DYNAMIC REGISTRATION USING ULTRASOUND FOR ANATOMICAL REFERENCING

Haydar Talib[1], *Martin Styner*[2], *Tobias Rudolph*[1], *and Miguel A. González Ballester*[1]

[1]MEM Research Center, Institute for Surgical Technology and Biomechanics, University of Bern, Switzerland
[2]Departments of Computer Science and Psychiatry, University of North Carolina, Chapel Hill, USA

ABSTRACT

This paper proposes methods to circumvent the need to attach physical markers to bones for anatomical referencing in computer-assisted orthopedic surgery. Using ultrasound, a bone could be non-invasively referenced, and so the problem is formulated as the need for dynamic registration. A method for correspondence establishment is presented, and the matching step is based on three least-squares algorithms: two that are typically used in registration methods such as ICP, and the third is a form of the Unscented Kalman Filter that was adapted to work in this context. A simulation was developed in order to reliably evaluate and compare the dynamic registration methods.

Index Terms— Point-to-point registration, Kalman filtering, ultrasound, computer-assisted surgery, anatomical referencing

1. INTRODUCTION

Surgeons who regularly perform computer-assisted orthopedic surgery (CAOS) justify its use by reporting more accurate and faster interventions, as well as enhanced visualization and reduced invasiveness [1]. One of the enabling technologies of CAOS is surgical navigation, whereby surgical tools and the bones to be operated are localized in real-time by a tracking system. Localization, or referencing, is made possible by using dynamic reference bases (DRBs), which are markers that are rigidly attached to the objects in question.

DRBs are fixed onto bones using either clamps, screws, or pins. Their placement is often limited to surgically exposed areas, but in some cases additional incisions may be required for attaching DRBs, as in the case of navigated kyphoplasty [2]. The attachment of DRBs to bones is an invasive process, and even when no additional incisions are needed, it must be ensured that DRBs remain rigidly attached to the objects being referenced. Movement of DRBs attached to bones occurs frequently, causing frustration to surgeons and prolongation of surgeries [3]. Some solutions have been explored for the fixation of DRBs, but they focussed on using different materials, or placing DRBs at different locations [4].

The need arises at this point to explore alternative solutions to the direct attachment of DRBs to bones. We therefore propose the use of ultrasound (US) imaging as a means to non-invasively reference bony anatomy. Two-dimensional US imaging systems are common, and yield numerous advantages: they do not have detrimental effects on patients' health; they can provide real-time images during surgery; they are non-invasive; and they can provide information from surgically inaccessible areas.

Consequently, in the scope of CAOS, a DRB can be attached to the US probe, which is then referenced in the same manner as other tools used for the surgery. If we consider the use of US for anatomical referencing, the only positional information that would be available is located in the US images. We will also consider that the CT scan of the anatomy of interest is available.

We therefore formulate the problem as the need for dynamic registration. From the partial, underdetermined information provided by navigated 2D US, the aim is to provide the 3D position of the anatomy of interest by registering, for instance, a 3D surface model obtained by having segmented the CT scan prior to surgery.

For non-invasive referencing in CAOS, accurate calibration of the US is required, whereby the coordinates of the US image plane are known with respect to the DRB attached to the probe. Furthermore, a rigid registration of the CT to the anatomy would be needed for initialization. More information regarding calibration and registration can be obtained in [5].

This paper presents the first attempt at dynamic registration using US, and to this end three methods for least-squares problems were chosen and adapted to the application. The main selection criteria for the methods were that they should be computationally efficient and that they should be accurate. As such, we chose the methods of Horn [6], Arun et al. [7] and the Unscented Kalman Filter (UKF) for least-squares rigid matching of two point sets [8].

Computer-assisted kyphoplasty could be an initial application for the methods that we present here, since the attachment of a DRB causes the procedure to be more invasive. We therefore begin by exploring non-invasive referencing of the lumbar vertebrae.

