

Registration of Planar Film Radiographs with Computed Tomography

Lisa M. Gottesfeld Brown *
lisab@watson.ibm.com
I.B.M T.J. Watson Research Center
Hawthorne, NY 10532

Terrance E. Boulton
teb6@Lehigh.edu
LeHigh University
Bethlehem, PA 18015

Abstract

In this paper we describe a method to register Computed Tomography (CT) data with planar film radiographs. Previous methods applied to the problem of CT-radiograph registration rely on determining the correspondence between occluding contours of the 3D surface in the CT data with 2D contours in the projection image. These methods implicitly assume that the correspondence is accurate, ignoring fundamental nonlinear differences in the underlying measurements. In contrast, our emphasis has been to directly exploit the relationship between imaging devices. This is performed by registering radiograph data with intensity-corrected simulated radiograph data derived from CT measurements. We will show that by exploiting the physical relationship between CT and radiograph measurements we can significantly improve registration accuracy. Concomitantly, we detail the relationship between CT and radiograph measurements and the primary factors influencing discrepancies between simulated and real radiograph data.

1 Introduction

Registration of medical data from different imaging devices has proven to be an important tool for extracting additional information for diagnosis, therapy, and surgery. For example, high resolution, three dimensional, structural medical images, such as data from X-ray computed tomography (CT) and magnetic resonance imaging (MR), are capable of clearly delineating many anatomical structures. These images may be taken prior to a surgical operation for diagnosis and localization. The surgeon may use such images to plan a surgical procedure. Then during surgery, 2D ultrasound images or fluoroscopy data may be used to guide the surgeon through his plan. To effectively execute the plan, the intra-operative images need to be registered with the pre-operative

images. However, the poor resolution of the intra-operative images, the differences in the sensor characteristics, and the changes in the patient's position and state make this a challenging task.

Historically, this registration has been performed using stereotactic frames, external markers, and 3D positioning devices. These interventions pose a burden on the surgeon and patient and limit the accuracy and generality of the registration. The ideal solution is anatomy-based patient registration. Several methods have been proposed to perform this type of image-guided registration, but in general, these methods are sensor-independent and do not address one of the most significant distortions between data sets: distortions due to differences in sensor measurements.

This paper specifically addresses the problem of registering X-ray CT data with planar film radiographs. Our emphasis has been to directly exploit the relationship between imaging devices. This is performed by registering radiograph data with intensity-corrected simulated radiograph data derived from CT measurements. We will show that by utilizing the physical relationship between CT and radiograph measurements we can significantly improve registration accuracy. Concomitantly, we detail the relationship between CT and radiograph measurements and the primary factors influencing discrepancies between simulated and real radiograph data. This relationship is useful, not only in improving registration, but also to enhance our understanding of the measurements of the individual modalities, their distortions and sensor-dependent information. We also intend to use this relationship to aid our understanding of the more complex relationship between CT and fluoroscopic images.

We start off by describing previous work in this area and the strategy we employed to improve the accuracy currently achieved. In Section 3, we describe the acquisition of our data and the calibration procedure we implemented to evaluate our registration technique. In Section 4, we examine the relationship between radiographic and CT data. CT data is used to simulate radiographic data and the factors which influence the differences between real and simulated data are

*This work is supported in part by NSF CISE grant #CDA-90-24735, in part by NSF Presidential Young Investigators Awards program #IRI-90-57951, and in part by ARPA contract DACA-76-92-C-007.

discussed.

Sections 5 and 6 describe the registration technique and the results. The registration method uses the results of Section 4 to exploit the full sensor relationship. The transformation between a single radiograph and a 3D CT data set is determined using the simulated radiographs generated from the CT data. The optimal transformation is the X-ray system configuration which generated the simulated radiograph which is maximally correlated with the original radiograph. An error analysis is conducted based on the results of the calibration study. In the last section, we summarize the results of this research and offer suggestions for future work.

2 Related Work

Registration methods designed for determining the X-ray system parameters which relate CT data with one or more X-ray projection images, can be split into two major categories. In the first category, surfaces in the CT data are initially extracted. Then, corresponding features in X-ray images, such as edges or contours, are found. The registration transformation is then found by minimizing a cost function which evaluates the proximity between the projected 3D surface and the 2D contours. This strategy was used by [7] and [3]. In [8] a similar approach was taken except only a small number of anatomical features, namely, boundaries along skull landmarks, were used.

