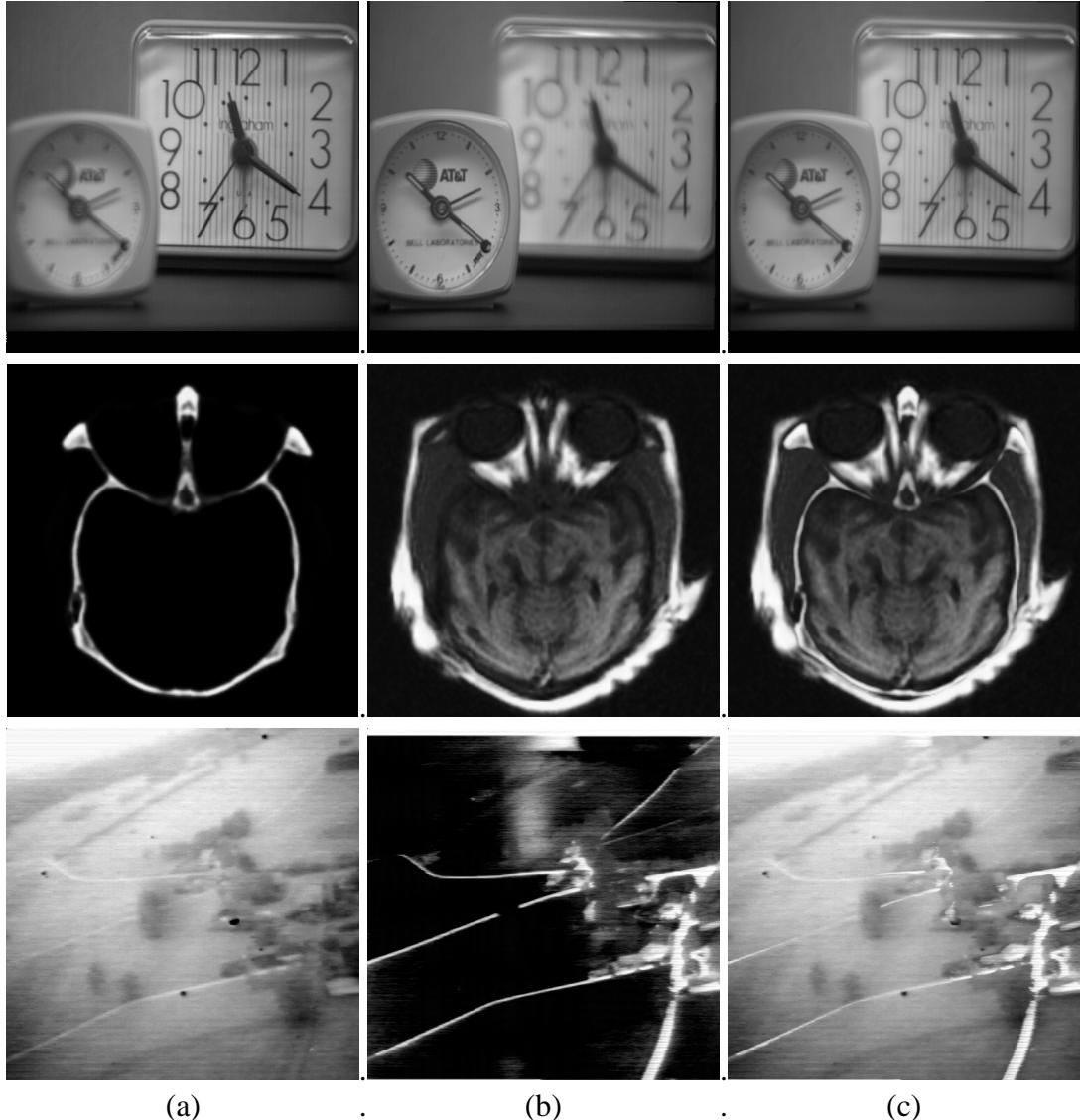
$$\begin{aligned}\hat{\mathbf{f}} &\sim p(\mathbf{f}|\mathbf{g}, \hat{\mathbf{q}}, \hat{\mathbf{s}}, \hat{\boldsymbol{\mu}}) \propto p(\mathbf{g}|\mathbf{f}) p(\mathbf{f}|\mathbf{q}) p(\mathbf{f}|\mathbf{s}, \boldsymbol{\mu}) \\ \hat{\mathbf{q}} &\sim p(\mathbf{q}|\mathbf{g}, \hat{\mathbf{f}}, \hat{\mathbf{s}}, \hat{\boldsymbol{\mu}}) \propto p(\mathbf{f}|\mathbf{q}) p(\mathbf{q}) \\ \hat{\mathbf{s}} &\sim p(\mathbf{s}|\mathbf{g}, \hat{\mathbf{f}}, \hat{\mathbf{q}}, \hat{\boldsymbol{\mu}}) \propto p(\mathbf{f}|\mathbf{s}, \boldsymbol{\mu}) p(\mathbf{s}) \\ \hat{\boldsymbol{\mu}} &\sim p(\boldsymbol{\mu}|\mathbf{g}, \hat{\mathbf{f}}, \hat{\mathbf{s}}, \hat{\mathbf{q}}) \propto p(\mathbf{f}|\mathbf{s}, \boldsymbol{\mu}) p(\boldsymbol{\mu})\end{aligned} \quad (33)$$