In §2, we present a simulation that was made in order to reliably evaluate the three methods, which are described in §3. §3 is divided into two subsections: §3.1, where we present a method for determining correspondence and §3.2, where we outline the three least-squares approaches. A comparison of the three dynamic registration methods is presented and discussed in §4 and §5. A list of some of the variables used throughout this paper is provided below:

	Rigid transformation parameters consist of the Euler angles and translations for the x-, y- and z-axes, $[\alpha, \beta, \gamma, t_x, t_y, t_z]^T$
S	The surface at the true position
S_E	The surface at the prior estimated position
S'_E	Surface at estimated position after 2D estimation
S''_E	Surface at estimated position after 3D estimation
y_k	Measurement obtained from S at time step k
$\bar{y_k}$	Points obtained from S_E at time step k
θ	Phase of the centroid in 2D, given by $\theta = \arctan\left(\frac{\mu_y}{\mu_x}\right)$

2. CONSTRUCTING A SIMULATION

In the non-invasive referencing scenario, the principal aim is to dynamically extrapolate the position of bony anatomy with respect to the tracked US probe. During surgery, the anatomy will be displaced due to some patient motion as well as perturbations from the various tools used in the procedure.

To simulate the motion of a bone with respect to an US imaging probe, we begin with a surface model of one L4 lumbar vertebra,

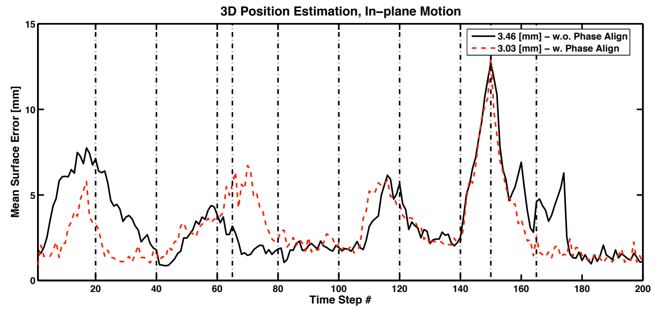

Fig. 1. The mean surface error between S_E and S throughout the simulation, illustrating the effects of phase and centroid alignment for Horn's method. Vertical lines indicate changes in motion.

which was obtained from the CT scan of a plaster vertebra model. The US image plane is defined to have a roughly realistic shape and size (fan-shaped, 6 [cm] base width, and a depth of 4 [cm]), and is initially placed orthogonal to the vertebra as it would be on a patient. As a consequence, the spinous process and the facet joints will be the most visible structures of the vertebra.

The set of points y_k, representing the measurement at time k, is defined as the intersection of the US image plane with S, with a small thickness along out-of-plane directions. We used an out-of-plane thickness of 2 [mm], and all the intersecting points from the surface are projected onto the simulated US plane. White Gaussian noise was then added to the measurements ($N(0, 1[mm^2])$). The resulting points can be regarded as analogous to the result of a threshold-based segmentation applied to 2D US images of the vertebra immersed in water, but it should be noted that more information from the anatomy will be available than in real ultrasound images. To speed up computation of the methods in §3, y_k is evenly sampled to consist of no more than 50 points.

Over time, the vertebra surface model S undergoes varying rigid motion, and y_k is computed at each time step, providing dynamic input information. To simplify the presentation of our methods, we will consider out-of-plane motion to occur as changes along the z-axis. Alignment of the US imaging plane to the xy-plane of world coordinates can be ensured by using the change of coordinate matrix.

3. DYNAMIC REGISTRATION USING ULTRASOUND

As the vertebra moves, a new y_k is provided through the US images. $\bar{y_k}$ is determined from S_E, the surface at the prior estimated position, by computing the intersection between the US plane and S_E. The process is similar to determining y_k, but no Gaussian noise is added.

The aim at this point is to determine the rigid transformation that would match the position of $\bar{y_k}$ to that of y_k. The same transformation could then be applied to S_E so that it matches the unknown position of S. Since S has undergone unknown rigid motion, and y_k contains error, the first step is to establish correspondences between y_k and $\bar{y_k}$. The rigid transformation is then determined by minimizing the sum-of-squares error between y_k and $\bar{y_k}$.

We will exploit in-plane information at every time step by making an initial estimate considering only in-plane motion. Applying the 2D rigid transformation, we obtain S'_E and assume that it has been successfully aligned to S in the xy-plane. A new $\bar{y_k}$ is generated from S'_E. We then hypothesize that differences in $\bar{y_k}$ and y_k will be mostly due to out-of-plane motion, and then compute the correspondence and 3D rigid transformation to obtain S''_E, the surface that should adequately match S.