These methods presume the correspondence between CT data and X-ray projection images is inherent, or minimally, that a prior stage has performed some kind of calibration to make this presumption true. The underlying assumption is that surfaces found in the CT data that are tangent to perspective rays of the X-ray system configuration, will correspond to edges in the projection images. There are several limitations of this approach and its general implementation. In particular, the formal ray integral equation is not used. The fact that radiographic data is a negative exponential of an integral is ignored. Similarly, the different sensor resolutions and measurement sensitivities are overlooked. Only the simpler relationship, between occluding contours of 3D surfaces and their projections onto the radiograph image, is used. Information is limited to surface boundaries [7, 3] or feature extraction [8]. Grey value information is only used at the level of 3D segmentation and typically this difficult problem is solved independently to the registration problem. Usually only specific anatomic structures are found. This has advantages that only structures which are known to be rigid are used. However, contours in projection images do not necessarily arise from 3D surface tangents. View-dependent supposition of anatomic structures may effect projection contours in complex ways. When ray integration is not performed, only a small subset of the original grey value information is utilized. In structure/contour matching

methods, outliers are often a significant source of error.

Because the underlying relationship between the two data sets is not exploited, the relationship between the resolution of the two data sets is also not exploited. Typically 3D segmentation is performed without regard to either CT or X-ray detector resolution, or the particular X-ray system configuration. Surfaces are improperly smoothed and simplified. Both over and under-sampling are common and partial volume effects cannot be modeled. Methods are often tested on simulated data or verified by dependent information. In [6] video camera images were used to simulate fluoroscopy. In [3] accuracy is measured in average distance of matched points, which is not necessarily conclusive. These methods are advantageous in that they are potentially fast and less sensitive to changes in patient state or sensor domain. They perform well in applications where the sensor differences between images are minimal. If sensor differences are significant, they may be usefully applied for fast initial estimate prediction.

The other major approach to this problem uses voxel or pixel similarity and includes our method. Methods which use this strategy do not rely on higher level extraction of features or regions. Such extraction is typically application-dependent, sensor-dependent, and often only semi-automated, and therefore subjective. Examples of multimodal registration methods of this kind are given in [12, 5]. In [12] a filtering technique is used to find "sensor-invariant" ridges in image space. Correlation is then used to find the optimal transformation between data sets. This is successfully applied to MR and CT data. In [5] MR and CT data were also registered. Instead of maximizing a correlation function, they minimize the coefficient of variation of intensity ratios between the two images and devise a similarity measure sensitive to this metric. Both investigations attempt to find features and matching metrics which use the full data set but are invariant to sensor differences. However, both investigations are general-purpose; neither exploits the particular sensors or sensor relationship.

In our approach, voxel/pixel similarity is performed by simulating radiographic data and then optimizing the match between simulated and real radiographs. Two other research teams have been investigating this approach concurrently with our own. In [11, 1] image-guided radiosurgery is performed by correlating an orthogonal pair of radiographs with precomputed radiographs. The details of this project have not been published and are proprietary[1a]. In [9] registration was performed of CT data with a stereo pair of radiographs. Their work closely follows our own. However, our objective has been to scrutinize the underlying relationship between the radiographic and CT data, to achieve the maximal accuracy in registration, while the research of [9] was directed more towards achieving a feasible solution which could be implemented and tested more efficiently. They did

not directly consider the intensity and resolution relationship between simulated and real radiographs. On the other hand, they present a more elaborate procedure for finding an initial estimate and for more efficiently optimizing the correlation between the CT data and a stereo pair of radiographs.

3 Experimental Setup

In our experiment, planar X-ray film was taken of a dry femur bone rigidly attached to a precisely machined calibration object. The calibration object was composed of two 3/8 inch plexiglass sheets, attached at their ends at a right angle. Each sheet contained a grid of 6x15 embedded stainless steel balls with 5/32 inch (3.97mm) diameter.

