Super-Resolution Registration Using Tissue-Classified Distance Fields

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

Abstract—We present a method for registering the position and orientation of bones across multiple computed-tomography (CT) volumes of the same subject. The method is subvoxel accurate, can operate on multiple bones within a set of volumes, and registers bones that have features commensurate in size to the voxel dimension. First, a geometric object model is extracted from a reference volume image. We use then unsupervised tissue classification to generate from each volume to be registered a super-resolution distance field—a scalar field that specifies, at each point, the signed distance from the point to a material boundary. The distance fields and the geometric bone model are finally used to register an object through the sequence of CT images. In the case of multiobject structures, we infer a motion-directed hierarchy from the distance-field information that allows us to register objects that are not within each other's capture region.

We describe a validation framework and evaluate the new technique in contrast with grey-value registration. Results on human wrist data show average accuracy improvements of 74% over grey-value registration. The method is of interest to any intrasubject, same-modality registration applications where subvoxel accuracy is desired.

Index Terms—Distance fields, joint kinematics, medical imaging, registration, tissue classification, wrist.

I. INTRODUCTION

AS RESEARCH areas that employ image registration techniques focus on ever-smaller features, they require higher registration accuracy. *In vivo* kinematic analysis of small joints, such as the wrist, exemplifies the need for highly accurate intrasubject, same-modality registration. A common way to analyze joint kinematics is by CT-imaging the joint bones in several different positions and registering them across all volume images. While early studies have focused on retrieving bone pose and orientation, recent research focuses on measuring how more subtle features like interbone spacing change with motion. In the first case, errors on par with the image sampling-step size, like those introduced by existing tracking systems, may be acceptable, while in the latter case errors as small as 0.5 mm can compromise the study by introducing interbone collisions. At the same time,

*G. E. Marai is with the Department of Computer Science, Brown University, Providence, RI 02912 USA (e-mail: gem@cs.brown.edu).

D. H. Laidlaw is with the Department of Computer Science, Brown University, Providence, RI 02912 USA.

J. J. Crisco is with the Department of Orthopaedics and the Division of Engineering, Brown Medical School/Rhode Island Hospital, Providence, RI 02912 USA.
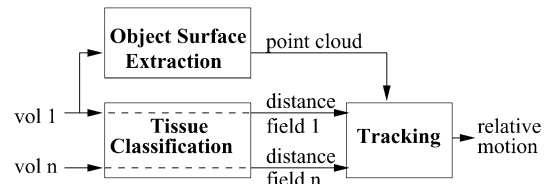
Fig. 1. Our registration method works in three steps: to register an object across two volume images, we first extract the object surface from one volume; we generate through tissue classification a localized distance field from each volume; we then use the object surface and the distance fields to track the object.

decreasing the image sampling-step results in increased imaging cost and time. We need a subvoxel-accurate method for registering features whose size is on par with the image sampling step.

We describe in this paper an automated intrasubject same-modality registration method that attains subvoxel-accuracy. The method is of interest to any registration applications involving datasets where the image sampling step is larger than features of interest.

II. REGISTRATION METHOD

A. Overview

The registration method works in three steps on a series of volume images. First, we extract the surface of the object to be registered from an arbitrarily selected reference image. Next, in order to obtain an accurate localized distance field for registering the object, we classify the tissues in each volume image using a probabilistic approach. Last, we register the object by automatically adjusting its position and orientation, thereby minimizing a distance-field derived cost function (Fig. 1).

In the case of multiobject structures (e.g., joints in the human body) we infer from the distance field an object hierarchy that expands the capture range of our procedure beyond the capabilities of previous registration methods. The capture range represents the range of positions from which a registration algorithm can converge to the correct minimum or maximum.

We validate our method using CT data from a cadaver with external markers, an *in vivo* volunteer, and forty subjects participating in a wrist-motion study. We compare the performance of our method against a manually aided segmentation-based method as well as a standard grey-value-registration method.

B. Object Surface Extraction

Through manual segmentation, thresholding, and user interaction, we extract in this first step an object surface from a reference CT volume image [1]. Summarizing this reference, the contours defining the outer cortical bone surfaces of each object are extracted using thresholding and image algebra processes
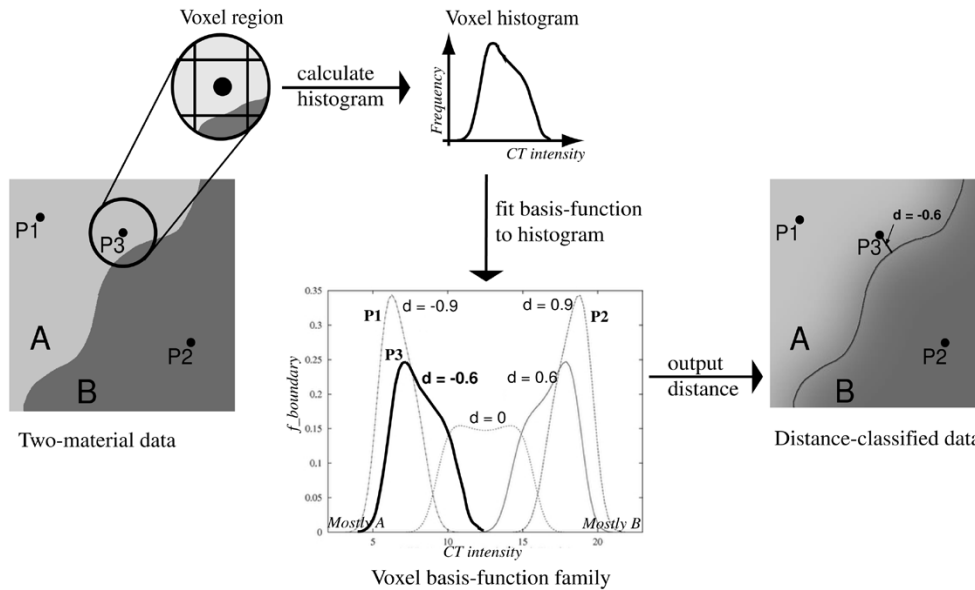
Fig. 2.   The classification algorithm computes distances from sample points to material boundaries. Points P_1 and P_2 lie inside regions of a single material, either A or B. Point P_3 lies near the boundary between A and B. We treat each voxel as a region, by subdividing it into 8 subvoxels, and taking into account information from neighboring voxels. We evaluate the image intensity and its derivative at the center of each subvoxel. The resulting voxel histogram is then fit to a family of basis functions (*f_boundary*), whose shapes reflect the estimated distance d to the material boundary. The best-fit basis instance is selected through a maximum likelihood process ($d = -0.6$ fits best the histogram of point P_3). The result is a localized distance field, that specifies, at each point, the signed distance from the point to the material boundary.

with a three-dimensional (3-D) imaging software package (Analyze AVW 2.5; Biomedical Imaging Resource, Mayo Foundation, Rochester, MN). Each contour is then assigned to the appropriate object using Matlab custom code, which designates contours based on the contiguity of their centroids. Contour lines are output as collections of discrete points, which are distributed densely in the contour plane and sparsely between different contours.