## SIMULATION RESULTS

In the following we show a few simulation results showing the feasibility of data and image fusion in different situations. The details of the algorithm implementations are omitted.

## FUSION OF REGISTERED IMAGES

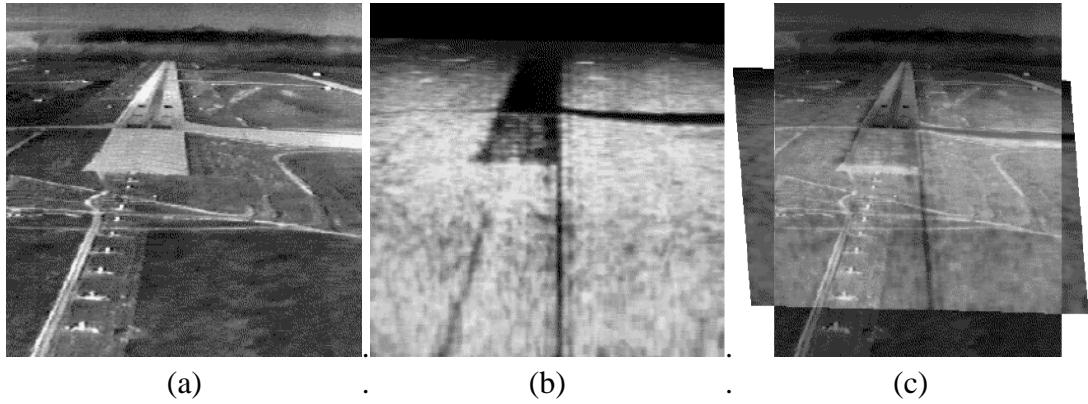
As we mentioned, image fusion when the images have been registered is an easy task. There are many simple techniques such as mean, median, PCA which are very easy to implement. There are also multi-resolution and pyramidal representation techniques which are based on the basic idea of doing fusion in each resolution or scale level before coming back to the original space. The following figure shows a few results obtained with these simple methods which are obtained using the Matlab package (<http://www.rockinger.purespace.de/indexp.htm>) developed by [15, 16, 17].



**FIGURE 6.** Fusion of registered images: The images in column (c) are the results of the fusion of images in column (a) and (b).

## Registration and fusion

When the images have not been registered, the main difficult part is the estimation of the registration parameters which needs a global optimization. Here also there are multi-grid, multi-resolution and pyramidal representation techniques. The basic idea here is to estimate these parameters in a coarser level and use them as initialization values when going to a finer level. The following figure shows one result obtained using a multi-grid optimization technique. The data have been taken from the following reference (<http://vision.arc.nasa.gov/personnel/al/hsr/fusion/97.html>) [18, 19].



