

# Generalized Image Models and Their Application as Statistical Models of Images

Miguel Ángel González Ballester, Xavier Pennec, and Nicholas Ayache

Epidaure Project, INRIA, Sophia Antipolis, France  
Miguel.Gonzalez@sophia.inria.fr  
<http://www.inria.fr/epidaure>

**Abstract.** A generalized image model (GIM) is presented. Images are represented as sets of 4-dimensional sites combining position and intensity information, as well as their associated uncertainty and joint variation. This model seamlessly allows for the representation of both images and statistical models, as well as other representations such as landmarks or meshes. A GIM-based registration method aimed at the construction and application of statistical models of images is proposed. A procedure based on the iterative closest point (ICP) algorithm is modified to deal with features other than position and to integrate statistical information. Furthermore, we modify the ICP framework by using a Kalman filter to efficiently compute the transformation. The initialization and update of the statistical model are also described.

## 1 Introduction

This paper introduces a Generalized Image Model (GIM) aimed at unifying the representation of images, landmarks, point sets, and other shape representations. The statistical nature of data, reflecting uncertainty on measurements due to noise and device resolution, as well as the explicit representation of anatomical variability across patients and/or acceptable ranges for data values, is given special attention. In this paper, we describe how the GIM framework can be used to seamlessly model images and statistical models of image data in the same representation. A registration method based on the iterative closest point (ICP) algorithm [2] is given. We follow the principle of the ICP-4D algorithm [3] in regards to taking into account intensity as well as position. More importantly, the method has been extended to deal with the statistical nature of the data contained in the GIM. Thus, it applies without modification to the registration of either two images, an image and a statistical model, or two models. Optimization of the registration criterion is improved by substituting the classical ICP framework by a Kalman filter, thus obtaining considerable gain in computational time and ensuring convergence. Finally, the initialization and update of the statistical model is dealt with by computing simple statistical measures on the Voronoi tessellation defined by the elements of the GIM.

A description of the GIM and its use to represent in the same framework images and statistical models of image data are given in section 2. The ICP-based image-model registration algorithm follows in section 3, and the use of Kalman filters to efficiently find the solution is described in section 4. Finally, the initialization and update of the statistical model is given in section 5. Section 6 shows some preliminary results, and discussion and conclusions are provided in section 7.

## 2 Generalized Image Model (GIM)

The essence of the GIM is the explicit representation of position along with intensity information, and the inclusion of a covariance term to model uncertainty and correlation. In particular, an image  $A$  is modeled as a set of  $N_A$  sites  $\mathbf{v}_k$ , each one with an associated covariance matrix  $\Sigma_k$ :

$$A = \{(\mathbf{v}_k, \Sigma_k)\}_{k=1..N_A} \quad (1)$$

Each site contains position and intensity information. In the general case, the position component of a site  $\mathbf{v}_k$  is a  $D_p$ -dimensional (typically  $D_p = 3$ ) vector  $\mathbf{p}_k$ . Similarly, the intensity component can be generalized as a vector  $\mathbf{i}_k \in \mathfrak{R}^{D_i}$ , with  $D_i = 1$  in the case of scalar-valued images. Thus:

$$\mathbf{v}_k = (\mathbf{p}_k, \mathbf{i}_k) \quad (2)$$

The covariance matrix  $\Sigma_k$ , of dimensions  $(D_p + D_i) \times (D_p + D_i)$ , models the variance in position and intensity, and their possible correlation, at each site.

The choice of this image model is part of an attempt to seamlessly integrate information coming from different shape representations. Thus, an image is a particular case of GIM in which the positions are situated in a regular grid. Additionally, images with different resolutions can be combined in a natural fashion, as the uncertainty components can be used to model voxel sizes. More importantly, this representation can be used to include other models, such as point sets, landmarks or meshes (cf. discussion).

The use of the GIM to build statistical models of images is described in this paper. The same GIM framework is used to represent both an image  $A = \{(\mathbf{v}_k, \Sigma_k)\}_{k=1..N_A}$  and a statistical model  $M = \{(\bar{\mathbf{v}}_k, \bar{\Sigma}_k)\}_{k=1..N_M}$ . The key difference between them lies within the nature of the covariance matrices. The covariance matrix for an image models the uncertainty in the localization of the information (e.g. voxel size) and in the intensity information (e.g. noise in the imaging process). On the other hand, the covariance matrix for a model contains information about the variability in position and intensity of the set of samples used for training, i.e. it contains the shape variation model.

