Robust Automated Registration of Wrist Bones Using Tissue Classification and Distance Fields

G. Elisabeta Marai, David H. Laidlaw, and Joseph J. Crisco

Manuscript received ; revised . This work was supported in part by NSF CCR-0093238 and NIH AR44005.

*G.E. Marai and D.H. Laidlaw are at the Department of Computer Science, Brown University, Providence, RI 02912. E-mail: gem@cs.brown.edu.

J.J. Crisco is with the Department of Orthopaedics and the Division of Engineering, Brown Medical School/Rhode Island Hospital, Providence, RI 02912.

Abstract

We present an intra-subject method for automated registration of wrist bones from computed tomography (CT) volume images. The method makes possible the automated extraction of joint kinematic information from sets of volume images of wrists in different poses. The images may be acquired at relatively low resolution, reducing the total acquisition time per human subject and general scan storage requirements. The method is likely to be applicable to bones of other joints.

Our registration method works on a sequence of CT images of the same wrist in two steps: in the first step, each CT volume is processed by a tissue classifier that generates a localized distance field. The distance field approaches zero at bone boundaries. In the second step, the distance fields and a geometric bone model are used to track bone motion through the sequence of CT images. The distance value of each vertex in the geometric model is looked up in the corresponding distance field. The bone position and orientation are then automatically adjusted to minimize these distances. We infer a motion-governing bone hierarchy from the distance information. The derived hierarchy may have biological significance. We use the joint hierarchy to speed up the registration process.

We validate the method on synthetic, *in vitro*, and *in vivo* data. Our method maintains the stability of the manual approach, while reducing the amount of user interaction by a factor of N when applied to sets of N volume images. Accuracy is within manual registration variation. Validation against marked *in vitro* data shows registration errors under 1 deg (rotation) and under 1 mm (translation).

Index Terms

Medical imaging, registration, joint kinematics, wrist, distance fields, tissue classification.

I. INTRODUCTION

Analyzing *in vivo* joint kinematics in normal and injured wrists requires the measurement of motion among a large number of wrist poses. One way to achieve motion measurements is through the segmentation of wrist bones from 3D CT images and their registration over various poses. However, the segmentation and registration of wrist bones from CT images are notably challenging tasks. Factors that make these processes difficult include the small size of tightly packed bones (Fig. 1), which may move relative to one another; the complex shape of these bones; the presence in volume images of partially scanned bones (forearm and metacarpals); the non-uniformity of wrist bone tissue, ranging from dense, compact cortical bone to textured spongy bone; the blurring and partial-volume effects inherent to CT imaging technology (Fig. 2); the wide range of motion of the wrist (Fig. 3). To further complicate matters, in multiple wrist

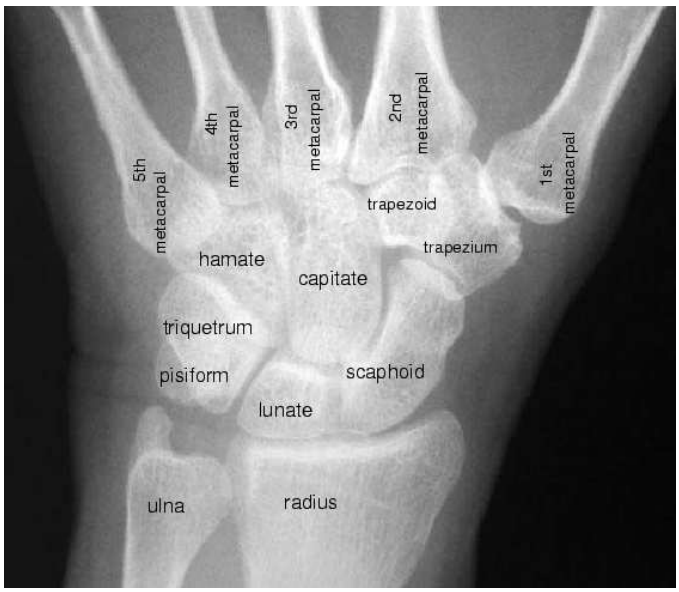


Fig. 1. The wrist comprises the distal end of the two forearm bones (radius and ulna), and eight small, tightly packed carpal bones. In this X-ray view the five metacarpals are also included. Note the complexity of the joint and also the presence of incompletely scanned bones like the forearm bones and the metacarpals.

pose studies the in-plane resolution of acquired images is often lowered from 0.2 mm to almost 1 mm in order to attain acceptable examination times per human subject.

Multiple attempts have been made to automate the wrist segmentation–registration process. However, existing automated techniques require extended tuning for each individual CT scan, as the character of the volume images changes with subject, wrist pose, relative amounts of soft and hard tissue, and imaging protocols. Thus, medical experts are often relied upon to segment and register the volume images. Although labor intensive, the manual approach is often preferred due to its unmatched accuracy and stability across subjects and wrist poses.

In this paper we present an intra-subject method for the robust automated registration of wrist bones. Our method integrates segmentation and registration; the registration accuracy is thus not limited to the accuracy of the segmentation step. One CT volume image per subject needs to be manually segmented. The segmented model is then automatically mapped to the remaining CT scans of that subject. Wrist motion information through a sequence of CT volume images is automatically extracted. When applied to sets of N wrist poses, the method reduces the user interaction time by a factor of N, while demonstrably maintaining the accuracy and stability of