At each time step, the aim is to determine S's position using y_k:

(a) The new measurement y_k is obtained from S, and $\bar{y_k}$ in 2D is extracted from S_E (§2). The in-plane correspondence is determined between y_k and $\bar{y_k}$ (§3.1).
(b) An estimate is made for the 2D rigid transformation, which is applied to S_E, yielding $S'_E = R_{2D} * S_E + t_{2D}$ (§3.2).
(c) A new $\bar{y_k}$, in 3D, is extracted from S'_E (§3.1). The correspondence is determined in 3D (§3.1).
(d) The correspondence from (c) is used for estimating the 3D rigid transformation, which is applied to S'_E, yielding $S''_E = R_{3D} * S'_E + t_{3D}$. At the next time step, $S_E = S''_E$ (§3.2).

3.1. Establishing Correspondence

Correspondence in iterative registration algorithms is usually determined by minimizing the Euclidean distance between two point sets [8, 9]. Although iterative methods may perform quickly, it is unlikely that they would be suitable for real-time applications, and so the need arises for a simple and yet more useful approach to dynamically finding correspondence.

In-plane correspondence: Motion occurring within the US plane can be described by three parameters: one rotation, γ and two translations, t_x and t_y. Using y_k and $\bar{y_k}$, initial guesses can be made about the parameters.

Assuming that motion between time steps will be small, y_k and $\bar{y_k}$ should be very similar. The rotation parameter between the two can be obtained through the phase of each centroid with respect to the origin in world coordinates. Once the phase is obtained, $\gamma = \theta - \theta_E$. To obtain initial guesses for t_x and t_y, we use the centroids of each point set, and the difference of centroids provides the two translation parameters.

Applying the 2D rigid transformation matrix using the guesses for $[\gamma, t_x, t_y]^T$ should roughly align the two data sets, and the correspondence can then be obtained by minimizing the Euclidean distance. This transformation is only applied in order to determine the correspondence, so $\bar{y_k}$ remains in its position until an estimate is made using one of the methods in §3.2. Nevertheless, it is possible to provide examples when using this approach for aligning $\bar{y_k}$ to y_k can be erroneous. To account for such cases, we compute the Euclidean distance with and without alignment, and select the correspondences that are determined by the method that has smaller distances. This solution does not necessarily guarantee that the suitable method will be selected, but it reduces the likelihood of an erroneous result.

Fig. 1 illustrates one example of the improvement due to this approach for establishing correspondence, compared to simply minimizing the Euclidean distance.

3D correspondence: Once the in-plane correspondence has been established, an estimate is made for the 2D rigid transformation using one of the three methods described in §3.2. The transformation is then applied to S_E, yielding S'_E, and a new $\bar{y_k}$ is generated.

To account for out-of-plane motion, $\bar{y_k}$ is extended orthogonally to the US image plane. Now the intersection between the plane and S'_E is computed as before, except the out-of-plane thickness is taken to be 5 [mm] instead of 2 [mm]. Here, however, we do not apply any form of alignment, and instead only minimize the Euclidean distance between the sampled y_k and $\bar{y_k}$.

The key difference at this step is that although the correspondence is established using only the xy-coordinates, the z-coordinates of $\bar{y_k}$ will be retained for use in the 3D rigid transformation estimate.

3.2. Least-squares Matching

The three methods we chose for rigid least-squares matching were all originally formulated with the intent of minimizing computation, and they are usually applied iteratively in registration problems. The methods of Horn [6] and Arun [7] were derived to work with large sets of points, so they can be applied straightforwardly. The UKF used for registration [8] was formulated to function with an incrementally increasing number of points, so for our purposes it was adapted to use full point sets for each estimate.

In our dynamic registration scenario, least-squares matching is required twice at each time step (once for in-plane estimation, and a second time for the full 3D estimation). In the case of in-plane estimation, both y_k and y_k^- lie on the xy-plane. In the case of 3D estimation, although y_k has zero-valued z-coordinates, y_k^- will likely have non-zero values for z.

Horn and Arun's Methods: Horn's method is usually used in ICP algorithms, and was favored by Besl [9] over Arun's method to avoid the risk of reflections due to coplanarities in the data. Both methods are based on similar principles, with the main difference being that Horn uses a unit quaternion-based derivation for the rigid transformation matrix, and Arun provides the rotation matrix itself.