The following two sections describe an algorithm for the registration of images represented using the GIM, with a particular emphasis to the registration of an image and a statistical model. Section 5 will deal on the construction of such statistical models.

### 3 ICP-Based Registration of Images and Statistical Models

In this section we will describe how to compute the affine transformation that puts into correspondence two images represented following the GIM described in section 2. Due to the generality of the GIM, this method applies without modification for the registration of two images, an image and a statistical model, or two statistical models. The registration problem, illustrated for the case of the registration of an image and a statistical model, is as follows. For each site in the image,  $\mathbf{v}_k \in A$ , we aim at finding the index  $k'$  of the site  $\bar{\mathbf{v}}_{k'} \in M$  that best matches  $\mathbf{v}_k$ . Let us keep in mind that the information at each site contains not only location but also intensity (and potentially other features).

Our method is based on the Iterative Closest Point (ICP) [2]. This method was extended to 4-D as described in [3] in order to jointly match location and intensity (a technique known as ICP-4D). We modify this technique in order to handle the statistical information contained in the GIM, thus allowing it to deal with statistical models.

#### 3.1 Criterion to Be Minimized

The first approach is to replace the classical Euclidean distance based minimization criterion by a criterion based on the Mahalanobis distance between the current estimate of  $\mathbf{v}_k$  and each point in the model (which defines a Gaussian distribution). Thus:

$$k' = \arg \min_{k'} (\bar{\mathbf{v}}_{k'} - \mathbf{v}_k)^T \bar{\Sigma}_{k'}^{-1} (\bar{\mathbf{v}}_{k'} - \mathbf{v}_k) \quad (3)$$

Alternatively, the uncertainty in the location of  $\mathbf{v}_k$  can also be included in the criterion by measuring the distance between the distributions defined at each location of the model and the image. Several alternatives are available, depending on the distance measure of choice. Here we list three criteria based on a modified Mahalanobis distance, the Bhattacharyya distance, and the Kullback-Leibler divergence, respectively:

$$k' = \arg \min_{k'} (\bar{\mathbf{v}}_{k'} - \mathbf{v}_k)^T (\bar{\Sigma}_{k'} \Sigma_k)^{-1} (\bar{\mathbf{v}}_{k'} - \mathbf{v}_k) \quad (4)$$

$$k' = \arg \min_{k'} \left[ \frac{1}{8} (\bar{\mathbf{v}}_{k'} - \mathbf{v}_k)^T \left[ \frac{\bar{\boldsymbol{\Sigma}}_{k'} + \boldsymbol{\Sigma}_k}{2} \right]^{-1} (\bar{\mathbf{v}}_{k'} - \mathbf{v}_k) + 0.5 \log \frac{\frac{|\bar{\boldsymbol{\Sigma}}_{k'} + \boldsymbol{\Sigma}_k|}{2}}{\sqrt{|\bar{\boldsymbol{\Sigma}}_{k'}| |\boldsymbol{\Sigma}_k|}} \right] \quad (5)$$

$$k' = \arg \min_{k'} \frac{1}{2} (\bar{\mathbf{v}}_{k'} - \mathbf{v}_k)^T [\bar{\boldsymbol{\Sigma}}_{k'}^{-1} + \boldsymbol{\Sigma}_k^{-1}] (\bar{\mathbf{v}}_{k'} - \mathbf{v}_k) + \frac{1}{2} \text{tr}(\bar{\boldsymbol{\Sigma}}_{k'}^{-1} \boldsymbol{\Sigma}_k + \boldsymbol{\Sigma}_k^{-1} \bar{\boldsymbol{\Sigma}}_{k'} - 2\mathbf{I}) \quad (6)$$

A description and comparison of these distance measures can be found in [5].