C. Localized Distance Fields

In the second step of our method, we classify the tissues in each CT volume image probabilistically in order to generate a localized distance field. Our tissue classifier uses the partial-volume technique described by Laidlaw *et al.* [3]. This method identifies distances from material boundaries and creates 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 is assumed to contain either a pure material or two pure materials separated by a boundary (Fig. 2).

We treat each voxel as a region, by subdividing it into 8 subvoxels, and evaluating the image intensity and its derivative at the center of each subvoxel. The intensity is interpolated from the discrete data using a tricubic B-spline basis that approximates a Gaussian. Thus, intensity and derivative evaluations can be made not only at sample locations, but anywhere between samples as well. From this intensity and derivative information we infer a histogram of each voxel, accumulating the contributions from all subvoxels. This gives us a more refined histogram than we would obtain by evaluating only the intensity values at the same number of points. Histograms are next fit by basis functions, each basis function corresponding to either one material or a mixture of two materials.

Pure material basis-functions are Gaussians whose parameters are the mean CT grey-scale value and standard deviation for that material. Mixture basis functions have an additional parameter d, describing the distance from the center of the voxel to the boundary between materials. As the distance parameter changes, the shape of the basis function changes (Fig. 2). The basis-function shape that best fits each mixture voxel histogram is chosen through a maximum likelihood process. The derivation of the basis-function formulas and the description of the optimization process are presented in detail in [3]. We repeat the fitting procedure for each material, and select the material basis function that fits each voxel histogram best.

For each tissue type, the sole input required by our tissue classifier is an initial estimate of its CT grey-scale value's mean and standard deviation. We estimate these measures from sets of approximately one hundred voxel samples, one set per tissue type. We consider three distinct pure materials: air, soft-tissue, and bone. Soft-tissue is present both outside bones and inside bones (as bone marrow). Material samples are collected only once, from the same *in vivo* dataset. We consider two instances of the mixture basis function: one modeling mixtures along air and soft-tissue boundaries, the other modeling mixtures along soft-tissue and bone boundaries. We initialize the basis function parameters to the same values throughout all the datasets, including the *in vitro* datasets.

Through this basis-function tissue-classification process, we generate a localized distance field. The distance field is a scalar 3-D 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 as far as gridpoints located within a five-voxel band around the material boundary. Distances between gridpoints are approximated through tricubic interpolation.

a)   b)   c)
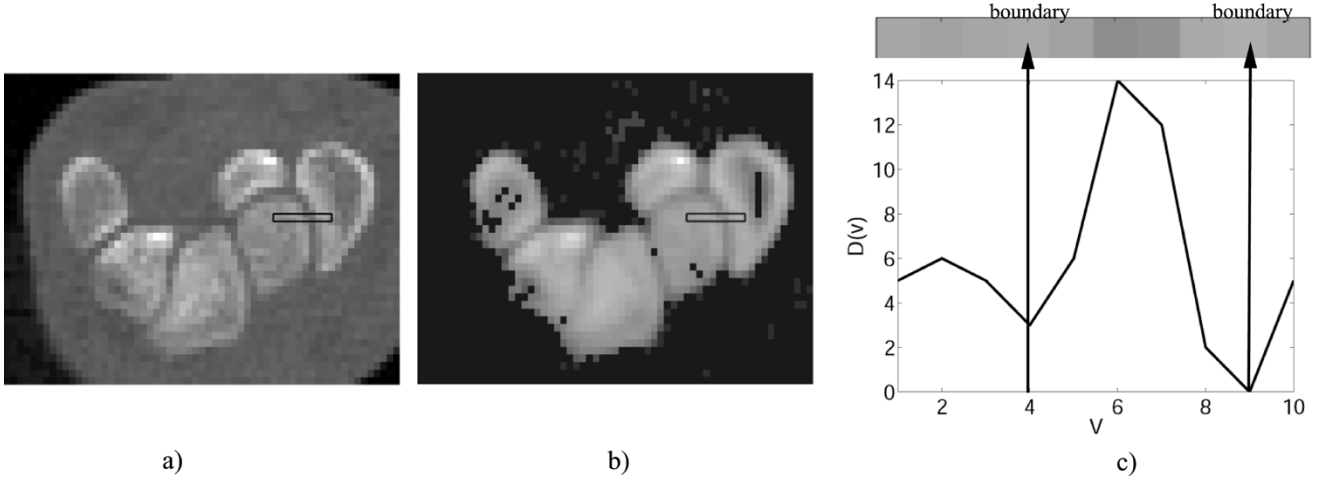
Fig. 3.   Tissue-classified distance-fields quantify the distance from the center of each voxel to the closest boundary. (a) One slice from a low-resolution (0.9×0.9 mm) wrist CT volume image. (b) Localized distance field corresponding to bone material. Dark pixels have been classified as either pure 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 grey, the closer to a bone boundary the voxel is. (c) Plot of the distance values along the strip on top. 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 classification of a wrist volume image produces one distance field per material type. We use the distance field corresponding to bone material (Fig. 3) in the tracking stage of our registration method.

D. Tracking Procedure

In the third step of our method, we register an object through a sequence of CT volume images classified using the process described in Section II-C. For each bone, we recover the rigid body transformation between the reference image that generated the geometrical model and a target image. The rigid body transform is expressed as a rotation around the bone's center of mass, and translation.

An object's geometric model is registered with a target image of the object when its signature in the reference distance field, D_R, is most similar with its signature in the target distance field, D_T. We measure this similarity with a sum-of-squared-differences cost function that takes into account the reference and target distance-field values of the vertices in the geometric model. The sum is weighed by the number of vertices that are still inside the target distance field after applying the current transform to the model. The cost function is thus

$$F = \frac{1}{V}\Sigma_{j=1}^{n}(D_R(p_j) - D_T(p'_j))^2$$

where p_j are points in the geometric model, p'_j are the 3-D points obtained by applying the current translation and rotation to p_j, n is the number of points in the geometric model, and V is the number of points that are still inside D_T after rotation and translation. Whenever p'_j is outside D_T, $D_T(p'_j)$ returns an approximation of the distance from p'_j to the D_T volume, obtained by projecting p'_j on the closest face of the volume. This expands the cost-function gradient outside the volume to register, in order to accommodate partially scanned bones. Note that, by incorporating D_R in the cost function, we compensate for the

Fig. 4.   Two-dimensional illustration of the tracking procedure. In this example, we search for the optimal location of the 2-D boundary of a bone (shown in white) using a 2-D bone and soft-tissue distance field (shown in grey). (Left) In a highly unlikely neighborhood, the cost function F has a high value; the bone boundary may become trapped in local minima. (Center) 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. (Right) At the correct location and orientation the cost function F should be close to zero.

small errors in boundary-point location that occur occasionally during segmentation of the geometric model. Because this cost function attempts to match distance-field signatures, geometric model vertices that diverge slightly from the true bone boundary due to segmentation-errors will be off by the same amount in the registered image.