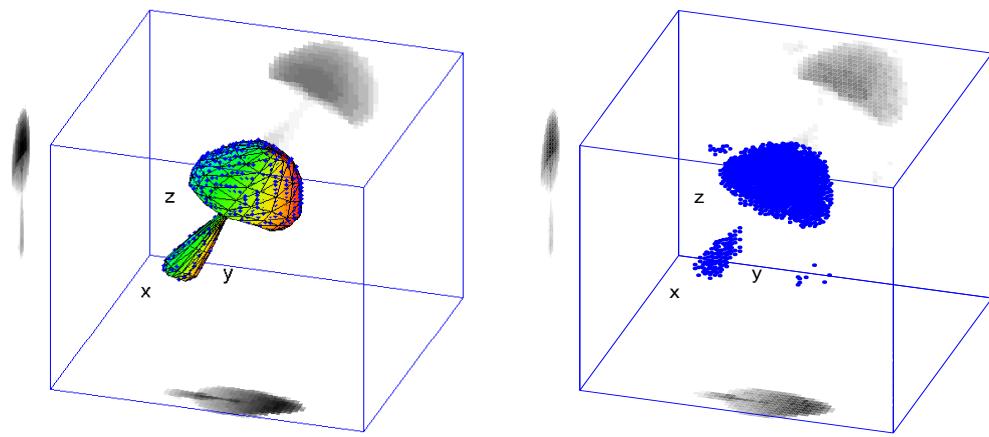
**FIGURE 7.** Fusion of unregistered images: column (c) is the result of the fusion of images in column (a) and (b).

## Shape from X-ray projection data

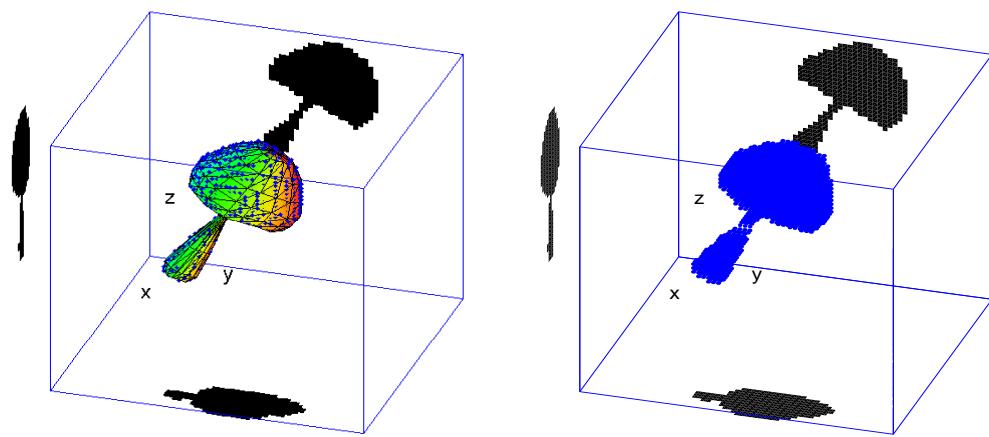
In computed tomography, it is possible to reconstruct any 3D volume from a finite set of its radiographic data. When the 3D volume consists of a compact and homogeneous body inside a homogeneous background, it is still easier to recover its shape from a more restricted number of its radiographic data in few directions. In the following figure we show an example of shape recovery from only three orthogonal projections. Note that, even if the shape is not convex, it is still possible to recover it from its radiographic data in a few directions.

## Shape from shadow

If we assume that the body is illuminated by a plane homogeneous light, the corresponding shadows are just the support functions of the radiographic data in the previous section. So, the problem of shape recovery from its shadows is more difficult (more ill-posed) than the previous problem of shape from radiographic data. However, it is still possible to recover the shape from shadows using the same kind of techniques of CT. The following figure shows an example of shape recovery from only three orthogonal shadows.