Eventually, additional information can be included at each site. For example, curvature can be used to further constrain the ICP evolution. A particularly interesting case is that of an independent curvature term, which for the case of the Mahalanobis distance gives the following criterion:

$$k' = \arg \min_{k'} \left[ (\bar{\mathbf{v}}_{k'} - \mathbf{v}_k)^T \bar{\boldsymbol{\Sigma}}_{k'}^{-1} (\bar{\mathbf{v}}_{k'} - \mathbf{v}_k) + \frac{\kappa_{\bar{\mathbf{v}}_{k'}} - \kappa_{\mathbf{v}_k}}{\sigma_{\bar{\mathbf{v}}_{k'}}} \right], \quad (7)$$

where  $\kappa_{\mathbf{v}}$  is the curvature at site  $\mathbf{v}$  and  $\sigma_{\mathbf{v}}$  its standard deviation in the model.

### 3.2 Computation of the Transformation via ICP

The classical ICP framework [2] works in an iterative fashion repeating the following three phases: 1) establish correspondences between points in the two data sets being registered; 2) compute the least squares estimate of the transformation; and 3) apply the transformation. Although simple, this method has proven to be very effective in practice. Typically, ICP is used to compute a rigid transformation between point sets. Here, we use it to compute an affine transformation that will act jointly on position and intensity.

A modification of the classical approach for its use with statistical models could be envisaged. Searching for the closest point would no longer rely on Euclidean distance, but it would utilize one of the distance measures listed above. However, the minimization of a term encompassing one of such distances provides no closed form solution, and thus one has to rely on costly iterative gradient methods [4, 1]. An alternative, efficient method to minimize the criterion is given in the following section.

## 4 Solving for the Transformation via Kalman Filtering

The computation of the transformation in the classical ICP algorithm is based on the full set of point correspondences found for every point in the data set. This is clearly not optimal from a computational point of view, as the incremental updates of the estimate of the transformation at each iteration are released at a very slow rate. Here,

we propose a method for updating the transformation every time a correspondence is found, thus speeding up the computation (see [6] for a similar approach applied to tracking).

The computation of the transformation is formalized as a data assimilation task. A Kalman filter [7] is set up to this effect. Following the classical nomenclature for the description of Kalman filters, we identify the following constituent elements:

- the parameters  $\mathbf{x}$  of the affine transformation  $\mathbf{T}_k$  form the *internal state*, i.e. the parameters to estimate. For the case of a 2D scalar image the transformation is:

$$T_k = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 & \mathbf{x}_4 \\ \mathbf{x}_5 & \mathbf{x}_6 & \mathbf{x}_7 & \mathbf{x}_8 \\ \mathbf{x}_9 & \mathbf{x}_{10} & \mathbf{x}_{11} & \mathbf{x}_{12} \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (8)$$

- the image data point  $\mathbf{v}_k$  is the *measure*;
- $\Sigma_k$  models the incertitude of the measure;
- $\bar{\mathbf{v}}_k$ , and  $\bar{\Sigma}_k$ , are external parameters.

Each data point  $\mathbf{v}_k$  is treated individually, and an estimate of the transformation is refined each time a new point is treated. The system evolution is similar to the one employed for the estimation of a constant:

$$\mathbf{x}^{(t+1)} = \mathbf{x}^{(t)} + \mathbf{q} \quad (9)$$

where  $\mathbf{q}$  is drawn from a zero mean Gaussian distribution with covariance matrix  $\mathbf{Q}$ , modeling the *process noise*. The measurement function is as follows:

$$\hat{\mathbf{v}}_k(\mathbf{x}; \bar{\mathbf{v}}_k) = \mathbf{H}(\bar{\mathbf{v}}_k) \mathbf{x} + \mathbf{r} \quad (10)$$

where the matrix  $\mathbf{H}(\bar{\mathbf{v}}_k)$  is set up such that  $\mathbf{H}(\bar{\mathbf{v}}_k) \mathbf{x} = \mathbf{T}_k \bar{\mathbf{v}}_k$ , and  $\mathbf{r}$  is drawn from a zero mean Gaussian distribution with covariance matrix  $\mathbf{R}$ , modeling the *measurement noise*.

The matrix  $\mathbf{Q}$  is taken to be diagonal with very small values for variances. In fact, it could simply be removed, as we assume that the transformation being estimated remains constant throughout the process; however, we found that initializing the process variance to a very small value helps reaching convergence faster. As for the matrix  $\mathbf{R}$ , it models the noise in the measurement process. Therefore it seems natural that the covariance matrix  $\Sigma_k$  should be used in its place.