Our tracking procedure searches for the position and orientation of each bone that results in maximal distance-field similarity at registration, i.e., the rotation and translation that minimizes F (Fig. 4). We use a quasi-Newton algorithm to solve the optimization problem [4]. The distance volume serves as a smooth local gradient field, which leads to rapid convergence when the search starts from a point where at least a few geometric model vertices are within the capture region of the localized distance field. In practice, we begin by applying to all the bones a rough alignment translation M_{com}. The alignment translation is based on the center of mass of the bottom five slices of a joint and the center of mass of the bony points in the first five slices of the distance field. This approximation suffices as a search start point.

Fig. 5. The human wrist is a complex structure comprising the distal end of the two forearm bones, and eight small, tightly packed carpal bones. In this X-ray view, the five metacarpals are also included. Figure reproduced with permission from [5].

Fig. 6. Wrist hierarchy induced from distance field information. We consider four layers in ascending order from the forearm: *uln*a and *rad*ius; *lun*ate and *sca*phoid; *pis*iform, *triq*uetrum, *ham*ate, *cap*itate, and *trapezium*; *m*etacarpals and *trapezoid*. 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 *rad*ius and *sca*phoid may be governing the motion of the other bones.

The quasi-Newton method is fast and robust; however, like most optimization procedures, it is susceptible to being confined to sub-optimal local solutions. Consequently, we use 64 perturbed start positions for each bone and choose the solution that yields the smallest value of the error function. Multiple searches per bone can be performed in parallel. The optimization procedure is stable with respect to perturbations in the space of possible rotations. This is consistent with the fact that rotations around an object's center of mass are not likely to change the object's original capture region. The perturbed start positions were, 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 returned the same minimum. The alternative local minima were at least one order of magnitude higher (expressed in squared millimeters).

E. Hierarchical Approach

The distance field formulation allows us to apply the tracking procedure hierarchically, expanding the capture range of our method. We derive a hierarchy empirically, based on a trial-and-error analysis of the start values of the cost function F on a few separate sequences of volume images. For a complex structure like the human wrist (Fig. 5), we use three *in vivo* sequences of volume images. Each sequence consists of ten different wrist poses, each of which corresponds to a different human subject. All possible tree hierarchies starting from the radius and ulna and branching toward the metacarpals were considered; we chose the one which generated best start values of the cost function across all sequences.
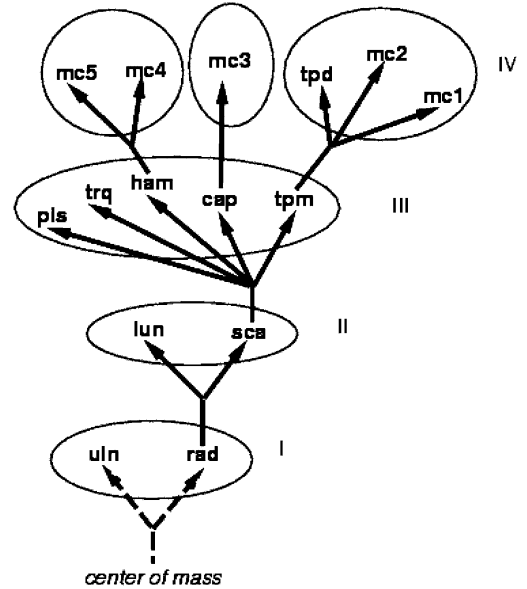
We run the optimization procedure on successive layers of the wrist bones, starting with the forearm bones, as shown in Fig. 6. We iterate through bones: once we detect the motion of bone b_i through cost function optimization (*optimization* transform), we propagate the motion to all the bones that have b_i as an ancestor in the tree hierarchy (*propagation* transform), then we move on to the next bone. Optimization and propagation transformations are accumulated for each bone.

The hierarchical approach ensures that we always start an optimization step from a reasonable neighborhood, thereby boosting the capture range of the registration procedure from less than 5° rotational pose increments to a full range of wrist motion (about 180°), as shown in Fig. 7.

III. VALIDATION METHOD

In this section, we describe a series of experiments where the method was used to register wrist intrasubject CT images. In order to compare our method's performance with earlier results reported in the literature, we begin by evaluating our method's accuracy on high-resolution, marked *in vitro* data. We then examine our method's robustness with respect to practical issues such as image-resolution and perturbations in the registration start-point. We take validation one step further by examining our method's performance on *in vivo*, unmarked data. Finally, we evaluate our method's robustness with typical *in vivo* factors such as variation in image subject and object pose.

A. Data Acquisition

Four different datasets (Table I) were used in our experiments. All datasets were acquired using CT technology (Hispeed Advantage, General Electric Medical System, scan parameters: 80 kV, 80 mA). All images consist of axial slices,
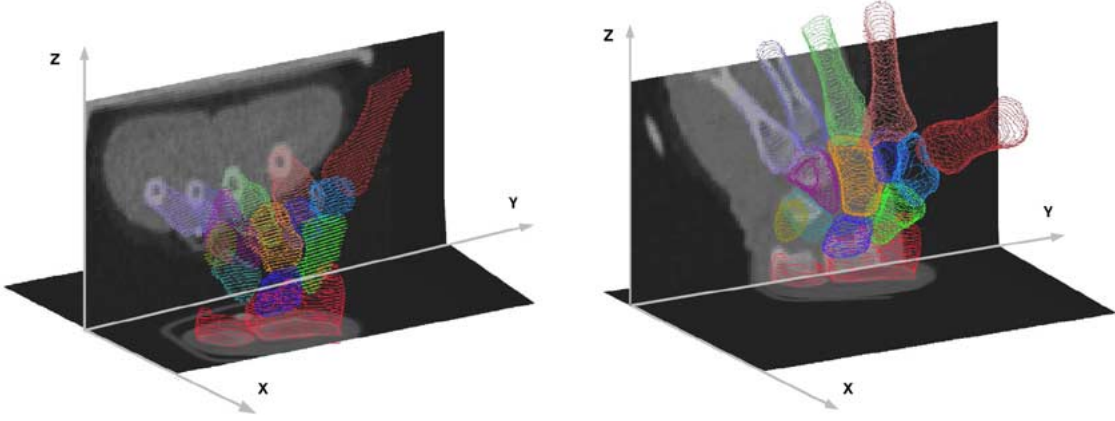
Fig. 7. Imaged wrist-poses do not necessarily come in small motion increments. The images show the same geometric wrist model, after registration, in two different poses. The orthogonal greyscale planes correspond to vertical and horizontal sections through the CT volume images (darker grey areas correspond to soft-tissue, brighter areas to bones). Note the significant differences in bone posture, orientation, and overall wrist location in the scan volume. As shown here, two subsequent instances of the same wrist can be outside each other's capture region.

