A System for Multimodality Image Fusion of the Spine

Paul F. Hemler, Thilaka Sumanaweera, Ramani Pichumani
Petra A. van den Elsen, Sandy Napel, John Drace, John Adler
Stanford University Medical Center, Stanford University, Stanford CA

Abstract

This paper describes a semi-automatic system for registering and visualizing CT (Computer Tomography) and MR (Magnetic Resonance) images of the cervical spine. Registration requires identifying similar objects or structures in each image set. Identifying similar structures for this application is complicated because complementary imaging modalities are used. Following structure identification, the mapping transformation from one image set to the other can be determined using a non-linear optimization procedure. The final step is visualizing the results of the registered images sets.

Introduction

By image fusion we mean the registration and visualization of different sets of images of a given object or scene. Image fusion is important when diagnosing disorders of the spine because spinal pathology typically involves the interaction of different tissue types which are best imaged with different devices. For example, Computerized Tomography (CT) provides the best description of hard or dense structures, such as bone, while Magnetic Resonance (MR) provides better discrimination of soft tissues. The spinal surgeon routinely orders both CT and MR image studies, views each image set independently, and then mentally merges the image information. This implicit image fusion is subjective and prone to error. The goal of this work is to reduce the probability and magnitude of this error by replacing the mental image fusion process with a semi-automatic fusion system.

The system was originally designed for fusing images of the brain and has been augmented for the spine. Image fusion of the spine is more difficult than the brain because the spine is not a rigid body, but rather, an articulating structure. Additionally, some of the elements of the spine are small with respect to the resolution of the imaging devices, making accurate localization a difficult task.

The registration process begins by automatically identifying (segmenting) 3-D surfaces of similar anatomical structures in each image set. Reliable automatic segmentation of CT and MR images is difficult because of limited spatial and contrast resolution, noise and partial volume effects. Therefore, this system provides an interactive tool to visualize and edit the automatically detected structures. A novel 3-D surface matching technique is then applied to determine the optimal (in the least mean-squared sense) transformation that maps one set of structures to the other. This surface matching technique only requires a sampling of the surfaces of the same anatomical structure, and does not assume a correspondence between sample points. The resulting transformation is a rigid body motion which is appropriate if the structures are themselves rigid bodies and both image sets are free of geometric distortion. To satisfy these constraints, we use the vertebrae (bone) as the anatomic landmarks and correct the MR image set [Sumanaweera, 1992] for both gradient field non-linearities and magnetic susceptibility variations. Once the optimal rigid body transformation is determined, the system provides several means of visualization. We find that a composite image set formed by mapping bone points from CT to the MR image set is quite useful for diagnosis.

System Description

The fusion system consists of the following components: image acquisition, image processing, interactive editing, 3D surface formation, optimization, and visualization.

Image Acquisition: We use a spiral CT scanner to acquire high resolution images at $1mm$ collimation and $1mm/sec$ table speed and then retrospectively reconstruct image planes at $0.5mm$ spacing. This resolution insures that small objects such as the vertebrae are imaged properly and that little anatomic change occurs between images. We use a small field of view (FOV) resulting in an in-plane pixel size of $0.3mm^2$.

*This work was supported in part by the Veterans Administration and the Whitaker Foundation

ıere are many parameters that need to be tuned to ɔduce adequate MR images for our system. Spatial ;olution, signal-to-noise (SNR), tissue contrast mechisms and imaging time are all inter-related and suitle tradeoffs must be evaluated. Through extensive ɔerimentation, we have developed a reasonable protofor MR images of the cervical spine which highlight : disc/bone boundary. This boundary is of interest registration because it appears in CT images as well. : have found that a 3D fast gradient echo sequence :h a sagittal imaging plane, an in-plane resolution of mm^2, and a through-plane resolution of $1.2mm$ gives :eptable resolution, contrast, and SNR.

.age Processing: The registration process deterıes the transformation for mapping one image set to other by minimizing the total squared distance beıen similar anatomic surfaces in each image. Since ' reveals bone well and MR depicts soft tissue such disc, we choose to match the interface between the tebrae and the disc. Unfortunately, this surface may be visible in the axial view generated by the CT nner since the axial plane is often nearly parallel to ; surface. Instead, this surface is best depicted in a ittal or coronal views. We choose to use the sagittal w since it matches the imaging plane generated by :. The first step in processing the data is to reformat axial CT images along the sagittal plane.