The resulting Kalman filter is a linear one, which is solved efficiently (Extended Kalman Filters can be used for non-linear systems [7]). Furthermore, the size of the matrices to be inverted is very small, as each data point is treated separately. The update equations for Kalman filters can be found, for example, in [7].

## 5 Initializing and Updating the Statistical Model

The previous section showed how to perform the registration of a statistical model and an image. Once the registration has been successfully estimated, the information contained in the image can be used to update the model, or in the case of two models being registered, the information contained in both of them can be fused. In this section, we comment on the process of construction of statistical models as represented by the GIM.

The initialization of a model is obtained by registering a set of training images. To this aim, the method described in the previous section can be used (taking covariance matrices proportional to the voxel size of each image). Alternatively, any other registration method can be applied.

The mean position and intensity values  $\bar{\mathbf{v}}_k$ , and covariance matrices  $\bar{\Sigma}_k$  are then estimated from the set of registered images. The Voronoi tessellation defined by the elements of the GIM determines which sites of the images are used to compute each site in the model. Note that the transformation  $\mathbf{T}_k$  is also applied to the covariance matrix, i.e. the covariance associated to  $\mathbf{T}_k \mathbf{v}_k$  is  $\mathbf{T}_k \Sigma_k \mathbf{T}_k^T$ .

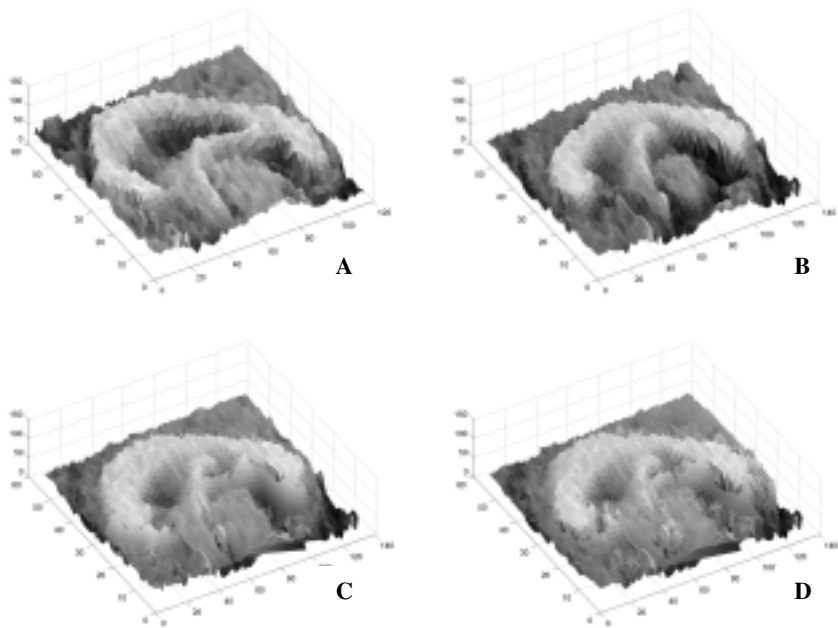
The reference image to which all images are registered can be an initial rough model (cf. the example in the following section), or one of the images in the training set. In this last case, in order to eliminate biases due to the choice of an arbitrary reference image, the process is iterated using the new average image as reference. This same procedure can be used to update the model upon arrival of a new registered image.

## 6 Results

An illustration of the relevance the covariance matrices play in the registration process is shown in figure 1. In this case, we consider the problem of finding the affine transformation that maps the corpora callosa (CCs) of two brains as seen on T1-weighted MRI. A rectangular image region containing the CC was taken from a central sagittal slice of each MRI. Figs 1A and 1B show surface plots, with height corresponding to intensity. Clearly, they are not aligned. We employed the ICP registration procedure described above with covariance matrices equal to the identity matrix (which is equivalent to a performing a standard ICP) to obtain the results in fig 1C. The two CCs were correctly aligned. Then, we modified the covariance matrices to artificially augment the variability in the posterior part of the CC, by assigning a different covariance matrix to this part (Mahalanobis distance was used). The result fits more closely to the anterior part of the CC, as seen on fig 1D.