TABLE I
DATASETS USED IN VALIDATION EXPERIMENTS

Data-set	Type	Number of images (subjects x poses)	Size	Voxel size (mm^3)
A	*in vitro*	1×4	$512 \times 512 \times 141$	$0.31^2 \times 1$
B	*in vitro*	1×4	$171 \times 171 \times 141$	$0.94^2 \times 1$
C	*in vivo*	1×2	$180 \times 180 \times 60$	$0.78^2 \times 1$
D	*in vivo*	80×12	$100 \times 100 \times 80$	$0.94^2 \times 1$

with the x axis oriented horizontally right to left, the y axis horizontally front to back, and the z axis vertically up, such that the image resolution is lowest in the z direction. The geometric-model point clouds have between 2000 and 8000 points, depending on the size of the bone.

B. Experiments

1) In Vitro Accuracy and Robustness Experiment: In this experiment, we evaluate *in vitro* accuracy against the ground truth yielded by external marker registration. We further compare our *in vitro* results with those generated by grey-value registration, implemented as described further below.

To enable comparison with earlier results reported in the literature, we use the high-resolution dataset A, consisting of four CT images of a fixed specimen (separated forearm and hand) in different poses. Each component was encased in plastic resin to prevent relative bone motion. To better reflect the *in vivo* scanning protocol the phantom forearm bones were only partially included in the scan field-of-view for three poses. Seven markers (ceramic spheres of various high-tolerance diameters) were rigidly glued to each specimen component, allowing us to establish the registration ground truth *in vitro*. Marker contours were extracted from each volume image by thresholding at 600 Hounsfield units. The contour images were then processed with a 3-D imaging software package (Analyze AVW 2.5). The centroids of the seven spherical markers (one set per specimen component) were used to calculate rigid-body motion by a method of least squares [6].

In both the tissue-classification method and the grey-value method, the optimization procedure is initialized with the ground truth. The resulting registration transforms should deviate from the given true transform due to each method's translation and rotation error. In both methods, we compute for each registered bone the error relative to the true transform. We report the relative error (mean and standard error obtained by cross-registration of the four images in dataset A) as a translation and rotation in *helical axis of motion* (HAM) coordinates [7]; HAM coordinates express rigid-body motion as pairs (θ, t) of rotations around and translations along a unique helical axis.

Next, we examine our method's robustness with image resolution, since in practice our *in vivo* data's resolution was limited by the large number of subjects participating in motion studies and the large number of images acquired per subject. To this end, we repeat the accuracy experiment on dataset B. Dataset B, designed to simulate lower resolution data, was obtained by smoothing and subsampling the images of dataset A.

Last, we examine the impact of initialization on the tracking procedure. We note that true transform data, as yielded by external markers, is usually not available *in vivo*. To simulate this situation *in vitro*, we perform in this experiment a perturbation study, in which the optimization procedure is restarted repeatedly from the ground truth yielded by external-marker registration, plus a small random rigid-body transformation. We perform a set of five trials, with a translational perturbation of 2 mm (approx. 2 voxels in image space) in a random direction, followed by a second set of five trials, with a translational perturbation of 5 mm in a random direction. Again, we report error relative to the ground truth transform, mean and standard error obtained by cross-registration of the four images in dataset B, for both our method and grey-value registration.

Grey-value registration implementation Grey-value registration was implemented as in Snel *et al.* [8], with several modifications to increase accuracy. First, all the points, as opposed to a random 10%, with a greyscale value greater than 600 Hounsfield units of each image were used in the calculation of the root-mean-square cost function. Second, the values in each target image were obtained by tricubic interpolation, as opposed to linear interpolation. Finally, we used a high-performance library

TABLE II
VALIDATION EXPERIMENTS

Validation experiment	Datasets	Results compared against
in vitro accuracy and robustness (image resol. and start-point perturbation)	A, B	grey-value registration
in vivo accuracy	C	grey-value registration segmentation-based registration visual inspection (collision detection)
in vivo robustness (image subject and object pose)	D	visual inspection numerical analysis

implementation [9], as opposed to a custom implementation, of the downhill simplex method of Nelder and Mead, with a maximum deviation from the initial transform values of $\Delta t = 6$ voxels per axis and $\Delta \theta = \pi/4$. To further boost this method's ability to deal with partially scanned bones, the original cost function was also slightly modified to approximate distance to the target volume whenever the model's points were outside the target image during matching (Section II-C).

2) In Vivo Accuracy Experiment: Because it is technically impossible to know the ground truth *in vivo*, we evaluate our method's accuracy by comparing results with the mean answer of several manual registration trials (described further below), and with the results generated by grey-value registration. In this experiment we use dataset C, consisting of two low-resolution CT images of the same *in vivo* left wrist, one with the wrist in a neutral pose (targeted by visually aligning the back of the hand with the back of the forearm and the third metacarpal with the long axis of the forearm) and one with the wrist extended.

Note that in this experiment we enhance the grey-value method with the hierarchical approach described in Section II-E. Without the hierarchical enhancement, the capture-range capabilities of the grey-value method are surpassed by the range of joint-motion in dataset C, rendering the method inapplicable. Results from all three methods—tissue-classification, manual, and grey-value are further verified using the following visualization method.

Visual validation is performed by superimposing the registered bone geometric wireframe models with vertical and horizontal slices of the volume image. Two sliders control the vertical and horizontal slice displayed. The registration results are automatically checked for potential erroneous collisions between objects that coexist in the same image, at the cost of further geometrical processing. To this end, a NURBS surface is fit to each object geometry (Raindrop GeoMagic, Research Triangle Park, NC), a level-set distance field representation is then generated from the NURBS representation [10], and the interobject distance is evaluated accurately for each vertex of the NURBS surface with respect to all neighboring objects [11]. Collisions are indicated by negative interobject distances and reported to the user. When collisions happen, each object surface is further color-mapped and iso-contoured according to the interobject distance, in order to create an informative visualization (see Section IV). Registration results are also evaluated numerically, by examining the final-fit cost function

values. Results are visually inspected in cases where fit values were abnormally high, i.e., above 0.01.

Segmentation-based registration Five medical school students, all familiar with the segmentation procedure and the anatomy of the wrist, manually segmented the wrist from each of the two volume images in dataset C. Each segmenter took several runs through the procedure, for a total of twelve runs. Registration of the carpal bones between two volume images was subsequently accomplished with an inertia-matching method [12]. Finally, relative motion to the radius was reported for each bone and run. Statistics on the registration results (rotation and translation mean and standard deviation, per bone) were collected; note that these statistics include intraobserver variation.

3) In Vivo Robustness Experiment: Finally, we evaluate our method's *in vivo* robustness with respect to object pose and human subject data (dataset D). Dataset D consists of CT wrist images acquired from forty human subjects. Ten to twelve low-resolution volume images of both wrists were acquired per human subject, spanning a full range of wrist poses. The right and left wrists were subsequently split into separate volumes, for a total of 900 wrist volume images. Registration results are validated both visually and numerically with the method earlier described.