For the dynamic registration application, Horn's method was directly applied without modification. To account for reflections, Arun proposed computing the determinant of the resulting transformation matrix, with a value of -1 indicating an erroneous result.

In the case of in-plane estimation, y_k and y_k^- are clearly coplanar, so Arun's formulation for 3D rigid transformation cannot be directly applied. Instead, the method can be simplified to the 2D case, and the condition that two point sets should not be coplanar instead becomes that they should not be colinear along either the x- or y-axes.

The Unscented Kalman Filter: Moghari et al. [8] applied the UKF to estimate rigid transformation parameters in US-to-CT-surface registration. By their implementation, the UKF iterates N times, where N is the number of points in their y_k^-, and for each iteration, the number of points is gradually increasing. For the dynamic registration application, it would be necessary to consider the full set of points for each iteration of the UKF. The 3D formulation of the UKF can be reduced to 2D, so we only consider the 3D case here.

When using the UKF with more than one point at a time, one modification is necessary. y_k^- is normally a $3 \times N$ matrix, but now we vertically concatenate all the points such that it becomes a $3N \times 1$ column vector, $[y_{x1}^-, y_{y1}^-, y_{z1}^-, \ldots, y_{xN}^-, y_{yN}^-, y_{zN}^-]_k^T$. Consequently, the variables that depend on y_k^- similarly increase in size [10].

Taking all the points in y_k^-, P_{yy}, the predicted measurement covariance, becomes a $3N \times 3N$ covariance matrix. The increased dimension of P_{yy} in this way makes it susceptible to being ill-conditioned, which would create instability in the UKF. One way to ensure that P_{yy} remains well-conditioned is by strengthening the diagonal components. This can be done by adding its trace along the diagonal, or some multiple of the trace (we use a value of $1/3 * tr(P_{yy})$). This issue falls in the general category of ill-posed problems, and several approaches can be taken to dampen, or regularize P_{yy}.

4. EXPERIMENTS AND RESULTS

The simulation and the three dynamic registration methods were implemented using MATLAB. The methods performed equivalently, but were not optimized for speed. Each method required around 800 [ms] of computation per time step on a 1.67 [GHz] processor.

At the beginning of each simulation, S_E was initialized with a "perfect" registration, that is $S_E = S$. Then, over a period of at least 120 time steps, S undergoes varying rotations and translations. All motion was applied with respect to world coordinates.

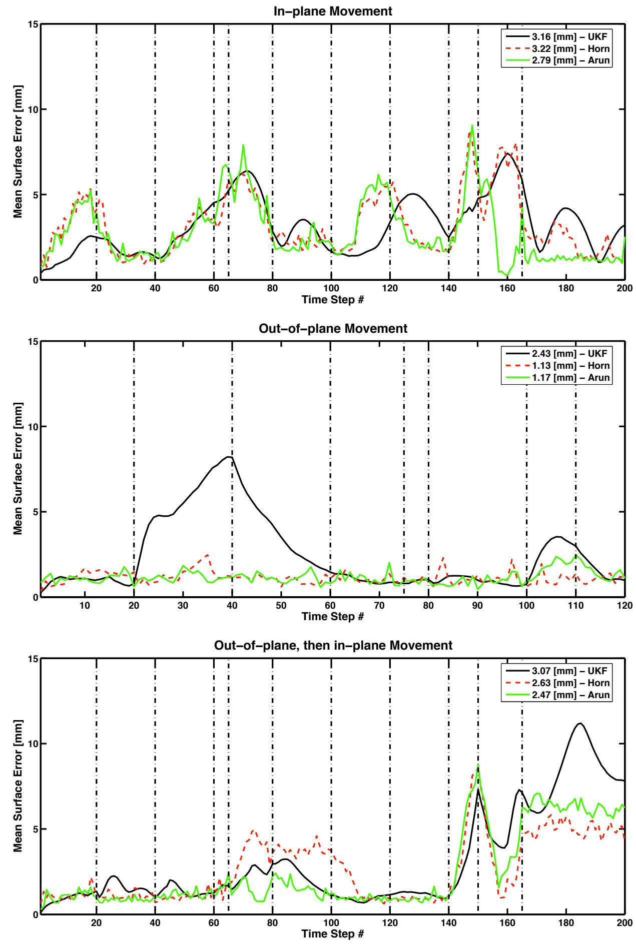

Fig. 2. The mean surface error between S_E and S throughout the simulation, comparing the three methods for dynamic registration. The graphs correspond to three different sequences of motion. Vertical lines indicate changes in motion.