X-rays were taken with a Siemens SIRESKOP4 with 125 KV tube voltage, at roughly 40 inches between the X-ray source and the screen film cassette. The film was Fuji Super HR-G 14x17 inch double emulsion and the intensifier screen was a Kyokko GH1 made by Kasei Optonix Ltd. The X-ray film was then digitized using IBM's Time-Delay-and-Integration Imaging System. This system is a high resolution digitizing scanner which can capture images at a spatial resolution of 3072x4096 pixels with a dynamic range of 69.65 DB and state-of-the-art noise suppression. The digitized images were roughly 2000x3000x12-bits (only 8-bits were ultimately used) with pixel sizes on the order of 0.078x0.075 mm². X-ray computed tomography data of the same femur/calibration object was taken on a GE9800 scanner. The CT data was stored in a 158x512x512x8-bit matrix, with slice thickness 3mm, and pixel size 0.390625x0.390625 mm².

To test our registration method, the markers of the calibration device were used to determine the parameters of the X-ray system. To perform this registration, we used the calibration procedure described in [8] and detailed for this application in [2]. This calibration determines, in addition to the rigid transformation between the CT and radiograph world, the intrinsic parameters of the X-ray system: the projection image x and y scale factors and offsets. The scale factors are implicitly related to the focal length of the X-ray system, or in other words, the distance between the X-ray source and the film. The root mean square error for the final calibration was less than 0.2 mm for both configurations. Monte Carlo simulations in which the data were perturbed using normal deviates, show that for Configuration 1, the calibration error is greatest along the optical axis. This suggests a possible error in the focal length.

4 The Relationship between Radiographs and Computed Tomography

In order to develop better methods to register radiographs with X-ray CT, we have simulated radiographs from CT data. The details of the radiograph simulation are given in [2]. In this section, we examine the intensity relationship between radiographs and simulated radiographs computed from CT. Understanding this relationship provides an important link which we will exploit to improve the methodology for the registration of radiographic and CT data. Several investigators have studied methods for registering CT and radiographs or X-ray fluoroscopy images but to our knowledge, the empirical relationship has yet to be delineated.

In the outputs shown in the top of Figures 1-2, simulated and real radiographs are shown for two data sets. The intensity values of the real radiograph represent the digitized values of the intensity transmitted by the film. The intensity values of the simulated radiograph represent the ray sum of the CT values along the ray from the X-ray source to the radiograph plane. These were computed from sufficiently small, uniformly spaced samples along the ray through the CT data, using trilinear interpolation. In this initial evaluation, we have not yet exploited the complete physical relationship between CT values and digitized radiograph measurements. Since the film is darkest where the most light strikes it, and the ray sum, by itself, is just the sum of attenuation, the images appear in many ways similar - both are brightest where the least amount of X-rays were transmitted. However, as we will show in this section, there are several important factors that can be modeled to improve this similarity.

We have determined four primary factors effecting the intensity relationship between the simulated and real radiographs. They are from the:

1. formal intensity relationship: the ray integral equation, linearly-scaled measurements, initial and final intensities, and film characteristics,
2. differences in domain,
3. limitations in CT resolution and partial volume effects,
4. variations in X-ray source spectrum and the Heel Effect.

We will describe each factor, its effect on the intensity relationship, and how it effects the design of a registration method for radiographic and CT images.

4.1 Formal Intensity Relationship

The initial simulations shown in the top right of Figures 1-2 are based on the simple ray sum of the CT values. However

this does not fully model the physical generation of digitized film measurements. In particular, assuming a narrow, monochromatic X-ray beam passing along a ray from the X-ray source to the detector, the relationship between the input and output intensity, is the ray integral of the linear attenuation coefficients $\beta(x, y, z)$:

$$I_{out} = I_{in} e^{\left\{-\int_{source}^{detector} \beta(x, y, z) dx dy dz\right\}}$$