Table II summarizes the datasets and validation methods used for each experiment.

IV. RESULTS

1) In Vitro Accuracy and Robustness Experiment: In the *in vitro* accuracy experiment our tissue-classification registration method demonstrated super-resolution accuracy, and generally had smaller translational errors than grey-value registration. For all registrations the mean tissue-classification translational error was less than 0.3 mm, compared to a mean grey-value translational error of 0.5 mm. The mean rotational error was less than $0.4°$ in both methods (Fig. 8). The grey-value registration results are consistent with those reported by Snel *et al.* [8]. A one-sample t-test ($\alpha = 0.05$) on the difference between the mean results produced by the two methods confirmed the tissue-classification translational accuracy improvement was statistically significant ($p = 0.007$).

Decreased image resolution impacted the accuracy of our method less than the accuracy of the grey-value registration method (Fig. 8). Our method introduced a mean translational
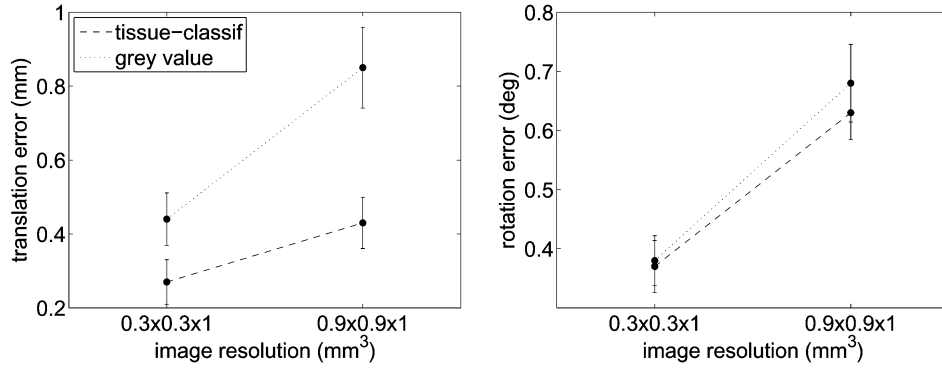
Fig. 8. The accuracy of tissue-classification and grey-value registration, and the influence of image resolution on both methods. For each resolution, we plot the mean and standard error obtained by registering ten carpal bones across four volume images. The tissue-classification method introduces smaller translational errors than grey-value registration. Tissue-classification accuracy degrades more gracefully than the accuracy of grey-value registration.
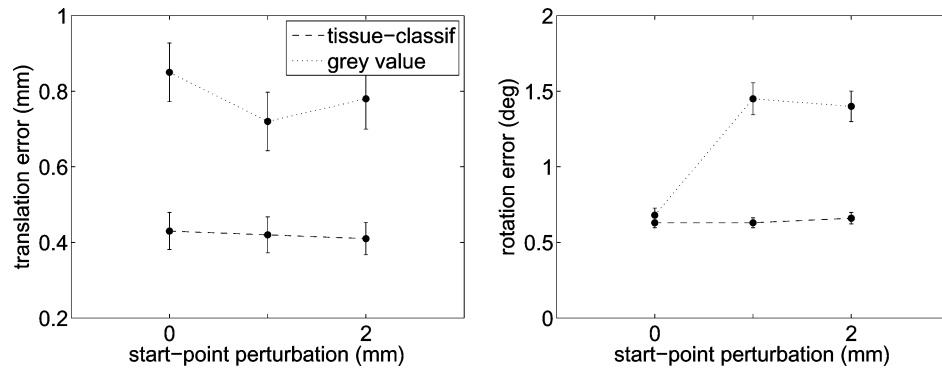

Fig. 9. The influence of start-point perturbation on tissue-classification and grey-value registration accuracy. The amount of perturbation increases on the horizontal axis from 0 mm to 5 mm. At each point, we plot the mean and standard error obtained by registering ten carpal bones across four volume images, in five perturbation trials. The tissue classification method is stable with perturbation, while grey-value registration is not. Increasing the amount of perturbation from 2 mm to 5 mm does not impact further the accuracy of either method.

error of less than 0.4 mm (a 30% translation accuracy decrease when image resolution drops to one third), compared to 0.9 mm translation error in the grey-value method (a 80% accuracy decrease when image resolution drops to one third). The mean rotational error increased to 0.6° in the tissue-classification method, and 0.7° in the grey-value method, respectively. A one-sample t-test ($\alpha = 0.05$) on the difference between the mean results produced by the two methods confirmed the tissue-classification translational accuracy improvement was statistically significant ($p = 0.001$).

The last part of the *in vitro* experiment showed that the tissue-classification registration method maintains super-resolution accuracy with perturbations in the optimization start position (Fig. 9). The tissue-classification method was practically insensitive to perturbation (less than 0.4 mm, 0.6° mean error, 0% degradation with perturbation), while the grey-value method's mean rotational error doubled (0.8 mm, 1.5°), as shown in Fig. 9, middle column. A one-sample t-test ($\alpha = 0.05$) on the 2 mm perturbation results confirmed that the tissue-classification accuracy improvement was statistically significant, with respect to both translation ($p = 0.01$) and rotation ($p = 0.007$). Increasing the amount of perturbation from 2 mm to 5 mm did not further impact the results of either method (Fig. 9). When distortions as small as 2 voxels are present in the optimization start point, the tissue-classification

method significantly outperforms grey-value registration, with an average combined (translation and rotation) accuracy improvement of over 74% (Fig. 10).

2) In Vivo Accuracy Experiment: The *in vivo* accuracy experiment showed good correlation between the results returned by the three methods tested: tissue-classification, grey-value, and segmentation-based (Fig. 11). Since the ground truth is not available *in vivo*, this consistency with the expert segmentation-based results, along with visual inspection, were the best indications of accuracy available. The visual inspection of the grey-value registration results revealed collisions between several pairs of bones. Fig. 12 illustrates a collision detected between the radius and scaphoid bones. No collisions were detected in the results generated using the tissue-classification method. These results indicate that the accuracy of grey-value registration is insufficient when measuring small features such as interbone spacing.