Fig. 2 illustrates the results of three simulations that were used to compare the dynamic registration methods. The mean surface error, plotted with respect to time, was computed as the mean point-to-point distance between S_E and S. Since S_E and S represent the same object, albeit in different positions, their exact correspondence at each time step is known, enabling reliable error computation.

For the top graph of fig. 2, only in-plane motion was simulated, though the three methods were estimating 3D transformations. Transitions between different motions are indicated by the vertical dotted lines, and only γ, t_x and t_y were modified. The sequence consisted of (each motion was applied per time step): $t_y = 0.25[mm]$; then S stopped moving; $t_x = 0.3[mm]$ and $t_y = -0.25[mm]$; a rotation of $\gamma = 0.03[rad]$; $\gamma = -0.03[rad]$; S stopped; $t_x = -0.3[mm]$; S stopped; $\gamma = 0.05[rad]$; then $\gamma = -0.05$; and finally S stopped.

Similar sequences of motion were used for the other graphs of fig. 2. For the middle graph, the motion of S was exclusively out-of-plane. For the bottom graph, the sequence began with out-of-plane motion, but after the 20$^{\text{th}}$ time step, the motion was strictly in-plane.

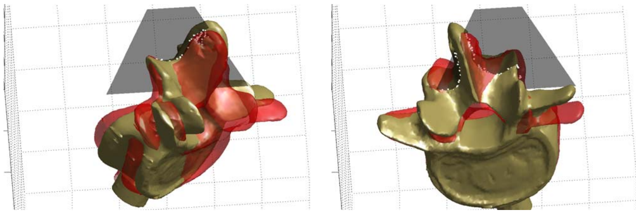

Fig. 3. In the simulation, S_E is displayed translucent, and S is opaque. (left) When the measurement data is insufficient the algorithms can get caught. (right) Another case of insufficient measurement data, as the vertebra moves out of the US image plane.

The simulation represented by the bottom graph consisted of rotating the vertebra out of the US image plane such that the anatomy became less visible. This simulation was performed to test the robustness of the methods when only a few, tightly-clustered points are available from the bone that's being referenced.

In the top graph of fig. 2, the three methods perform equivalently, with Arun's method having a slightly lower mean surface error (averaged over time) of 2.79 [mm], compared to the UKF (3.16 [mm]) and Horn's method (3.22 [mm]).

In the middle graph, a clear distinction can be made between the UKF and the two other methods. Horn and Arun's methods performed equivalently, with average mean surface errors slightly above 1.1 [mm]. The UKF, however, had an average mean surface of error of 2.43 [mm] and had a peak near the 40^{th} time step, which occurred when a change in out-of-plane rotation was applied to S.

For the bottom graph, the three methods once again performed equivalently, and Arun's method had a slightly lower average mean surface error with a value of 2.47 [mm]. All three methods had high mean surface errors towards the end of the simulation.

Fig. 3(left) illustrates the situation occurring at the end of the simulation from the bottom graph of fig. 2. Fig. 3(right) shows another instance when estimation is inadequate, but referring to the top graph of fig. 2, at the peaks occurring near the 70^{th} time step, the methods are able to recover and hence provide improved estimation.

5. DISCUSSION

Throughout the experiments, the algorithms performed well in estimating the position of S using y_k so long as there was a suitable view of the vertebra. The UKF did not perform as well as Horn and Arun's methods for the case of out-of-plane motion, though it was able to recover, as can be inferred from the middle graph of fig. 2. The case where the methods could not suitably estimate S's position was due to a minimal view of the vertebra (see fig. 3(a)).

The measurement should span the object sufficiently that changes in motion can be ascertained from changes in y_k. This is an important aspect of the problem, and will be more serious as we begin examining dynamic registration using real US data in the future.