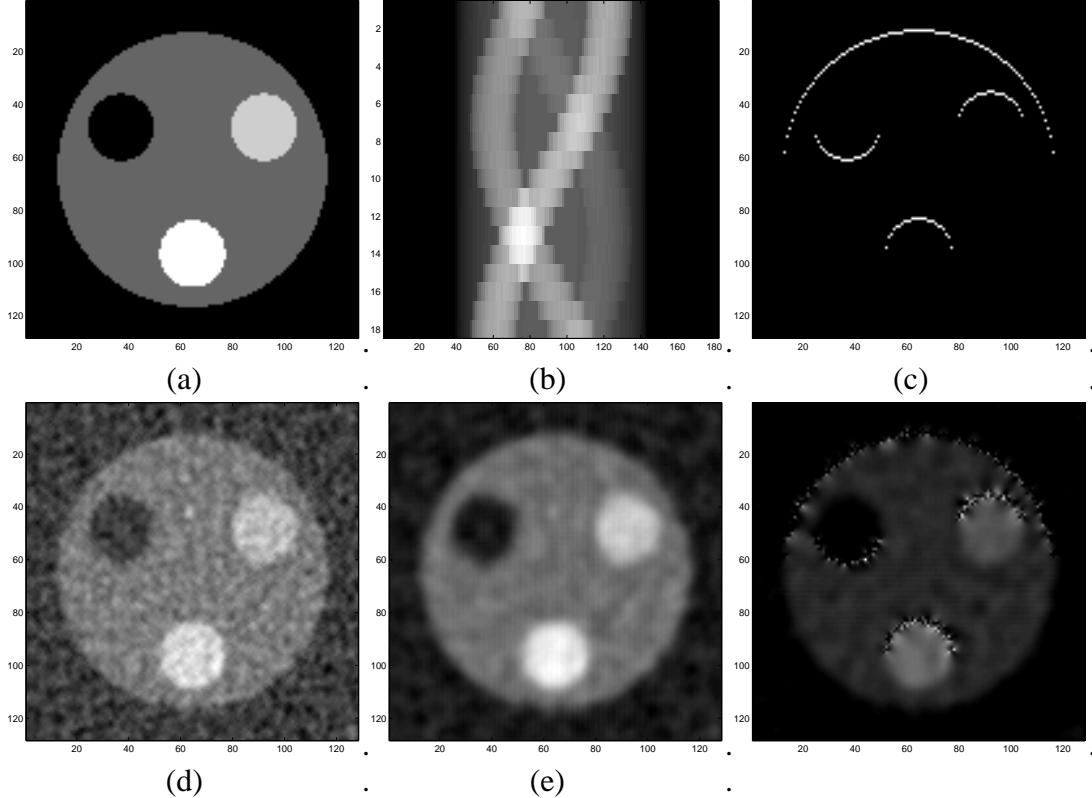
**FIGURE 8.** Shape recovery from X ray radiographic data (computed tomography): The left figure shows a 3D compact body and its three radiographic data. The right figure shows the shape recovered and its corresponding three radiographic data.



**FIGURE 9.** Shape recovery from shadow: The left figure shows a 3D compact body and its three shadows. The right figure shows the shape recovered and its corresponding shadows.

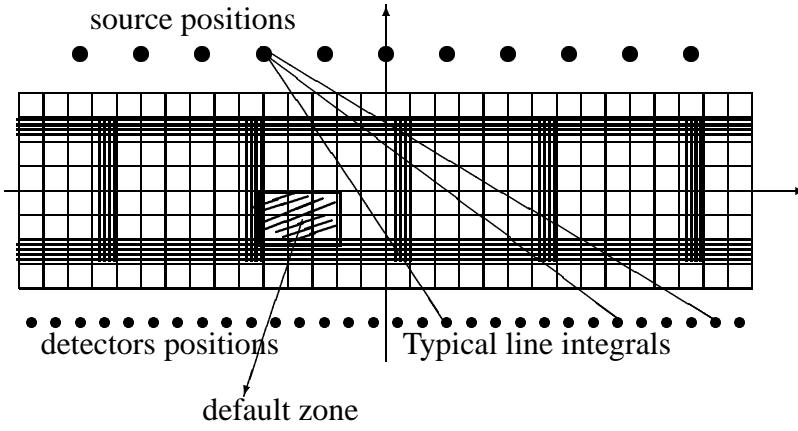
## Fusion of X-ray radiographic and ultrasound echo-graphic data

Here we give two examples of data fusion in computed tomography (CT) image reconstruction. The first example concerns X-ray projection data and ultrasound echo-graphic data. Here, the ultrasound data have first been processed to obtain the data in (c). Then it has been used in X-ray CT.