3) In Vivo Robustness Experiment: In the *in vivo* robustness experiment, more than 13 500 bones were registered through the tissue-classification method. Visual and numerical validation showed that the method was stable with both object poses and human subjects. In less than 0.1% of cases (8 bone instances), numerical validation indicated suspicious fit values. For each of these cases, further visual inspection revealed that the abnormal fit values resulted from errors in the scanning pro-
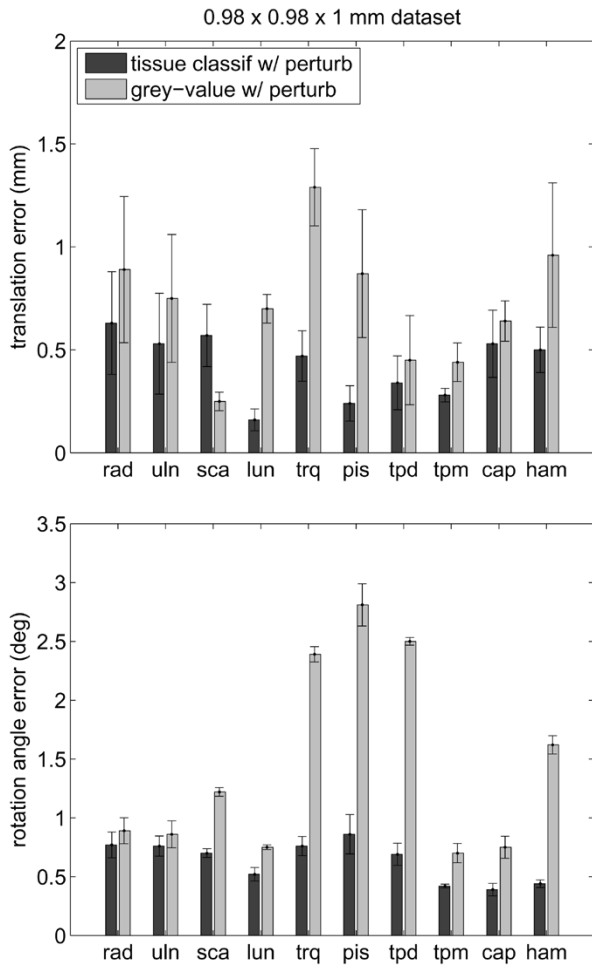
Fig. 10. *In vitro* tissue-classification registration error and grey-value registration error with a 2 mm random perturbation in the optimization start point (top—translation, bottom—rotation). We register each bone across four volume images. Each registration is performed five times, each trial corresponding to a different initial translational perturbation of 2 mm in a random direction. For each bone, we plot the mean and standard error, thus, obtained.
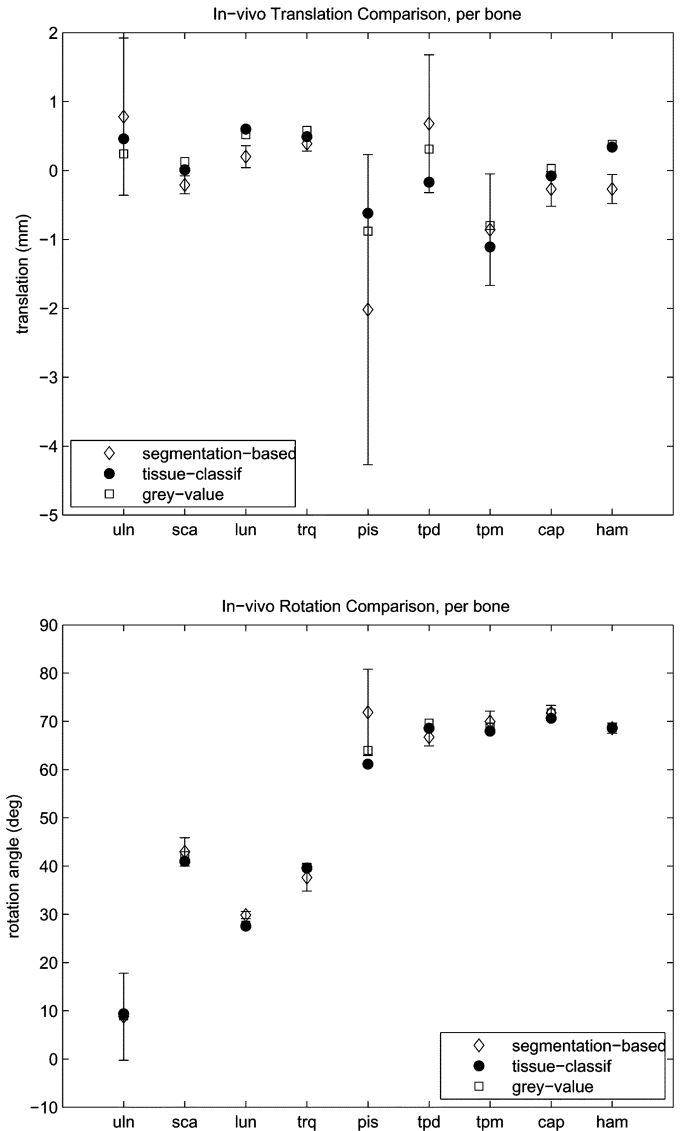

Fig. 11. Tissue-classification registration results versus segmentation-based registration and grey-value registration results. For each bone, we plot the mean and standard deviation obtained by manual registration in a total of twelve runs, the tissue-classification registration result, and the grey-value registration result. Both tissue-classification and grey-value registration results are generally within one standard deviation of the mean expert segmentation-based registration results.

cedure, whereby the respective bone was only partially included in the target scan. With occasionally as much as half of a bone missing from the volume image, visual verification showed that the bone was still being registered correctly using the information available.

V. DISCUSSION

Our results indicate that tissue-classification registration consistently attains subvoxel accuracy. The method maintains subvoxel accuracy despite decreasing image resolution, and is stable with perturbations in the initial optimization start position. Furthermore, visual and numerical validation during clinical application (*in vivo* robustness experiments) shows that the method is robust with varying object poses and subjects.

Tissue-classification should be given primary credit for our method's accuracy, because the object boundary estimated through tissue classification has super-resolution accuracy. The comparison with grey-value registration shows that without the super-resolution boundary estimate, the matching process is effectively reduced to using voxel-wide estimates, which

results in lower accuracy. The resulting distance field representation further ensures that registration results converge to the same value, regardless of perturbations in the start position. The distance field acts as a gradient guiding the search to the global minimum. Without this gradient, the search can easily be trapped in local nearby minima, as our *in vitro* perturbation-stability experiments with grey-value registration show.