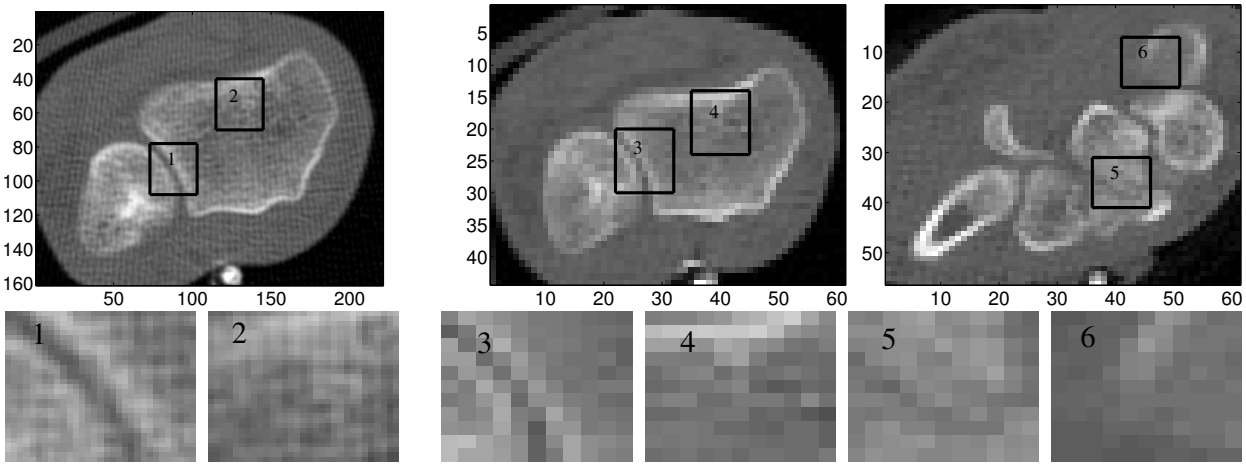


Fig. 2. The influence of acquisition resolution and partial volume effects on wrist image quality. From left to right: one slice from a high-resolution CT image (0.3 x 0.3 mm), and two slices from a low-resolution CT image (same wrist, 0.9 x 0.9 mm). Each area of interest is detailed in the bottom row. Box 1 and 3, note the extremely narrow inter-bone distances. Boxes 2 and 4, note the differences in bone tissue texture; high-resolution images reveal a distinctive texture pattern (box 2), while detail is lost in low-resolution images (box 4). Box 5, note how boundaries are blurred due to partial volume effects. Box 6, note the soft or diffused bone boundary.

the manual approach.

Lack of validation hinders the use of many registration methods in widespread clinical application. We describe our validation framework in Section IV.

A. Method Overview

Our method works in two steps. In the first step, a tissue classifier processes each CT volume image. The result is a localized distance field that will be used to track bone motion. A *localized distance field* is a scalar 3D grid that specifies at gridpoints the distance to the closest cortical bone surface. The tissue classifier takes into account the blurring or partial-volume effects inherent to CT imaging. The algorithm does so by interpolating the CT data to produce a continuous function and by considering the entire function over small regions. These regions contain more information than single point samples and give a more accurate estimate of where the cortical bone surface lies within the CT data [10].

In the second step, given a sequence of CT images processed by the tissue classifier and a geometric model of each wrist bone, we track the motion of each bone between the volume

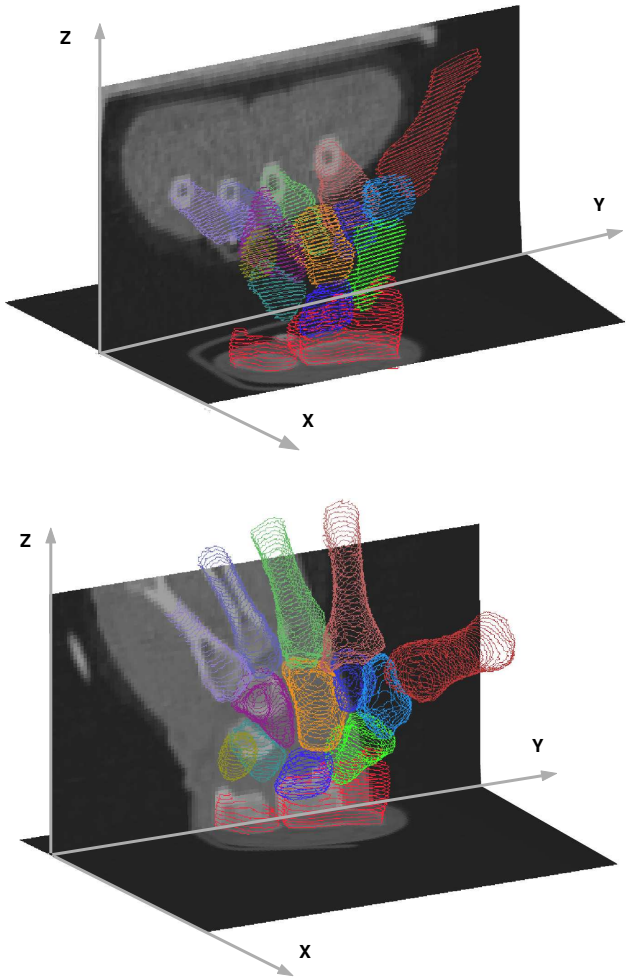


Fig. 3. The analysis of joint kinematics from CT scans requires the measurement of bone motion between different wrist poses. The images show the same geometric wrist model, after registration, in two different poses. The orthogonal greyscale planes correspond to vertical and horizontal sections through the CT volume images (darker gray areas correspond to soft-tissue, brighter areas to bones). Note the significant changes in bone posture and orientation between the two poses.

image I_M that generated the geometric model and any other volume image I_k. The distance value of each vertex in the geometric model is referenced in the processed volume corresponding to I_k. The tracking procedure automatically adjusts a bone's position and orientation in order to minimize these distances. The distance field serves as a local gradient that automatically guides the adjustment procedure.

We evaluate the accuracy of our registration method at different levels: on a synthetically

generated test-case, *in vitro* — on marked cadaver data, on an *in vivo* volunteer data set, and on several subjects participating in a wrist motion study.

The remainder of this paper is organized as follows: we discuss related work in Section II, detail our methods in Section III, present the validation framework and discuss the results obtained in Section IV; we conclude in Section V.

II. Related work

A. Wrist registration methods

In general, wrist registration methods can be regarded as either segmentation-based, or voxel property-based [9].