Although this paper only presented simulations using a vertebra, the methods were formulated to work for other areas as well. We could not make any assumptions about the shape or nature of the measurement data, nor could any assumptions be made about the anatomy's motion. Accordingly, y_k could only be treated as a point cloud, and we've seen that in the current formulation, points in y_k do not need to be close to one another, or form a visually coherent object, but rather the points should sufficiently span the anatomy. In addition to y_k, the CT surface model of the vertebra was used as *a priori* information in order to dynamically simulate $\bar{y_k}$.

The initial registration of the CT surface model to the bone should be adequate, but as can be inferred from the results, the algorithms are able to recover. The assumption was made that y_k was akin to the result of a segmentation algorithm. Despite the added noise, however, the measurement data was usually reliable, though this will not be the case in a real scenario. With real US data, the onus is placed on the need for a more sophisticated, and yet computationally efficient method for dynamic correspondence establishment. It would also be interesting to explore how the methods presented here could be applied directly to the greyscale information, rather than using points obtained by a segmentation step [11].

Overall, the UKF did not perform as well as Horn and Arun's methods. Although the UKF was less accurate, its estimations yielded smoother motion. Looking closely at the graphs of fig. 2, we can observe that Horn and Arun's estimations contain high-frequency variance, so the vertebra has "shaky" motion in some instances. At this stage, a combination of the UKF with either Horn or Arun's method may provide a solution that joins the smooth estimates of the former to the accurate estimates of the latter.

Acknowledgments: The authors wish to thank the NCCR Co-Me, BrainLab AG and the AO Foundation for funding this project. The authors also thank Ségolène Tarte (UCL) and Mauricio Reyes, and extend a special thanks to Professor Lutz Dümbgen of the Institute for Mathematical Statistics, University of Bern.

6. REFERENCES

[1] LP Nolte and R Ganz, Eds., *Computer-Assisted Orthopedic Surgery (CAOS)*, Hogrefe and Huber Pub., 1st edition, 1999.

[2] JAK Ohnsorge, M Weisskopf, E Schkommodau, JE Wildberger, A Prescher, and CH Siebert, "The benefit of computer-assisted kyphoplasty," in *4th An. Meeting of CAOS-International*, 2004, pp. 367–370.

[3] T Hüfner, D Kendoff, J Geerling, V Look, C Krettek, and M Citak, "Biomechanical stability of different attachment arrays of reference markers in navigated orthopaedic surgery," in *6th An. Meeting of CAOS-International*, 2006, pp. 225–226.

[4] ML Swank, "Low incidence of post operative complications due to reference array pin placement in image free computer assisted total knee replacement," in *6th An. Meeting of CAOS-International*, 2006, pp. 513–515.

[5] DC Barratt, GP Penney, CSK Chan, M Slomczykowski, TJ Carter, PJ Edwards, and DJ Hawkes, "Self-calibrating 3d-ultrasound-based bone registration for minimally invasive orthopedic surgery.," *IEEE Trans Med Imaging*, vol. 25, no. 3, pp. 312–23, 2006.

[6] BKP Horn, "Closed-form solution of absolute orientation using unit quaternions," *Journal of the Optical Society of America A*, vol. 4, no. 4, pp. 629–642, April 1987.

[7] KS Arun, TS Huang, and SD Blostein, "Least-squares fitting of two 3-d point sets," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 9, no. 5, pp. 698–700, 1987.

[8] MH Moghari and P Abolmaesumi, "A novel incremental technique for ultrasound to ct bone surface registration using unscented kalman filter," in *Medical Image Computing and Computer-Assisted Intervention Proceedings*, 2005, vol. 3750, pp. 197–204.

[9] PJ Besl and ND McKay, "A method for registration of 3-d shapes," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 14, no. 2, pp. 239–256, 1992.

[10] B Ma and RE Ellis, "Surface-based registration with a particle filter," in *Medical Image Computing and Computer-Assisted Intervention Proceedings*, 2004, vol. 3216, pp. 566–573.

[11] GP Penney, DC Barratt, CSK Chan, M Slomczykowski, TJ Carter, PJ Edwards, and DJ Hawkes, "Cadaver validation of intensity-based ultrasound to ct registration," *Medical Image Analysis*, vol. 10, no. 3, pp. 385–395, 2006.