The comparison with grey-value registration reveals why tissue-classification maintains subvoxel-accuracy with decreasing image resolution. When the sampling step is sufficiently small with respect to the desired features, the raw volume image often offers rich information: object texture-patterns can be implicitly identified and used in the matching process. In this case, *more* raw information can be *better*. However, as the features of interest become smaller, imaging noise
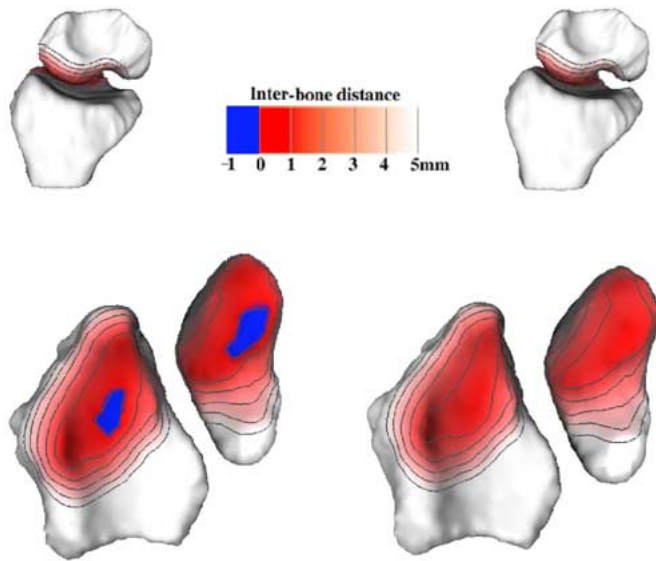
Fig. 12.   Visual inspection shows collisions between bones registered using grey-value registration. (Left) *In vivo* grey-value registration registration of the radio-scaphoid joint versus (right) tissue-classification registration of the same joint from dataset C. Bones are color mapped and contoured according to the distance between bones post-registration. The saturation of color on bone surfaces represents the distance to the nearest point on the opposite bone. Contour lines are drawn at 1 mm intervals. (Top) Bones in their correct anatomical context—note the two registration methods yield similar bone poses. (Bottom) Bones rotated to show articulated surfaces more clearly. Blue corresponds to negative interbone distances, indicating collision in the grey-value registration result.

effectively blurs the informational content of individual voxels (Fig. 13). In this case, *quality* processed-information—like accurate boundary-estimates obtained through unsupervised learning—begins to matter. As long as we guarantee super-resolution boundary estimates, registration accuracy stays within the subvoxel range.

In the process of developing our super-resolution accurate registration method, we have proposed new computational methods that are applicable to a broader scope of medical image processing. Our technique uses neighborhood information throughout a volume to generate localized distance fields directly from sampled datasets. No feature points need to be presegmented per scan or subject in order to generate the distance fields. Distance fields have been used before to expedite registration processes [13]; however, no previous work has generated these distance fields with super-resolution accuracy. Recovering material boundaries from sampled datasets, as well as generating distance fields once geometric models have been extracted, is a research topic in several fields [10], [14], [15]. We build on the work of Laidlaw *et al.* [2], who use Bayesian probability theory to classify accurately tissues in medical volume images. We are not aware of other work in generating super-resolution accurate distance fields directly from sampled datasets. Although we limited the distance field computation in our experiments to a five-voxel band around material boundaries, this distance computation could be either performed or propagated beyond this threshold [10]. However, we showed that the hierarchical approach described in Section II-C can overcome potential capture region limitations without the additional cost of extended distance computations.

Our matching procedure employs a new similarity measure that, unlike chamfer matching or iterative closest point measures, incorporates distance field knowledge from both the reference and target images. The measure also accounts for partially scanned objects. We believe this measure can improve the stability of other registration procedures with respect to errors introduced both by geometric model segmentation and partially scanned objects. While in recent literature a large number of sophisticated similarity measures have been proposed [16], we note these measures were developed mainly for intermodality registration, and likely do not have strong advantages over sums-of-squared-differences or root-mean-square matching when applied to images obtained through the same modality. In our tissue-classification registration approach, sum-of-squared-differences is in fact the natural similarity choice, since the reference and target distance-field intensities corresponding to registered points stand, by construction, in an identity relationship.

The tissue-classified distance field approach helped us identify a motion-directed multiobject hierarchy in the wrist-joint case. The potential physiological implications of this hierarchy are beyond the scope of this paper, but the hierarchy enabled our registration method to trace motion between wrist poses that were not within each other's capture region. A similar approach is likely to boost the capture-range capabilities of other local-search registration methods.

While accurate, our tissue classification method poses computational challenges. Cross-registering a series of twelve wrist volume images (15 objects per pose) takes 20 min on a multinode i686 cluster running Linux (AMD $Athlon^{TM}$ XP 2700+). We emphasize, however, that the focus of our work is boosting accuracy, and not minimizing running time. As no user interaction beyond the initial segmentation and potential final visual validation is required, the registration is, after all, run off-line.

While our approach addresses successfully some common problems found in medical image processing, it also relies on several assumptions. Our tissue-classification procedure assumes a simple tissue-structure, in which hand soft-tissue is assimilated with marrow-tissue. While modeling a more complex structure (for example, one that would distinguish between trabecular and cortical bone) would likely further improve our method's accuracy, our results indicate the simple model suffices for subvoxel-accurate registration. In the validation phase, we assume implicitly that *in vitro* accuracy is an indicator of *in vivo* accuracy. We note that, in fact, in our *in vitro* experiments the tissue-classifier misclassified several voxels. These misclassifications can be attributed to the differences between the material composition of a fixed specimen encased in resin with all soft tissue removed, and the *in vivo* tissue model we assumed. While the inability to determine the *in vivo* ground truth makes it difficult to compare *in vivo* and *in vitro* accuracy, it is reasonable to assume that errors in the classification process reflected negatively on the registration results. We speculate in this view that our *in vivo* results, in fact, surpass the accuracy of our *in vitro* results. Last, but not least, we note that we interpret collisions or the lack of collisions in our *in vivo* experiments not as a method validation measure, but as an indicator that error size matters.
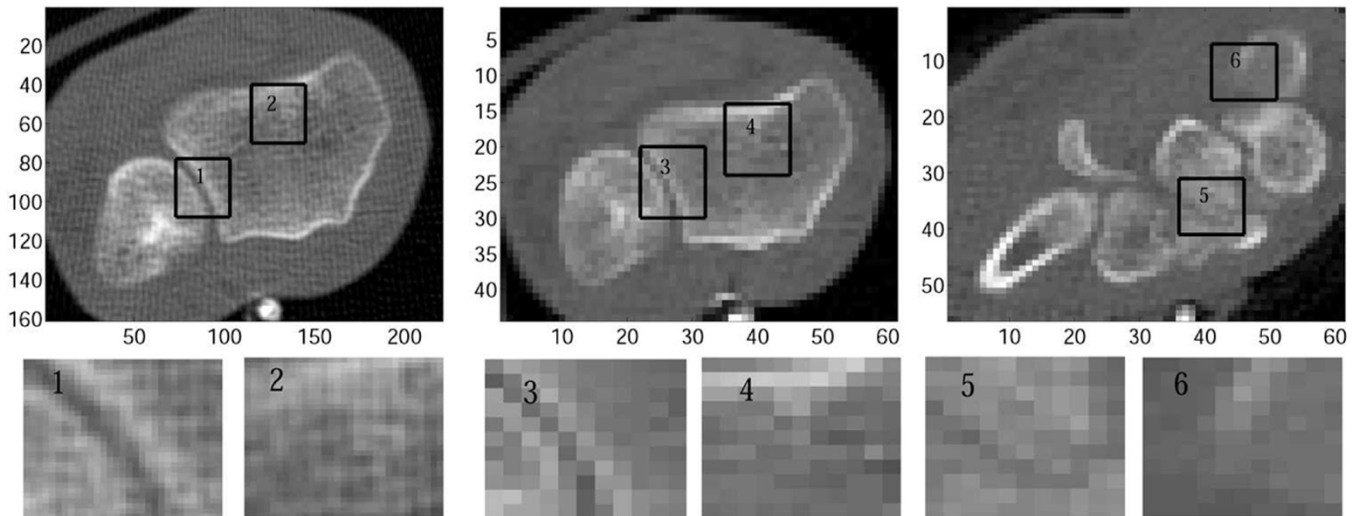
Fig. 13.   Image resolution impacts voxel information-value: as resolution decreases, "more" raw grey-values may deliver less information than "fewer" information-enhanced voxels. From left to right: one slice from a high-resolution CT image (0.3×0.3 mm), and two slices from a low-resolution CT image (same wrist, 0.9×0.9 mm). Each area of interest is detailed in the bottom row. Interbone distances that are narrow in high-resolution images (Box 1) become a mere voxel wide in low-resolution images (Box 3). 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, and Box 6, note the soft or diffused bone boundary. As shown in our *in vitro* experiments, accurate boundary estimates based on information throughout the scan contribute more information than such collections of unprocessed grey-values.