Film effectively measures the transmitted intensity, I_t . The optical density (*OD*) of the film is defined to be the $\log(I_{in}/I_t)$ and this is linearly proportional to the log of the relative exposure, for the straight-line portion of the H-D or characteristic curve of the film. Since exposure is intensity over time, relative exposure is equivalent to the relative intensity for the same exposure time. Thus, $\log(I_{in}^{radiographic}/I_t) \propto \log(I_{out}/I_{in}^{CT})$, where \propto implies linear proportionality. Because of the large bandwidth of our digitizer, we believe our digitized measurements are linearly related to the transmitted intensity of the film. Similarly, lack of CT calibration, bit reduction, and unit conversions from linear attenuation to Hounsfield units, are also all linear. Therefore we can derive the following relationship between the measured radiographic intensities, I_r and the measured attenuation coefficients β_m , of our CT data,

$$\log\left(\frac{1}{a_1 I_r + b_1}\right) = a_2 \log\left(a_3 \exp\left(-\sum_{ray} [a_4 \beta_m + b_4]\right)\right) + b_2 \quad (4.1)$$

where the constants a_i, b_i are from the following linear relationships:

[**a₁**, **b₁**] Optical density is the log of the inverse of the transmitted intensity *relative* to the intensity incident on film, and the digitization is linearly related to transmitted intensity.

[**a₂**, **b₂**] For the straight-line portion of the H-D curve of the film, the optical density is linearly related to the log of the relative exposure.

[**a₃**] The relative exposure, and therefore the relative intensity incident on the film depends on the initial intensity emitted from the source.

[**a₄**, **b₄**] The numbers output by the CT scanner are linearly related to the linear attenuation coefficients used in the line integral equation. This may be due to the need for calibration, conversion in units, or bit reduction. The constant offset b_4 can probably be ignored, since β_{air} is zero.

Equation(4.1) can be rewritten as,

$$\ln(I_r + k) \propto \sum_{ray} \beta_m$$

where k is a constant. The derivation and relations between k and the constants of linearity, and the constants a_i, b_i are given in [2].

Furthermore, if we would like to clip intensity values which are nonlinearly related because they are above or below the straight-line portion of the H-D curve, we can add the following cut-offs:

$$\text{if } \sum_{ray} \beta_m > c_l \text{ then } I_{sim} = c_l \text{ else } I_{sim} = \sum_{ray} \beta_m$$

$$\text{if } \sum_{ray} \beta_m < c_u \text{ then } I_{sim} = c_u \text{ else } I_{sim} = \sum_{ray} \beta_m,$$

where I_{sim} replaces $\sum_{ray} \beta_m$ in the previous equation. This is applicable only if we have a suitable signal to noise ratio. Notice, because the sum of the attenuation coefficients is inversely related to the amount of X-ray transmission the lower cut-off operates like an upper bound and similarly, the upper cut-off acts like a lower bound. For our radiographs, i.e., for the H-D curve of our film, only the upper cut-off c_u appears to be relevant.

We have written the equation in terms of a transformation of our radiographic data to our “simulated” radiograph for two reasons. First, since the resolution of our radiographic data is superior to our CT data, we would like to perform our calculations using the better data. Secondly, from a computational point of view, we would like to perform correction one time on the true radiograph, rather than for each simulation in the correlation tests. On the other hand, for visualization only, it may be useful to perform the inverse transformation, since there is greater familiarity with real radiographs.

4.2 Differences in Domain

A potentially significant difference between simulated and real radiographs can arise because of changes or differences in the patient or environment. Even in our controlled environment, where the environment was fixed, these differences were manifested in Configuration 2. The calibration object was not completely in the field of view of the CT scanner. As a consequence, both ends of the calibration plate are visible in the actual radiograph and not in the simulation. Domain differences are inevitable. Registration methods designed for this application must be able to withstand substantial domain differences to be effective.

4.3 Resolution Limitations

The third factor effecting the intensity relationship between the simulated and real radiographs arises from the limitations in CT resolution and the effect this has on the simulated radiograph. This is pertinent, not only in avoiding the matching

of blurred data with data at a higher resolution, but also in needlessly matching more data than can be usefully considered. In the simulated radiographs of the previous section, vertical blurring is apparent because, the vertical image axis, \vec{v} , is roughly aligned with the z-axis, or slice dimension, of the CT data.