Segmentation-based registration aligns structures such as feature points, principal axes, moments, curves, surfaces, or volumes. Segmentation-based methods may use deformable models, in which case an extracted structure from one image is elastically deformed to fit the second image. They may also use rigid models, in which case the same structures are extracted from any images to be registered [9]. Wrist segmentation-based registration methods are particularly tedious and ineffective. Deformable-model methods [6], [7] converge poorly on wrist bone boundaries [1], leading to poor registration accuracy. Even more advanced methods like the hybrid SCDM [1] rely on unique assumptions about wrist imaging characteristics. In our experience, such methods require considerable tuning from one volume to another in order to avoid collisions between the deformable models, rendering them less reliable than manual segmentation. Current methods used in clinical applications are, to the best of our knowledge, rigid model-segmentation based, and rely on some form of manual correction [4]. This last approach is time consuming. Although the advent of new wrist segmentation techniques [3] may decrease the amount of user interaction, registration accuracy is limited to the accuracy of the segmentation step.

In contrast, voxel-property-based registration methods operate directly on the image grey values, without prior data reduction by the user or segmentation: the full image content is used, via different paradigms like neighborhood statistics, cross-correlation, Fourier analysis, optic flow, and so forth, throughout the registration process [9]. Voxel property-based methods integrate segmentation and registration and are thus likely to generate accurate results, but their use has been limited by the considerable computational costs. A voxel-property-based method successfully used for wrist registration is the technique of Snel et al. [2], who use chamfer

matching and texture characteristics to track 3D wrist motion. Given a sequence of CT volume images acquired at the same high resolution, but at different levels of radiation, and a surface model of the wrist bones, this method attempts to find the optimum location of each bone across the volume images. In order to exhibit typical bone interior texture characteristics, the images need to be acquired at a relatively high resolution (320 x 320 matrix, 0.2 mm resolution), resulting in prolonged acquisition time per human subject and increased storage requirements. Additionally, the initial bone position needs to be adjusted closely — within a 1 cm neighborhood of the solution. In this respect, the motion is built incrementally from one wrist pose to another, requiring a pre-established ordering of the CT volume images and very similar positioning of the wrist on the scan table through a sequence of scans. These requirements are not, however, readily satisfied: sequences of volume images are seldom labeled with the pose index, and CT volumes are often cropped, to reduce storage space; consequentially bones do not appear in the same area of the volume between two poses.

In contrast, our method can use relatively low resolution images of the wrist (100 x 100 matrix, 0.97 mm resolution). As a result, the total examination time per human subject decreases from hours to minutes, and storage space requirements are lowered by two thirds. This makes possible joint kinematics studies on large sets of joint poses (typically ten to twenty poses) involving large groups of subjects (typically hundreds). The volume images need not be ordered by pose.

B. Tissue Classification and Distance Fields

In its first stage our method classifies the tissues in a CT volume image in order to generate a local gradient field. We build on the work of Laidlaw et al. [10], [11], who use Bayesian probability theory to generate distance fields from medical volume images. Recovering material boundaries from sampled datasets is a research topic in several fields [12], [13], as well as generating distance fields from geometric models [14]. However, we are not aware of other work in generating distance fields directly from sampled datasets.

III. METHODS

For a given sequence $I_1, I_2, ...I_n$ of CT wrist volume images of the same subject, our method recovers the motion of each wrist bone between any two volume images I_i and I_j. The only

material	μ	σ
air	105.3	10.2
soft-tissue	1150.1	12.4
bone	1800.4	200.2

TABLE I

ESTIMATED WRIST PURE MATERIAL CHARACTERISTICS.

requirement is the existence of a geometric model of each bone, segmented from a reference volume image.

A. Geometric model of the bones

The registration method requires as input a geometric model for each bone we track. The geometric model is obtained through segmentation of one of the n volume images, and can be as simple as a collection of vertices (a cloud of 3D points). The geometric models we use were obtained through manual segmentation, via thresholding and user interaction [4].

B. Localized distance fields

In the first stage of our method, a tissue classifier processes each CT volume image in order to generate a localized distance field.

Our tissue classifier uses the partial-volume technique described by Laidlaw et al. [11]. The method identifies distances from material boundaries and creates signed distance fields for individual materials. The technique assumes that, due to partial-volume effects or blurring, voxels can contain more than one material, e.g. both cortical bone and soft tissue. Each voxel contains thus either a pure material or two pure materials separated by a boundary (Fig. 4).

The algorithm models pure material voxels with Gaussian basis functions. We estimate the mean and standard deviation defining these functions from sets of approximately one hundred voxel samples, one set per material.

The classification algorithm also models mixture voxels with basis functions (see [11] for the mixture basis function formula derivation). The mixture basis functions have an additional parameter d, describing the distance from the center of the voxel to the boundary between materials (Fig. 5).

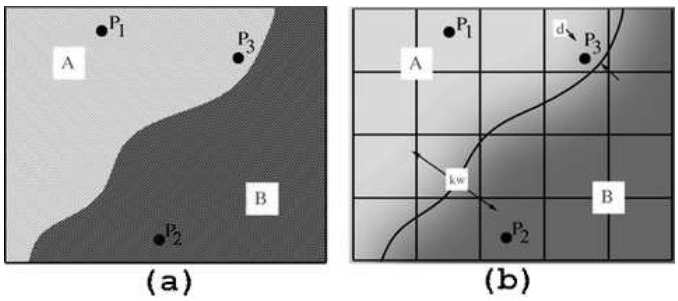


Fig. 4. The classification algorithm starts from the assumption that in a real-world object each point is exactly one material, as in (a). The measurement process creates samples that combine measurements of different materials, as in (b). From the samples we reconstruct a continuous, band-limited measurement function p(x). Points P_1 and P_2 lie inside regions of a single material. Point P_3 lies near a boundary between materials. The classification algorithm computes P_3's distance to the material boundary. In (b) P_3 is shown with a parameter d that indicates the distance from the center of the surrounding voxel to the boundary. The grid lines show voxel boundaries and how they relate to the regions.