:e the images have been reformatted into similar ıes, an edge detection algorithm is used to deterıe intensity variations occurring in the image. We a standard Canny [Canny, 1986] edge detection ɜme, which works well because of the high resoluı and contrast of the images. Once the edges are de.ed, they are tracked using a connected components senfeld and Kak, 1982] tracking algorithm. Finally, edges are segmented by using information about general location of the edges corresponding to the /vertebrae junction. We have developed an interacedge editing tool to view the segmented edges and ık or join the edges if necessary.

face Formation: Our system builds 3D surfaces ı a stacked set of 2D edges. The 2D edges are ɜely sampled and consist of many redundant points. remove some of this redundancy and improve the iency of the registration process, the edges on each : are approximated by a polygon. That is, we ace all edge elements which are close to a line ner, 1972] by the end points of the line.

ɜn fusing images of the brain, we found that objects rigidly related to the skull, for example, the ears, to be removed. We interactively edited the edges to ove these objects, resulting in broken or open con-s. We extended a classical surface reconstruction rithm [Fuchs *et al.*, 1977] to connect open edges ı one plane to open edges on the next plane. The 3D surfaces formed by this process are small triangular surface patches resulting in a surface tiling of the object. This tiling is a piecewise linear approximation to the actual surface and our algorithm generates a visibly correct description of the surface. Accurate surface representations [Pelizzari *et al.*, 1989, Jiang *et al.*, 1992] are fundamentally important when using surface information to compute mapping transformation.

Transformation Formulation: When determining the transformation which maps a rigid object in one set to a similar object in the other set, we assume that any geometric distortions have been corrected [Sumanaweera, 1992]. With this assumption, a rigid body transformation is sufficient to determine the mapping from one image set to the other. Mathematically, a rigid body transform can be represented by the following quaternion equation:

$$\dot{\mathbf{P}} = \dot{\mathbf{R}}\dot{\mathbf{P}}'\dot{\mathbf{R}}^* + \dot{\mathbf{T}}, \tag{1}$$

where $\dot{\mathbf{P}}$ is the quaternion representation of the transformed point, $\dot{\mathbf{R}}$ is a rotation quaternion, $\dot{\mathbf{P}}'$ is the quaternion representation of the original point, $\dot{\mathbf{R}}^*$ is the conjugate of the rotation quaternion, and $\dot{\mathbf{T}}$ is a translation quaternion. A quaternion can be considered a four element vector containing both real and imaginary parts. A set of operations, such as, addition, multiplication, are defined for a quaternions. For example, the quaternion $\dot{\mathbf{R}}$ is,

$$\dot{\mathbf{R}} = r_0 + r_x i + r_y j + r_z k \tag{2}$$

where r_0 is the real part and the remaining terms are the imaginary part. A three-dimensional point is represented as a quaternion by setting the real part of the quaternion to zero and letting $r_x = X$, $r_y = Y$, and $r_z = Z$. A rotation quaternion represents a rotation of θ degrees about and arbitrary axis. Specifically, a rotation of θ degrees about an axis $\mathbf{V}$ is,

$$\dot{\mathbf{R}} = \cos(\frac{\theta}{2}) + \sin(\frac{\theta}{2})\mathbf{V} \tag{3}$$

For a quaternion to be a rotation quaternion its 2-norm must equal one. That is, $\dot{\mathbf{R}}$ is a rotation quaternion if,

$$r_0^2 + r_x^2 + r_y^2 + r_z^2 = 1. \tag{4}$$

The quaternion representation for describing a rigid body transformation has several advantages over the matrix/vector representation. The quaternion representation requires four parameters to describe a rotation and the matrix/vector representation requires nine. Another advantage is that it is easier to determine if a quaternion is a rotation quaternion by determining if its 2-norm is equal to one. This calculation is somewhat simplier than determining if a matrix is orthonormal. Finally, an arbitrary quaternion can be converted into a rotation quaternion by normalizing it. Again, this is

considerably simpler than converting an arbitrary matrix into a rotation matrix.

In this application, the quantities $\dot{\mathbf{P}}$ and $\dot{\mathbf{P}}'$ are known and we seek to determine the rotation and translation quaternions, $\dot{\mathbf{R}}$ and $\dot{\mathbf{T}}$ such that the point $\dot{\mathbf{P}}$ maps to the point $\dot{\mathbf{P}}'$ in some optimal sense. Since the mapping relationship is non-linear and can not be solved in closed form, the parameters of the rotation and translation quaternions must be determined iteratively.