To compensate, in general, for this type of resolution discrepancy, the resolution limits of the simulated radiograph is determined based on the resolution of the CT data, $(res_x^{ct}, res_y^{ct}, res_z^{ct})$ in mm/pixel, and the initial estimate of the X-ray configuration for computing the simulation. Assuming orthogonal projection, the optimal resolution for an image whose optic axis is aligned with one of the axes of the world, is simply the resolution of the world in this direction. When the image intersects the world along some direction $\vec{u} = (u_x, u_y, u_z)$, we would like the resolution to reflect the projection of the world resolution onto this directional vector. We model this effect by assuming the world resolution is elliptical with principal axes aligned with the sampling axes and whose magnitudes are related to the respective half resolution of each axes. In particular, let the resolution of the 3D world be represented by the ellipsoid:

$$\frac{x^2}{(res_x^{ct}/2)^2} + \frac{y^2}{(res_y^{ct}/2)^2} + \frac{z^2}{(res_z^{ct}/2)^2} = 1.0$$

The resolution in the 2D projection space, is defined to be the diameter of the ellipsoid in each projection direction, (\vec{u}, \vec{v}) . This can be computed by finding the intrinsic parameters t_u and t_v such that the points $(u_x t_u, u_y t_u, u_z t_u)$ and $(v_x t_v, v_y t_v, v_z t_v)$ are on the ellipsoid. The resolution of the 2D projection space is then the diameters of the ellipsoids at these points,

$$res_u^{image} = 2 \left\| \frac{1}{\frac{(u_x t_u)^2}{(res_x^{ct}/2)^2} + \frac{(u_y t_u)^2}{(res_y^{ct}/2)^2} + \frac{(u_z t_u)^2}{(res_z^{ct}/2)^2}} \right\|$$

$$res_v^{image} = 2 \left\| \frac{1}{\frac{(v_x t_v)^2}{(res_x^{ct}/2)^2} + \frac{(v_y t_v)^2}{(res_y^{ct}/2)^2} + \frac{(v_z t_v)^2}{(res_z^{ct}/2)^2}} \right\|$$

where $\| * \|$ is the second norm. An elliptical resolution is reasonable, since the resolution along directions other than the principal axes should reflect the ability to interpolate using the values along both axes. This is similar to the multivariate case in which we know we have white noise with a given standard deviation along each axes; the standard deviation in 2D is elliptical. One might consider the resolution in 2D to be the resolution of the pixel quad. However, when the projection direction runs diagonally between the two resolution axes, this suggests that the resolution actually decreases, since the Euclidean distance of the diagonal of the quad is its largest cross-section. This does not take into

account the progressively more information that is obtained by neighboring pixels as the projection direction moves between the original orthogonal sampling directions for the case of trilinear interpolation.

In the real case, we do not have orthogonal projection, although, as previously discussed, the distance between the X-ray source and detector is relatively large compared with distances in the image. More importantly, the distance between rays upon entering the CT-data and exiting must be similar. This depends on the dimensions of CT data relative to the X-ray configuration, and ultimately on the size and location of the body. Generally a patient is placed as close to the detector plane as possible to minimize perspective distortion. We have chosen to ignore these complications and to use the worst case estimate based on orthogonal projection.

We would like to make one comment. This resolution limitation reflects the limitations of the CT data to clearly "see" 3D objects which are smaller than the sampling size. This is evident with our data set; it was difficult to accurately discern the location of the markers whose diameter size was only slightly greater than the slice spacing. This in turn, limits the resolution of the simulated radiograph. However, because the simulated radiograph is composed of ray sums, and these sums pass through the various voxels at different locations, in cases where there are multiple objects or objects whose boundaries are crossed multiple times in any one ray, the radiograph may provide information at a higher resolution than is reflected in the analysis here.

Alternatively, we could attempt to improve the resolution of the simulated radiograph using more sophisticated techniques to interpolate the CT data. Our approach has several advantages. Improving the interpolation requires assumptions that depend on the object characteristics and the acquisition. Furthermore, since our proposed matching scheme relies on simulating many projections, we would like to limit the size and time required to simulate each radiograph. Finally, depending on the assumptions that we are able to make, it is not clear to what extent we can add real information.