We also performed some preliminary experiments on the construction of a statistical image model of the CC. A central sagittal slice containing the CC was identified in 10 data sets, and a rectangular image region including the CC was extracted. The statistical model was initialized by smoothing one of the images and subsampling it to

a very small resolution. The model has only  $9 \times 18$  sites. Each image was registered to the model using the Kalman filter approach in section 4, with a covariance matrix for the image diagonal with values proportional to the voxel size and an initial covariance matrix for the model proportional to the size of the voxel region mapping to each site in the model. Then, the mean and covariance matrix of all the registered images were computed and assigned to each site of the model. Figure 2 shows the sites of the GIM as 3D points. A singular value decomposition of the covariance matrix at each site was performed, and the dominant eigenvector is represented in figure 2.

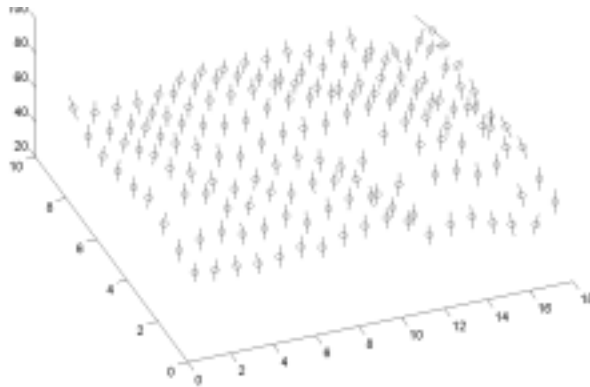


**Fig. 1.** **A)** and **B)** Surface plots of two image regions containing the corpus callosum (CC), extracted from two T1-weighted MR images; **C)** Affine registration of A to fit B using a standard ICP (the two surfaces are shown superimposed); **D)** Registration using the Mahalanobis distance with different covariance matrices in the posterior part of the CC, allowing for more variability. This results in a better fit to the anterior part of the CC.

## 7 Discussion and Conclusions

Notwithstanding the methodological contributions in the registration method, we consider that the most important element in our framework is the generalized image model presented in section 2. Conceptually simple, it allows for the seamless integration of different types of information, and we are but starting to evaluate its real possibilities.

The computation of a statistical model encompassing location and intensity information was presented. Here as well, several promising directions of future research are to be explored. The registration of point sets and meshes to images, as well as techniques for non-distorting resampling based on the GIM are being studied. The Kalman framework can be extended to model neighborhood interactions, and non-trivial system evolutions.



**Fig. 2.** 3D view of the statistical model. Height represents intensity, and the remaining dimensions are position components. The model has  $9 \times 18$  sites. The dominant eigenvector of the covariance matrix is shown.

## References

1. Batchelor, P. G. and Fitzpatrick, J. M.: A Study of the Anisotropically Weighted Procrustes Problem. In: *Procs. of the Workshop on Mathematical Methods in Biomedical Image Analysis* (2000) 212–218.
2. Besl, P. and McKay, N.: A Method for Registration of 3D Shapes. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 18(14) (1992) 239–256.
3. Feldmar, J., Declerck, J., Malandain, G., Ayache, N.: Extension of the ICP Algorithm to Nonrigid Intensity-Based Registration of 3D Volumes. In: *Computer Vision and Image Understanding*, Vol. 66(2) (1997) 193–206.
4. Pennec, X. and Thirion, J.-P.: A Framework for Uncertainty and Validation of 3D Registration Methods Based on Points and Frames. In: *Int. Journal of Computer Vision*, Vol. 25(3) (1997) 203–229.
5. Sooful, J. and Botha, E.: An Acoustic Measure for Automatic Cross-Language Phoneme Mapping. In: *Twelfth Annual Symposium of the South African Pattern Recognition Association*, Franschoek, South Africa (2001) 99–102.
6. Welch, G. and Bishop, G.: SCAAT: Incremental Tracking with Incomplete Information. In: *Procs. of SIGGRAPH'97*, Los Angeles, USA (1997).
7. Welch, G. and Bishop, G.: An Introduction to the Kalman Filter. Tutorial at SIGGRAPH 2001, Los Angeles, USA (2001) <http://www.cs.unc.edu/~welch/kalman/>