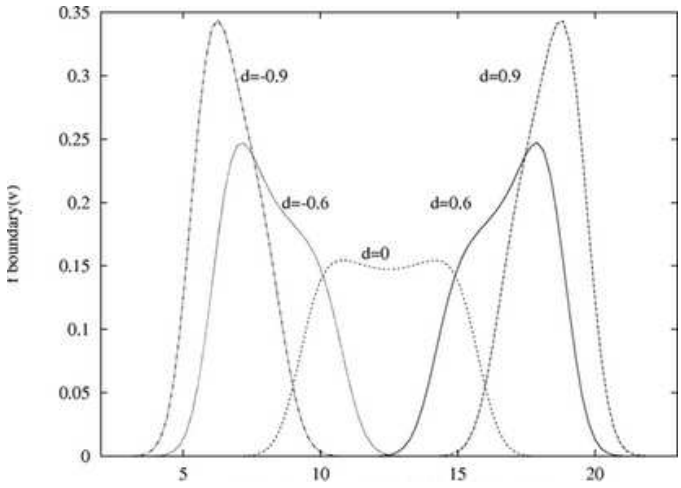


Fig. 5. The shapes of mixture basis functions, f boundary(v), are different for different values of d, the distance from the boundary to a voxel center. Note that the shapes approach normal distributions as d moves away from zero.

During classification, the basis functions are fit to each voxel neighborhood, and the most likely — in the Bayesian sense — basis is chosen probabilistically [10]. Bayesian theory is also used to estimate parameters describing the mixture basis functions, and in particular the distance from each mixture voxel to the nearest boundary between materials. The result of the classification is a localized distance field — a scalar 3D grid that specifies at gridpoints the distance to the closest boundary between two materials. The distance field is local in the sense

that the distance estimate is specified only at gridpoints located within about 5 voxels in each direction of the material boundary.

In the wrist case, we consider three distinct pure materials: air, soft-tissue, and bone. Soft-tissue is present both outside bones and inside bones (as bone marrow). All material samples were collected from the same *in vivo* dataset, yielding the parameter values shown in Table I. We consider two mixture basis function: one modeling mixtures along air and soft-tissue boundaries, the other modeling mixtures along soft-tissue and bone boundaries.

In our experiments this three material, two mixture tissue description was general enough to hold across different scans and different subjects. Accordingly, we initialized the basis function parameters to the same values throughout all the datasets, *in vitro* and synthetic dataset included.

Because it allows for mixtures of materials and treats voxels as regions, the Bayesian tissue classifier reduces errors that other classification methods can create along boundaries between materials [11].

The classification of a wrist volume images produces one distance field per boundary type. We use the distance field corresponding to mixtures of bone and soft-tissue (Fig. 6) in the tracking stage of our registration method.

C. Tracking

In the second stage of our method we track the motion of each bone through a sequence of CT volume images classified as described in section III.B. We register intra-patient data, therefore we need to recover the rigid body transform (i.e., rotation around its center of mass, and translation) per bone, between any two volume images I_i and I_j. In practice, it suffices to track the motion M_{Mk} between the reference neutral image I_M that generated the geometrical model and any other image I_k. With this information, the relative motion M_{ij} between any two volume images I_i and I_j can be computed as:

$$M_{ij} = {M_{Mi}}^{-1} \cdot M_{Mj}$$

The tracking procedure searches for the position and orientation of each bone that minimizes a cost function. The cost function evaluates the squared difference between the current distance field value of each vertex in the geometric model and the expected value of that vertex in the geometric model distance field. Let $D_k(p)$ denote the value of the distance field generated from

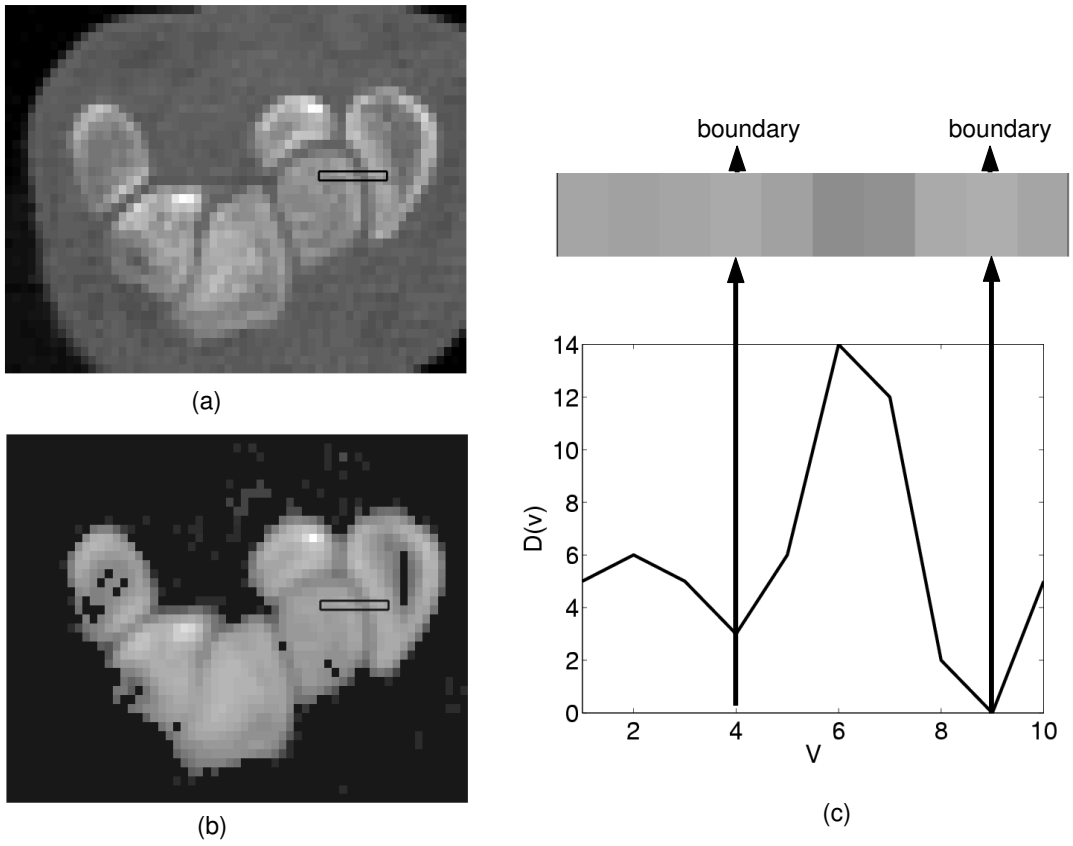
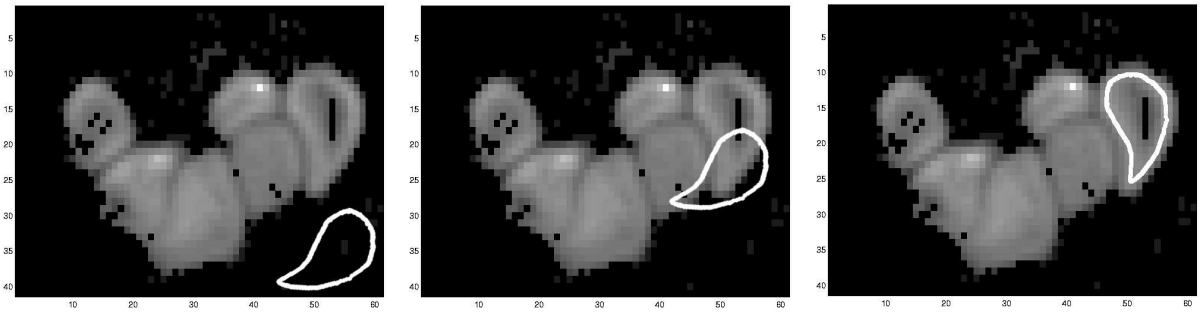