The accuracy results we obtained challenge opinions widely held about the superiority of voxel-property-based over segmentation-based registration methods [17]. Our tissue-classification registration can be regarded as a hybrid method, combining the strengths of the two approaches—voxel-property and segmentation-based. While our matching procedure evaluates the correlation between source and target image-values throughout a volume (like voxel-property methods do), through tissue-classification we nevertheless discard the original volume data that does not convey object-boundary information (like most segmentation-based methods do). The recovered object boundary can be thought of as having super-resolution, in that through our approach boundaries are detected with higher accuracy than an image's resolution allows. Nevertheless, due to image noise and modeling assumptions, this boundary is not perfect. Should one be able to recover more accurate object boundary information, generate a distance field from that boundary, then perform distance-field matching as described in Section II, registration accuracy would only increase. This suggests that, contrary to current knowledge, in the long run segmentation-based methods have the potential to surpass the accuracy of voxel-property methods.

VI. Conclusion

We presented in this paper a novel intrasubject method for subvoxel-accurate registration of objects from CT volume images. Results show average accuracy improvements of 74% over grey-value registration. The method is of particular interest to applications where collections of tightly packed, small objects need to be registered. To this end, we showed in a wrist data application that earlier registration methods can introduce false interobject collisions, while the new method does not.

The tissue-classification registration method maintains subvoxel accuracy with decreasing image resolution, and is stable with perturbations in the initial optimization start position. The method is also stable with respect to partially scanned objects, and with varying object pose and subject. Our approach should be of interest to any registration applications where super-resolution accuracy is desired.

References

[1] J. J. Crisco, R. D. McGovern, and S. W. Wolfe, "A noninvasive technique for measuring *in vivo* three-dimensional carpal bone kinematics," *J. Orthop. Res.*, vol. 17, no. 1, pp. 96–100, 1999.

[2] D. H. Laidlaw, K. W. Fleischer, and A. H. Barr, "Partial-volume bayesian classification of material mixtures in MR volume data using voxel histograms," *IEEE Trans. Med. Imag.*, vol. 17, no. 1, pp. 74–86, Feb. 1998.

[3] D. H. Laidlaw, "Geometric model extraction from magnetic resonance volume data," Ph.D., CalifornInst. Technol. (Caltech), Pasadena, 1995.

[4] NAG Fortran Library Routine Document. [Online]. Available: http://www.nag.co.uk/numeric/fl/manual/pdf/e04jyf.pdf

[5] M. L. Richardson. Radiographic Anatomy of the Skeleton. [Online]. Available: http://www.rad.washington.edu/RadAnat/WristPALabeled.html

[6] C. P. Neu, R. D. McGovern, and J. J. Crisco, "Kinematic accuracy of three surface registration methods in a three-dimensional wrist bone study," *J. Biomech. Eng.*, vol. 122, no. 10, pp. 528–533, 2000.

[7] M. M. Panjabi, M. H. Krag, and V. K. Goel, "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.*, vol. 14, pp. 447–460, 1981.

[8] J. G. Snel, H. K. Venema, T. M. Moojen, M. Ritt, C. A. Grimbergen, and G. J. den Heeten, "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.*, vol. 27, no. 9, pp. 2037–2047, 2000.

[9] NAG Fortran Library Routine Document. [Online]. Available: http://www.nag.co.uk/numeric/fl/manual/pdf/E04/e04ccf.pdf

[10] S. Mauch. A Fast Algorithm for Computing the Closest Point and Distance Transform. [Online]. Available: http://www.acm.caltech.edu/seanm/software/cpt/cpt.pdf

[11] G. E. Marai, D. H. Laidlaw, C. Demiralp, S. Andrews, C. M. Grimm, and J. J. Crisco, "Estimating joint contact areas and ligament lengths from bone kinematics and surfaces," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 5, pp. 790–799, May 2004.

[12] J. J. Crisco and R. D. McGovern, "Efficient calculation of mass moments of inertia for segmented homogeneous 3D objects," *J. Biomechan.*, vol. 31, no. 1, pp. 97–102, 1998.

[13] M. Van Herk and H. M. Kooy, "Automatic three-dimensional correlation of CT-CT, CT-MRI, and CT-SPECT using chamfer matching," *Med. Phys.*, vol. 21, pp. 1163–1178, 1994.

[14] K. S. Bonnell, M. A. Duchaineau, D. A. Schikore, B. Hamann, and K. I. Joy, "Material interface reconstruction," *IEEE Trans. Vis. Comput. Graphics*, vol. 9, no. 4, pp. 500–511, Oct.-Dec. 2003.

[15] R. Whitaker, D. Breen, K. Museth, and N. Soni, "Segmentation of biological volume datasets using a level set framework," in *Volume Graphics 2001*. Vienna, Austria: Springer, 2001, pp. 249–263.

[16] J. West, J. M. Fitzpatrick, M. Y. Wang, C. R. Maurer, R. M. Kessler, R. J. Maciunas, C. Barillot, D. Lemoine, A. Collignon, F. Maes, P. Suetens, D. Vandermeulen, P. A. van den Elsen, S. Napel, T. S. Sumanaweera, B. Harkness, P. F. Hemler, D. L. G. Hill, D. J. Hawkes, C. Studholme, J. B. A. Maintz, M. A. Viergever, G. Malandain, X. Pennec, M. E. Noz, C. Q. Maguire, M. Pollack, C. A. Pellizari, A. Robb, D. Hanson, and R. P. Woods, "Comparison and evaluation of retrospective intermodality image registration methods," *J. Comput. Assist. Tomogr.*, vol. 21, pp. 554–566, 1997.

[17] J. B. A. Maintz and M. A. Viergever, "A survey of medical image registration," *J. Med. Image Anal.*, vol. 2, no. 1, pp. 1–36, 1998.