Limitations in the resolution of the projections used to reconstruct CT data may cause an artifact known as the partial volume effect[4]. In our domain, this may be the cause of an under-estimation of the intensity of the markers in the CT data, since the marker size is on the order of magnitude of the slice spacing. This should also be corrected for by the limited resolution of the simulated projection which re-enacts the same type of averaging as occurred originally. However, it is also possible, that the marker brightness is due to a greater sensitivity of the film to high attenuation than the CT scanner, or a consequence of the different energy levels used in the two mediums.

4.4 Variations in X-ray Source and the Heel Effect

The fourth primary factor effecting the intensity relationship between the real and simulated radiographs is apparently due to differences in the energy spectrums of the X-ray sources of the two modalities. This is extremely difficult to quantify. Since the local attenuation coefficient varies with energy, we need more information than we have from a single CT scan to model the effects of different energies. This is further complicated if we would like to consider the full energy spectrum and the changing energy spectrum as the X-rays traverse the medium. It does not appear necessary to model this level of complexity for the purposes of registration.

However, there is significant spatial variation in intensity apparent in the real radiograph that may be more easily explained. In both real radiographs, as compared with their respective simulations, the images appear to get brighter as you move from the bottom to the top of the image. We attribute this to the anode heel effect. In addition to the polychromatic nature of X-ray sources, the intensity of many diagnostic X-ray sources varies from one side of the source to the other. This is a physical property of the X-ray source because of the way it is constructed. Electrons produced by heating a tungsten filament strike a target which absorbs the electrons and emits X-rays. Because the target, or anode assembly, is mounted at an angle to the filament, or cathode assembly, the X-ray intensity decreases from the cathode to the anode side of the beam. This relative decrease may be as large as 25 percent.

This effect is seen vertically throughout the images. It appears more significantly where the bone and markers are present, but it is harder to quantify since its effects depend on the variations in attenuation coefficient with energy. This effect is not evident in the simulated radiograph. We expect that the best way to correct for this phenomena would be through calibration of the X-ray system using film obtained with no objects in the field of view. This information could then be used to replace a_3 in equation[4.1]. However, since our objective is to design registration methods, we note that the effect is, for the most part, monotonic and gradual across the image. This implies that although raw image intensities are effected, metrics such as edge detectors would still be effective since they are invariant to gradual intensity fluctuations.

5 Registration of Radiograph and CT data

In this section, we describe a method to register radiograph and CT data using simulated radiographs. We use the model developed in the previous section to improve the intensity

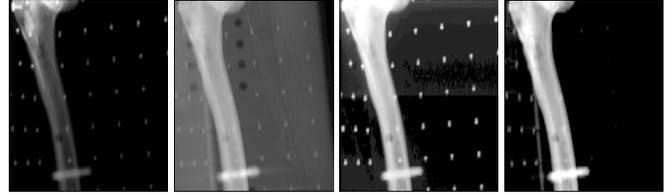


Figure 1: Intensity Matched Real and Simulated Radiographs - Configuration 1, Left image pair is the original and simulated radiographs. Right image pair is corresponding intensity corrected original and simulated radiographs.

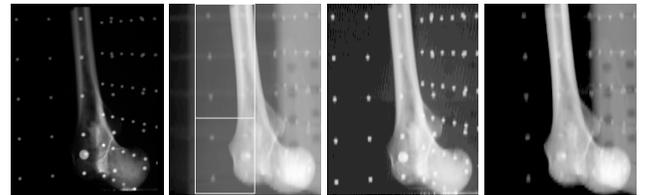


Figure 2: Intensity Matched Real and Simulated Radiographs - Configuration 2

relationship between real and simulated radiographs. We then analyze the accuracy of this approach compared with the calibration results. Finally, we examine the robustness of our approach by testing the sensitivity to intrinsic parameters and resolution and by examining the smoothness of our cost function over the search space.

5.1 Intensity Correction

The right pair of images in Figures 1-2 show the results of correcting the intensity relationship and resolution between the original and simulated radiographs for the two configurations using the formulations described in the previous section. The left pair of images are the original (far left) and simulated (second from left) radiographs as discussed in the previous section. On the right, the intensity matched original and simulated radiographs are shown, including the film cut-off and sampling of the original radiograph. Notice how segmentation performed without intensity matching would have a difficult time matching the femur outlines in both images. Even if adaptive thresholding is capable of finding the femur in the original real radiograph, which is possible, this outline is narrower than the outline in the simulated radiograph because the peak of the intensity gradient is effected by the nonlinear intensity scaling. This is also clearly seen in the marker size, which changes from the original to the intensity-scaled version. In this case, the resolution matching also plays a role, since the vertical resolution in the real radiograph is reduced.