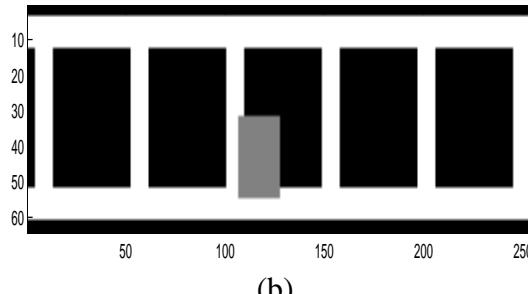


**FIGURE 10.** Data fusion in computed tomography: (a) original object (b) X-ray data, (c) ultrasound echo-graphic data, (d) reconstruction from X-ray data only by backprojection, (e) reconstruction from X-ray data only by proposed method, (f) reconstruction using both X-ray and echo-graphic data.

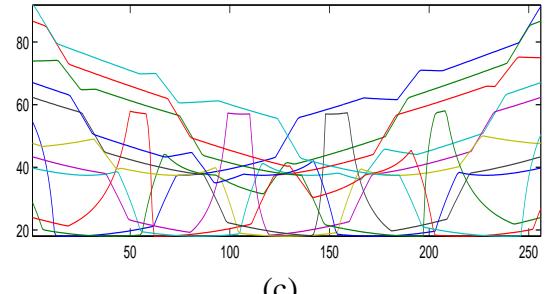
The second example concerns limited angle CT of sandwich structures such as an airplane wing. Here we use not only the X-ray data but also some geometrical knowledge of the structure.



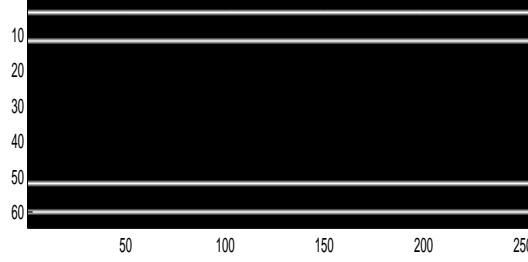
(a)



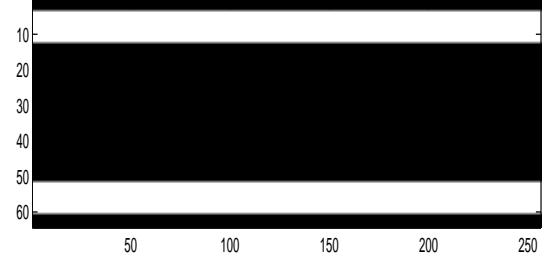
(b)



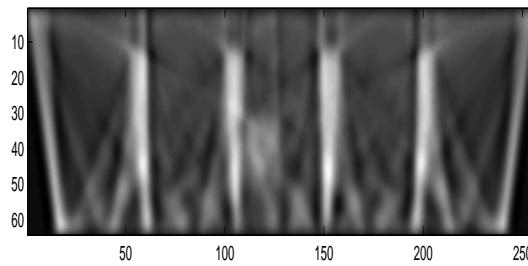
(c)



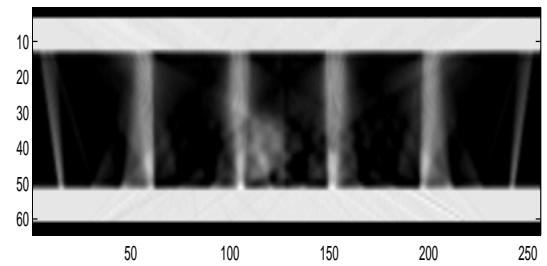
(d)



(e)



(f)