Optimization: Non-linear optimization techniques can be used to solve a set of non-linear equations. These techniques are iterative, where successive steps in the iteration seek better solutions. Solutions are evaluated based on an error function or cost function. The optimal solution is the one that minimizes the error function.

There are many non-linear optimization techniques ranging from an exhaustive search of all possible solutions to those based on the gradient of the error function. Exhaustive searches become infeasible when there are more than a few parameters. Gradient based methods require a good initial estimate of the optimal solution and that the error surface be quadratic near the minimum. At each iteration the parameters are adjusted according to the direction of the gradient, so successive solutions follow the "downhill" slope of the error surface, converging at the minimum in a few iterations. Although it is difficult in general prove the quadratic nature of an arbitrary error surface, good initial estimates of the optimal parameters can be derived easily. Specifically, the patient is in roughly the same orientation during both CT and MR scanning resulting in no apparent rotation. The translation is estimated by determining the difference in the location of the center-of-mass of each image set.

We use a non-linear optimization algorithm developed by Marquardt [Press *et al.*, 1992] which automatically adapts between the gradient descent and the inverse Hessian technique. When the solution is far from the minimum, the technique closely resembles gradient descent, and when the solution is near a minimum, it is similar to the inverse Hessian technique. In general, this technique has very good convergence properties.

The error function we use is the total squared distance between surface points in one space and the approximated surface in the other space. Let p_i' be a point in one image space. Then,

$$p_i(\mathbf{a}) = T(p_i') \tag{5}$$

is the corresponding point in the other image space. In equation 5, $p_i(\mathbf{a})$ is written as an explicit function of the transformation vector $\mathbf{a}$, a seven element vector defined as,

$$\mathbf{a} = [r_0, r_x, r_y, r_z, t_x, t_y, t_z]^T. \tag{6}$$

Let q_i be the point on the surface that is closest to the point $p_i(\mathbf{a})$. The distance between the transformed point and its neighboring surface can be written as a function of $\mathbf{a}$:

$$d_i^2(\mathbf{a}) = (q_i - T(p_i'))^2 = (q_i - p_i(\mathbf{a}))^2 = \|(q_i - p_i(\mathbf{a}))\|^2 = (q_i - p_i(\mathbf{a})) \cdot (q_i - p_i(\mathbf{a})). \tag{7}$$

The total squared distance between all transformed points and the corresponding surface can be expressed as:

$$D^2 = \sum_{i=1}^{N} d_i^2(\mathbf{a}) = \sum_{i=1}^{N} (q_i - p_i(\mathbf{a}))^2. \tag{8}$$

Equation 8 measures the total squared distance between a set of transformed points $p_i(\mathbf{a})$ and surfaces q_i. The transformed points are determined from the original points p_i' and the set of parameters $\mathbf{a}$. The points q_i are the closest points on the surface to the points $p_i(\mathbf{a})$. The computation of these points is simplified by the piecewise linear approximation to the actual surface. Specifically, the surface of the object is represented by a triangular tiling. Without loss of generality, let $P1, P2$, and $P3$ be the three vertices of a triangle in the tiling. A point q on the plane of the triangle can be described in relation to these points using,

$$q = \lambda P_1 + \mu P_2 + (1 - \lambda - \mu) P_3. \tag{9}$$

The shortest distance between a point p and the triangle will be some distance along a perpendicular line to the triangle. That is, the vector joining points p and q will be perpendicular to all the lines in the plane. Two constraints on the location of q can be determined using two of the edges of the triangle. Consider the edges between points P_1 and P_2 and between points P_1 and P_3, then the constrains can be written as,

$$(p - q) \cdot (P_1 - P_2) = 0 \tag{10}$$

and

$$(p - q) \cdot (P_1 - P_3) = 0. \tag{11}$$

After substituting for q and rearranging the terms in equation 10 we arrive at an equation involving all known quantities and the two unknown quantities λ and μ. The equation is:

$$(p - P_3) \cdot (P_1 - P_2) = \lambda[(P_1 - P_2) \cdot (P_1 - P_3)] + \mu[(P_1 - P_2) \cdot (P_2 - P_3)]. \tag{12}$$