Fig. 6. The distance-to-bone field image quantifies the distance from the center of each voxel to the closest boundary. (a) One slice from a low-resolution (0.9 x 0.9 mm) wrist CT volume image. (b) Localized distance field corresponding to bone–soft tissue boundaries. Dark pixels have been classified as either pure soft-tissue, pure air, or soft-tissue and air mixture. The area of interest in the box crosses two bony boundaries and is detailed on the right. Each voxel in the field codifies the distance from the voxel center to the closest bony boundary (the lighter the gray, the closer to a bone boundary the voxel is). (c) Plot of the distance values along the strip on top. Three distance units correspond approximately to the size of a voxel. Note the two dips in the plot corresponding to the two bone boundaries. In this particular case the bone cortex is very thin (1 voxel wide); consequentially there are no samples inside the bone cortex, hence the distance function D(v) does not take negative values.

the k-th CT volume image at the 3D point p. Let $D_M(p)$ denote the value of the distance field generated from the geometric model volume image at the 3D point p. Let $\mathrm{R} \cdot \mathrm{T} \cdot \mathrm{v}$ denote the 3D point obtained by applying the rotation R and translation T to the vertex v. The cost function is then:

$$\mathrm{F} = \frac{\Sigma_{j=1}^{m} (D_k(\mathrm{R} \cdot \mathrm{T} \cdot \mathrm{v_j}) - D_M(\mathrm{v_j}))^2}{V}$$



(a) High F value (b) Lower F value (c) Zero F value

Fig. 7. 2D illustration of the tracking procedure. In this example we search for the optimum location of the 2D boundary of a bone (shown in white) using a 2D bone and soft-tissue distance field (shown in gray). (a) In a highly unlikely neighborhood the cost function F has a high value; the bone boundary may become trapped in local minima. (b) In the neighborhood of the solution the cost function F has a lower value, as some boundary points overlap with lower distance field values. The distance field serves as a local gradient: F decreases smoothly as the location and orientation of the white boundary approaches the correct solution. (c) At the correct location and orientation the cost function F should be close to zero.

where $\mathrm{v}_{1:\mathrm{m}}$ are vertices in the geometric model, and V is the number of points that are still inside the CT volume after rotation and translation. Incorporating in the cost function knowledge about the expected distance values compensates for the slight errors in boundary-point location that occur occasionally during manual segmentation of the geometric model. The V argument ensures that partially scanned bones like the radius and ulna or the metacarpals are not too heavily penalized for partial presence in the CT volume. It is often the case that such bones are scanned up to different lengths from one scan to another. The $D_k(p_i)$ function returns a penalty value proportional with the distance to the volume image, whenever p_i is outside the CT volume. This formulation allows partially scanned bones to find their optimal location inside the CT volume.

We pose the registration process as an optimization problem, where we find for each bone the rigid body transform (R,T) that minimizes the cost function F. We use a quasi-Newton algorithm to solve the optimization problem [8]. The distance volume serves as a smooth local gradient field, which leads to rapid convergence when the search starts from a reasonable point, i.e. at least a few geometric model vertices are within the capture zone of the localized distance field (Fig. 7). In practice, we apply to all the bones a rough wrist alignment translation M_{com} based

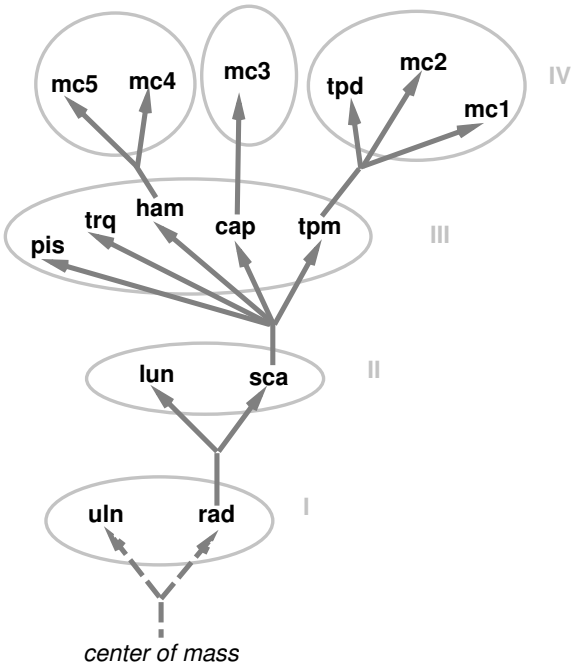


Fig. 8. Wrist motion-governing hierarchy. We consider four layers in ascending order from the forearm. During a propagation step the motion of a bone b_i is propagated to all bones in ascending levels that have b_i as an ancestor. The hierarchy indicates the radius (rad) and scaphoid (sca) may be governing the motion of the other bones.

on the center of mass of the first five slices of the forearm bones (radius and ulna) and the center of mass of the bony points in the first five slices of the distance field. This gross approximation suffices as a search start point.