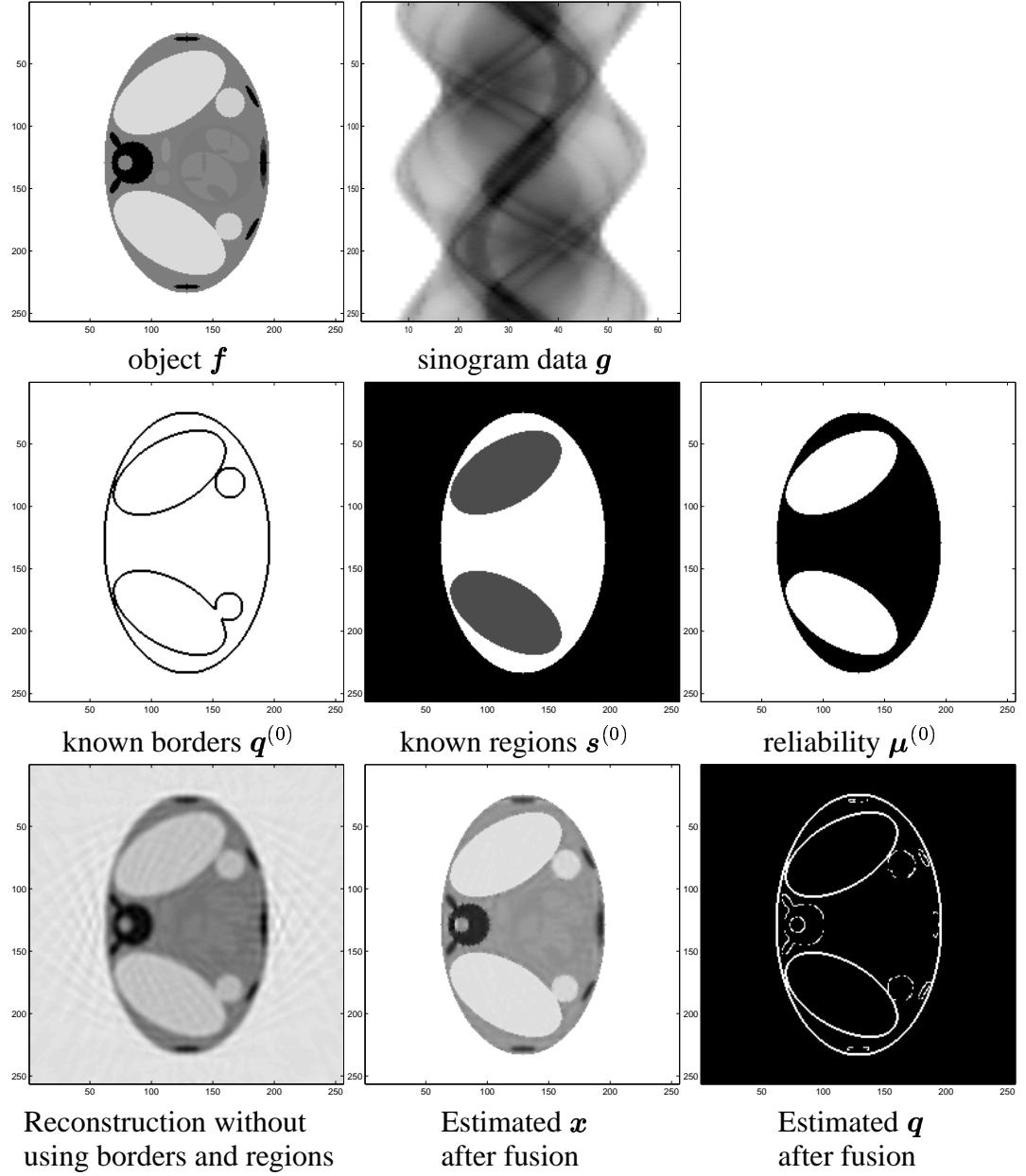


(g)

**FIGURE 11.** Data fusion in CND tomography: (a) is the data gathering system, (b) is the original object, (c) is the X-ray data, (d) is the known geometric borders data, (e) is the known geometric region values data, (f) is the reconstruction from X-ray data only, (g) is the reconstruction result using both X-ray and geometric data.

## Fusion of radiographic and anatomical data in medical imaging

Here we report a few examples of fusion results of radiographic and anatomical data in medical imaging.



**FIGURE 12.** Fusion of radiographic and anatomical data in medical imaging

## CONCLUSIONS

We presented a Bayesian estimation approach to image fusion and showed that simple models give simple data fusion algorithms. We then presented a general approach for heterogeneous data (different spaces) fusion. As an example, we considered X-ray radiographic and ultrasound echo-graphic data fusion for CT image reconstruction and presented some simulated results. We are working on the extension of this approach to 3D image reconstruction problem and will present some simulation results in final paper.