Similarly for equation 11,

$$(p - P_3) \cdot (P_1 - P_3) = \lambda[(P_1 - P_3) \cdot (P_1 - P_3)] + \mu[(P_1 - P_3) \cdot (P_2 - P_3)]. \tag{13}$$

This is a system of two linear equations in two unknowns. By substituting variables into the equations above, we get,

$$\begin{bmatrix} E \\ F \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} \lambda \\ \mu \end{bmatrix} \tag{14}$$

ere

$$A = (P_1 - P_2) \cdot (P_1 - P_3) \quad (15)$$

$$B = (P_1 - P_2) \cdot (P_2 - P_3) \quad (16)$$

$$C = (P_1 - P_3) \cdot (P_1 - P_3) \quad (17)$$

$$D = (P_1 - P_3) \cdot (P_2 - P_3) \quad (18)$$

$$E = (p - P_3) \cdot (P_1 - P_2) \quad (19)$$

$$F = (p - P_3) \cdot (P_1 - P_3). \quad (20)$$

e solution for the two unknowns in equation 14 are ily derived and are:

$$\lambda = \frac{(DE - BF)}{AD - BC} \quad (21)$$

l

$$\mu = \frac{-(CA + AF)}{AD - BC}. \quad (22)$$

e point on the plane q closest to the point p is found substituting the values of λ and μ into equation 9. e point q is in the triangle defined by points P_1, P_2, P_3

$$\lambda \geq 0, \mu \geq 0, (\lambda + \mu) \leq 1. \quad (23)$$

ial tests of our system show that this formulation ılts in a transformation which is stable with respect nitial conditions and noise. Our future work will be ed at quantifying our solution and comparing it to er approaches.

ualization: Once two image sets are registered the em provides several methods of visualization. One ınique permits the user to select a point of inter- in one of the image sets and the system then maps point onto the other image set. Two images are layed where one image shows an acquired slice of image set and the other image shows a resampling he other image set. The resampling is such that ı images display similar anatomy. Red markers are layed at the in-plane location of the selected point oth images. *Split screen* is a variation of this point ping technique where portions of an acquired image the resampled image are displayed simultaneously. height of both image planes is controlled interac- ly, providing motion cues which are sometimes use- n evaluating accuracy and in determining the prox- y of different structures. One final technique maps llection of points from one image set to the other, ıing a composite image set. We find it most useful ıreshold the bone in CT and insert these points into MR image set.

Conclusion

are building a system for fusing images of the cer- l spine. In this system we seek to minimize the unt of user interaction by automatically segment- the surfaces of certain anatomical landmarks. The marks are then matched and the transformation parameters are determined. We have used this system for fusing CT and MR images of the brain with excellent visual results and are in the process of extending it to the cervical spine.

References

[Canny, 1986] J. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6):676–697, November 1986.

[Fuchs *et al.*, 1977] H. Fuchs, Z.M. Kedem, and S.P. Uselton. Optimal Surface Reconstruction from Planar Contors. *Communication of the ACM*, 20(10):693–702, October 1977.

[Jiang *et al.*, 1992] H. Jiang, R. Robb, and K.S. Holton. A new approach to 3-D registration of multimodality medical images by surface matching. In R.A. Robb, editor, *Proceedings SPIE Vol 1808 Visualization in Biomedical Computing*, pages 196–213, Bellingham, WA, 1992. SPIE Press.

[Pelizzari *et al.*, 1989] C.A. Pelizzari, G.T.Y. Chen, D.R. Spelbring, R.R. Weichselbaum, and C.T. Chen. Accurate three-dimensional registration of ct, pet, and/or mr images of the brain. *Journal of Computer Assisted Tomography*, 13(1):20–26, 1989.

[Press *et al.*, 1992] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Numerical Recipes in C The Art of Scientific Computing.* Cambridge University Press, 1992.

[Ramer, 1972] U. Ramer. An iterative procedure for the polygonal approximation of plane curves. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1:244–256, 1972.

[Rosenfeld and Kak, 1982] A. Rosenfeld and A. Kak. *Digital Picture Processing.* Academic Press, 1982.

[Sumanaweera, 1992] T.S. Sumanaweera. *Segmentation and Distortion Correction in Medical Imaging.* PhD thesis, Stanford University, Dept. of Electrical Engineering, Stanford, CA 94305, 1992.