The tracking procedure is applied hierarchically, ensuring that all bones start from a viable start position. We run the optimization procedure on successive layers of the wrist bones, starting with the forearm bones, as shown in Fig. 8. We iterate through bones: once we detect the motion of bone b_i through cost function optimization (*optimization* transform), we propagate the motion to all the bones that have b_i as an ancestor in the tree hierarchy (*propagation* transform), then we move on to the next bone. Optimization and propagation transforms are accumulated for each bone.

The hierarchy was empirically derived based on a trial-and-error analysis of the start values of the cost function F on three separate *in vivo* sequences of volume images. All possible hierarchies starting from the radius and ulna were considered; we chose the one which generated best start values of the cost function across all sequences. The hierarchical approach ensures that we

always start an optimization step from a reasonable neighborhood, and thus speeds the search process considerably.

To further reduce the size of the search space, once the optimized location of a bone is identified, we remove the boundary of the bone from the volume image. The "removal" procedure wipes out the boundary by marking as non-bone material the volume values corresponding to the rigidly transformed coordinates of the points p_i. The areas in the volume image already occupied by tracked bones become in this manner unappealing to the search algorithm, leading to faster convergence.

The quasi-Newton method is fast and robust, but, like most optimization procedures, is susceptible to getting trapped in sub-optimal local solutions. Consequentially, we use sixty-four perturbed start positions for each bone and choose the solution that yields the least value of the error function. The multiple searches per bone are easily parallelized. We found that the optimization procedure was stable with respect to perturbations in the space of possible rotations. The perturbed start positions are therefore generated by sampling the space of possible initial translations on three concentric spheres of radius 2, 4, and 8 voxels respectively. In our experience, the majority of the repeated optimizations per bone return the same minimum. The alternative local minima are at least one order of magnitude higher (expressed in squared millimeters).

IV. Validation, Results, and Discussion

We evaluate the accuracy of our registration method at different levels: on a synthetically generated test-case; on *in-vivo* data; *in vitro* — on marked cadaver data; and, finally, on subjects participating in a wrist kinematics study.

The synthetic test case allows us to estimate the error introduced by the tissue classifier and the optimization procedure solely, given a mathematically perfect geometric model. The *in-vivo* test case — part of a repeatability study — measures the similarity between our results and the results obtained by expert manual segmenters. The similarity of the results to the manual registration results does not necessarily guarantee that our registration method is accurate. Human experts *can* be consistently wrong. In this respect, we validate the accuracy of our method against marked cadaver data, in the *in vitro* dataset study. Finally, we present results obtained in a clinical application. All the datasets used to generate the results presented in this paper — except for

the synthetic test case — were acquired using CT technology (scan parameters: 80kV, 80mA). Cortical bone and, in the *in vitro* case, marker contours were extracted from the CT images using thresholding and user interaction [4]. The synthetic contours were mathematically generated.

A. Synthetic dataset results

The synthetic dataset consists of a mathematical model of a 50 x 60 x 50 mm upright parallelepiped with the same material properties as cortical bone, and a computer-generated CT volume image of a copy of the parallelepiped, rotated around the Y axis by 0.4 radians, and subsequently translated by T(127.7783, 132.8651, 122.7857)mm. We used the automated registration procedure to align the original parallelepiped model with the translated and rotated parallelepiped.

In the total absence of image noise, and with a perfect geometric model of the parallelepiped, the registration procedure recovered a global rotation R(-1.18628e-06, 0.399996, 5.73388e-05) (expressed in radians as Euler rotation angles around the X, Y, and Z axis), and a global translation T(127.7790, 132.8662, 122.7860)mm. The synthetic test case serves as an estimate of the error introduced solely by the tissue classifier and the optimization procedure.

B. In vitro study results

The *in vitro* dataset consists of four high resolution (0.125 x 0.125 x 1 mm) CT images of a fixed specimen (separated forearm and hand) in different positions. Each component was encased in plastic resin to prevent relative motion between any bones. Seven ceramic spheres of various high tolerance diameters were rigidly glued to each specimen component. The centroids of the seven spherical markers allow for the calculation of rigid body motions by a method of least-squares and were used as the registration "gold" standard. Bone geometric models and sphere contours were extracted from each volume image.

We used the automated registration procedure to recover the individual bone transforms that mapped one set of bone geometric models to each specimen volume image. We subsequently computed the relative motion between each pair of volume images. We report the global transform results in *helical axis of motion* (HAM) coordinates [5]. HAM coordinates express bone motion as pairs (θ, t) of rotations around and translations along a unique helical axis.

Bone	$\Delta\theta$ (deg)	Δt (mm)
rad	0.06 $\pm$ 0.03	0.16 $\pm$ 0.22
uln	0.05 $\pm$ 0.04	0.40 $\pm$ 0.33
sca	0.17 $\pm$ 0.10	0.23 $\pm$ 0.13
lun	0.10 $\pm$ 0.11	0.50 $\pm$ 0.43
trq	0.24 $\pm$ 0.09	0.10 $\pm$ 0.10
pis	0.40 $\pm$ 0.31	1.24 $\pm$ 1.01
tpd	0.50 $\pm$ 0.50	0.26 $\pm$ 0.23
tpm	0.43 $\pm$ 0.32	0.49 $\pm$ 0.32
cap	0.06 $\pm$ 0.03	0.11 $\pm$ 0.07
ham	0.12 $\pm$ 0.05	0.07 $\pm$ 0.05

TABLE II

DIFFERENCES BETWEEN AUTOMATED REGISTRATION RESULTS AND THE GOLD STANDARD FOR THE IN VITRO DATASET (STATISTICS OBTAINED BY PAIRWISE REGISTRATION OF SIX IMAGES). THE ROTATIONAL COMPONENT AND THE TRANSLATIONAL COMPONENT COMPUTED BY THE AUTOMATED REGISTRATION METHOD ARE LESS THAN 1 DEG AND 1 MM AWAY FROM THE GOLD REGISTRATION RESULTS.