## REFERENCES

1. Knuth, K., "Bayesian source separation and localization", in *SPIE'98 Proceedings: Bayesian Inference for Inverse Problems, San Diego, CA*, edited by A. Mohammad-Djafari, 1998, pp. 147–158.
2. Knuth, K., and Vaughan JR., H., "Convergent Bayesian formulation of blind source separation and electromagnetic source estimation", in *MaxEnt 98 Proceedings: Int. Workshop on Maximum Entropy and Bayesian methods, Garching, Germany*, edited by F. R. von der Linden W., Dose W. and P. R., 1998.
3. Sharma, R., Leen, T., and Pavel, M., *Optical Engineering*, **40**, 1364–1376 (2001).
4. Daum, M., and Dudek, G., "On 3-D Surface Reconstruction Using Shape from Shadows", ????
5. Bourque, E., and Dudek, G., "Automated creation of imagebased virtual reality", in *Proc. SPIE Conference on Intelligent Systems and Manufacturing,(Pittsburgh,PA)*, 1997, pp. 292–303, URL [citeseer.nj.nec.com/bourque97automated.html](http://citeseer.nj.nec.com/bourque97automated.html).
6. Kirihara, S., and Saito, H., "Shape Modeling from Multiple View Images Using GAs", in *ACCV'98, Lecture Notes in Computer Science 1352*, 1998, vol. II, pp. 448 – 454.
7. Sato, Y., and Ikeuchi, K., "Recovering shape and reflectance properties from a sequence of range and color images", in *IEEE/SICE/RSJ International Conference on Multisensor Fusion and Integration for Intelligent Systems '96*, 1996, pp. 493–500.
8. Baker, S., Sim, T., and Kanade, T., *IEEE Transaction on Pattern Analysis and Machine Intelligence*, **25** (2003).
9. Mohammad-Djafari, A., "Hierarchical Markov modeling for fusion of X ray radiographic data and anatomical data in computed tomography", in *Int. Symposium on Biomedical Imaging (ISBI 2002), 7-10 Jul., Washington DC, USA*, 2002.
10. Gautier, S., Idier, J., Mohammad-Djafari, A., and Lavayssi  re, B., "X-ray and ultrasound data fusion", in *Proc. IEEE ICIP*, Chicago, IL, 1998, pp. 366–369.
11. Mohammad-Djafari, A., "Fusion of X ray and geometrical data in computed tomography for non destructive testing applications", in *Fusion 2002, 7-11 Jul., Annapolis, Maryland, USA*, 2002.
12. Mohammad-Djafari, A., "Bayesian approach with hierarchical Markov modeling for data fusion in image reconstruction applications", in *Fusion 2002, 7-11 Jul., Annapolis, Maryland, USA*, 2002.
13. Mohammad-Djafari, A., "Fusion of X ray radiographic data and anatomical data in computed tomography", in *ICIP 2002, 22-25 sep., Rochester, NY, USA*, 2002.
14. Mohammad-Djafari, A., Giovannelli, J.-F., Demoment, G., and Idier, J., *Int. Journal of Mass Spectrometry*, **215**, 175–193 (2002).
15. Rockinger, O., "Pixel level fusion of image sequences using wavelet frames", in *Proc. 16th Leeds Annual Statistical Research Workshop, Leeds University Press*, 1996, pp. 149–154.
16. Rockinger, O., "Image sequence fusion using a shift invariant wavelet transform", in *Proc. IEEE ICIP*, 1997, vol. III, pp. 288–291.
17. Rockinger, O., and Fechner, T., "Pixel-level image fusion: The case of image sequences", in *SPIE Proceedings*, 1998, vol. 3374, pp. 378–388.
18. Sharma, R., and Pavel, M., "Multisensor image registration", in *Society for Information Display*, 1997, vol. XXVIII, pp. 951–954.
19. Pavel, M., and Sharma, R., "Model-based sensor fusion for aviation", in *SPIE Proceedings*, 1997, vol. XXVIII, pp. 169–176.