Table II shows results obtained by registering each pair of volume images. Bone rotation errors are less than 1 deg. Bone translation errors are generally less than 1 mm. Values for the pisiform are somewhat larger, due to the almost spherical shape of the bone. Nevertheless, errors for other bones sensitive to inertia based registration — like the trapezoid, who have a symmetrical shape about the major axes — are still within the 1 deg, and 1 mm variation. This comes as an additional benefit of combining a surface-based registration method with a tissue classifier that often preserves more detail from the scan than expert segmenters. Also, our registration method appears stable with respect to incompletely scanned bones.

C. In vivo repeatability study results

The *in vivo* dataset for the "repeatability" study comprises two CT images of the same left wrist, one high resolution (0.273 x 0.273 x 1.0 mm) with the wrist in neutral position, and one low resolution (0.781 x 0.781 x 1.0 mm) with the wrist extended. Five medical school students manually segmented the wrist from the two volume images. The task was considerably more difficult in the case of the low resolution volume image, due to increased blurring and less

Bone	Manual mean±std. dev. θ (deg)	Automated θ (deg)	Manual mean±std. dev. t (mm)	Automated t (mm)
uln	8.78 ± 9.03	9.54	0.78 ± 1.14	0.33
sca	42.96 ± 2.93	43.91	-0.21 ± 0.13	-0.14
lun	29.85 ± 0.75	29.89	0.20 ± 0.16	0.27
trq	37.62 ± 2.79	38.01	0.39 ± 0.11	0.34
pis	71.88 ± 8.95	63.88	-2.02 ± 2.25	-1.51
tpd	66.72 ± 1.80	67.38	0.68 ± 1.00	-0.33
tpm	69.95 ± 2.16	67.93	-0.86 ± 0.81	-0.48
cap	71.77 ± 1.54	72.15	-0.27 ± 0.25	-0.20
ham	68.61 ± 1.05	67.97	-0.27 ± 0.21	-0.13

TABLE III

COMPARISON OF AUTOMATED REGISTRATION RESULTS AND MANUAL REGISTRATION RESULTS FOR THE REPEATABILITY *in vivo* DATASET. ALL ROTATIONAL AND TRANSLATIONAL COMPONENTS COMPUTED BY THE AUTOMATED REGISTRATION METHOD ARE WITHIN ONE STANDARD DEVIATION OF THE MANUAL REGISTRATION RESULTS.

detail in the scan. Subsequently, the motion of the wrist between the high resolution and the low resolution volume image was recovered with an inertia-matching method. Each segmenter took several runs through the procedure, for a total of twelve runs. Statistics on the results were collected. Table III summarizes our results and compares them with the manual-segmenter results. The resultant 3D motion is relative to the radius and is reported using HAM variables.

Note that the standard deviation of the manual translational component is higher in the case of certain bones like the ulna (incompletely scanned), or the trapezoid (symmetric shape with respect to a principal axis), reflecting shortcomings of the manual registration method.

The automated registration results are consistently within the variation of the manual results. HAM coordinates are extremely sensitive to slight changes in Eulerian angle rotations; differences between manual segmenter results, and indeed even the highest HAM difference between the automated registration and the expert segmenter results reflect insignificant perturbations in geometric model location and orientation (Fig. 9).

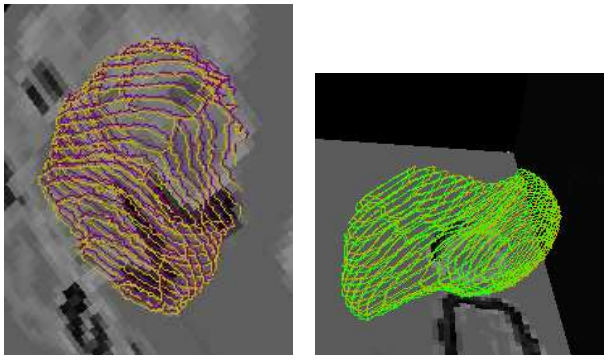


Fig. 9. Visual comparison of automated registration results against the furthest off manual registration results. Geometric models transformed by the automatically computed rigid motion are shown in purple and green. The same models transformed by the manually computed rigid motion are shown in yellow. Note how similar the resulted positions and orientations are.

D. In vivo intra-subject results

CT scans of both wrists were acquired from forty subjects. A high-resolution volume image (resolution 0.3125 x 0.3125 x 1.25 mm) with both wrists in neutral position was acquired for each subject. Scans of the wrists in various positions, spanning a full range of wrist motion, were acquired at a lower resolution (0.9375 x 0.9375 x 1.25 mm), in order to decrease the total examination time. Ten to twelve volume images were acquired per subject; the right and left wrists were subsequently split into separate volumes, for a total of nine hundred wrist volume images. Geometric models of the bones were extracted for each subject, then automatically registered across all wrist images of that subject.

Registration results were validated both visually and numerically. Visual validation was performed by superimposing the registered bone geometric models with vertical and horizontal slices of the volume image. Two sliders controlled the vertical and horizontal slice displayed. Numerical validation was performed by examination of the final-fit cost function values. In less than 1% of all the CT volume images the automated registration procedure failed to register correctly either the lunate or the trapezoid. Even in these cases all the other bones were registered correctly. These cases were associated with unusually sparse geometric models in conjunction with wrist poses in which significant parts of the surfaces of two different bones became parallel and very close. For example, lunate registration failed when sparse radius heads became parallel to one lunate side; similarly, trapezoid registration failed when sparse second metacarpal bottoms

became parallel to a trapezoid side. In such cases, boundary neighborhoods were too different between the neutral volume and the volume to register. Denser geometric models might have helped.

E. Discussion

As shown by our synthetic test-case experiment, in the absence of scanner noise and geometric model segmentation errors, our method introduces negligible errors: less than 10^{-3}deg rotational error, and less than 10^{-3}mm translational error.

As shown in the *in vitro* experiment, rotation errors are less than 1 deg, and bone translation errors are generally less than 1 mm even in the presence of noise and errors due to segmentation of the geometric bone model. An earlier accuracy study performed by Neu et al. [4] using the same *in vitro* specimen compared the performance of an inertia-matching registration technique against the performance of two surface matching techniques. The study found rotation and translation errors generally less than 2 deg and 1 mm, respectively. Values for certain bones like the trapezium, trapezoid, and pisiform were significantly higher. The inertia-based registration was not valid on incompletely scanned bones like the radius and ulna. These results indicate that our automated method is more accurate than earlier manually-aided segmentation-based registration methods.

We believe our method performs, in fact, better on *in vivo* than on *in vitro* data, although the absence of an *in vivo* gold standard makes this difficult to demonstrate. The tissue classifier misclassified several *in vitro* data voxels — as the material composition of a fixed specimen encased in resin, with all soft tissue removed, was rather different from the tissue model assumed by our classifier — based on an *in vivo* specimen. It is reasonable to assume that errors in the classification process reflected negatively on the registration results.

The *in vivo* repeatability study shows that the automated registration method attains results within manual registration variation, while reducing the amount of user interaction by a factor of N when applied to sets of N volume images.

Finally, visual and numerical validation during clinical application (*in vivo* intra-subject results) indicates the method is robust across different subjects and wrist poses.

The motion hierarchy we employ to speed up the registration process is an interesting topic in itself, as the principles governing wrist motion are currently little understood. The hierarchy

remains to be validated by further bioengineering studies.

Our C++ implementation of the registration method is parallelized and runs, off-line, under both Linux and Solaris. It takes about one and a half hours on a Linux i686 PC to register a wrist model with a low-resolution volume image. The total running time drops accordingly when the program is run on a multiprocessor machine. Approximately 75% of the running time is spent in the tissue classification phase. No user interaction beyond the initial segmentation and final visual validation is required.

V. Conclusion

We presented in this paper an automated method for intra-patient registration of wrist bones from computed tomography volume images. The images may be acquired at relatively low resolution, with a consequent reduction in the total radiation exposure time per human subject and on general scan storage requirements. Our approach reveals an unsuspected motion-governing bone hierarchy that may lead to better understanding of wrist kinematics.

In general, registration methods have a trade-off between human interaction, speed, accuracy, and robustness. While many automated registration methods offer increased speed and reduced amounts of user interaction, most are barred from meaningful clinical application by insufficient validation of their accuracy and stability across different datasets. Our validation results show that the method presented preserves the accuracy and stability of manual registration methods, while dramatically reducing user interaction time. In fact, the reported *in vitro* study results indicate improved accuracy over standard segmentation-based registration methods.

The registration method presented makes possible the automated extraction of joint kinematics information from sets of volume images of wrists in many different poses. The method may also be applicable to bones of other joints.

Acknowledgments

The authors would like to thank James Coburn, Sharon Sonnenblum, and Anwar Upal (Brown University Medical School) for feedback during the validation phase, Douglas Moore (Brown University Medical School) for improving our understanding of tissue structure, and Dr. Michael Richardson (University of Washington) for permission to reproduce Fig. 1. This work was supported in part by NSF CCR-0093238 and NIH AR44005.

REFERENCES

[1] Sebastian, T., Tek, H., Wolfe, S.W., Crisco, J.J., Kimia, B., Segmentation of Carpal Bones from 3D CT Images using Skeletally Coupled Deformable Models, MICCAI'98 Lecture Notes in Computer Science, 1496, pp. 1184–1194, Springer-Verlag, New York, 1998.

[2] Snel, J.G., Venema, H.K., Moojen, T.M., Ritt, M., Grimbergen, C.A. and den Heeten, G.J., Quantitative in vivo Analysis of the Kinematics of Carpal Bones from 3D CT Images Using a Deformable Surface Model and a 3D Matching Technique, Med. Phys., 27(9), pp. 2037–2047, 2000.

[3] Hahn, H., and Peitgen, H., IWT - Interactive Watershed Transform: A hierarchical method for efficient interactive and automated segmentation of multidimensional grayscale images, Proc. Medical Imaging, SPIE 5032, San Diego, Feb. 2003, in press.

[4] Neu, C.P., McGovern, R.D., Crisco, J.J., Kinematic Accuracy of Three Surface Registration Methods in a Three-Dimensional Wrist Bone Study, J.Biomech.Eng., 122(10), pp. 528–533, 2000.

[5] Panjabi, M.M., Krag, M.H., Goel, V.K., A Technique for Measurement and Description of Three-Dimensional Six Degree-of-Freedom Motion of a Body Joint With an Application to the Human Spine, J.Biomech.Eng., 14, pp. 447–460, 1981.

[6] Kass, M., Witkin, A., Terzopoulos, D., Snakes: activecontour models, Intern. J. of Computer Vision, 1, pp. 321–332, 1988.

[7] Cohen, L.D., and Cohen, I., Finite-element methods for active contour models and balloons for 2-d and 3-d images, IEEE Trans. Pattern Analysis and Machine Intelligence, 15(11), pp. 1131–1147, 1993.

[8] *NAG Fortran Library Routine Document*, http://www.nag.co.uk/numeric/fl/manual/html/e04jyf.html.

[9] Maintz, J.B.A., and Viergever, M.A., A survey of medical image registration, J. Medical Image Analysis, 2(1), pp. 1–36, 1998.

[10] Laidlaw, D.H., Fleischer, K.W., Barr, A.H., Partial-volume Bayesian Classification of Material mixtures in MR volume data using voxel histograms, IEEE Trans. Med. Imag., 17(1), 74-86, 1998.

[11] Laidlaw, D.H., Geometric Model Extraction from Magnetic Resonance Volume Data, PhD thesis , Caltech, 1995.

[12] Bonnell, K.S., Duchaineau, M.A, Schikore, D.A., Hamann, B., Joy, K.I., Material Interface Reconstruction, IEEE Transactions on Visualization and Computer Graphics, 9(4), pp. 500–511, 2003.

[13] Whitaker, R., Breen, D, Museth, K., and Soni, N., Segmentation of biological volume datasets using a level set framework, Volume Graphics 2001, pp. 249-263, Springer, Vienna, 2001.

[14] Mauch, S., A Fast Algorithm for Computing the Closest Point and Distance Transform, submitted to SIAM Journal on Scientific Computing.