Figure 3 shows 2D histograms of the intensity relationship between the original and simulated radiographs. In these

histograms, the intensities in each radiograph are linearly scaled to 8 bits. The two images are matched point-wise and a histogram of the number of points which have each pair of intensities, (I_{orig}, I_{sim}) is computed, where I_{orig} is the intensity in the original radiograph, and I_{sim} is the intensity in the simulation. The plots show this histogram, by indicating with color, the number of points with each intensity pairing. The smallest number is represented as dark blue and the largest by red.

In the top of Figure 3 the 2D histograms for Configuration 1 are shown. The left plot shows the original histogram before intensity correction, while the right shows the histogram after correction. Although significant information is lost due to the relatively poor resolution of the CT data, the corrected histogram is significantly more linear than the original.

In the lower portion of Figure 3 the 2D histograms for Configuration 2 are shown. The middle images show the original and corrected histograms as before. However, because of the differences in domain, the corrected histogram has a “shadow.” We verified this by computing the histogram for the part of the image without this problem. The histogram is shown in the lower left; the image section we used is shown by two boxes outlined on the simulated X-ray in Figure 2. The histogram still appears to saturate - the original radiograph seems to have higher intensities which are not represented in the simulation. We hypothesize that this is a manifestation of the Heel Effect. Again, we took only a subset of the images - this time limiting the image to lower box shown in Figure 2, the part of the image near the center of the image detector. The histogram for this subset is shown in the lower right. The results support our hypothesis. It is also interesting to infer from the 2D histogram the relative *intensity* resolutions of the two data sets. Notably, the simulated radiograph contains significantly less information.

5.2 The Similarity Metric

Since domain differences between images are potentially unavoidable, edges are extracted from the intensity matched radiographs. Images are normalized and filtered using a horizontal and vertical 3x3 Sobel filter. The output of this edge detection is two grey values, $[G_x, G_y]$, indicating the gradient in each direction. Gradient correlation is then performed using the following equation,

$$\sum_i (G_x^o(i) * G_x^s(i) + G_y^o(i) * G_y^s(i))$$

where the superscripts indicate (o)original or (s)imulated image, and i is a spatial index into the images. It is not correct to take the magnitude of the gradient dot product, since a negative gradient product indicates opposite gradient directions and therefore dissimilarity. On the other hand, since domain differences are expected, we do not want to penalize

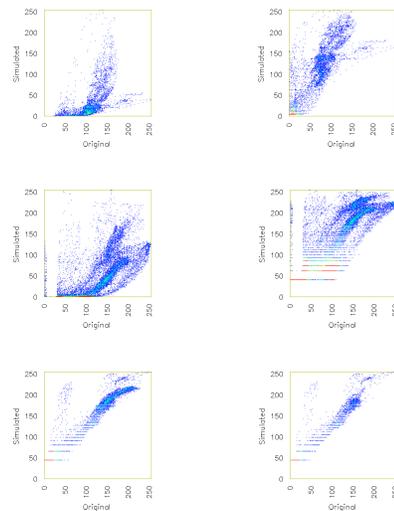


Figure 3: 2D Histograms of Intensity Relationship: Configuration 1, Top: original intensity relationship (left), corrected intensities (right) Configuration 2, Middle: original intensities (left), corrected intensities (right). Bottom: eliminating image portions with different domains (left), eliminating upper image where heel effect occurred (right).

a match because of the degree to which gradients do not agree. Therefore, we accumulate gradient product only if the product is positive indicating some similarity in the gradients at that location. Finally, we normalize our measure based on the sum of the gradient products of each image by itself, to obtain the absolute correlation coefficient.

To find the optimal rigid transformation, the maximum gradient correlation between the original and simulated, intensity-corrected radiographs was searched using Powell’s multidimensional direction set method [10]. We assume that a good initial estimate is available. This is reasonable since surface/contour matching techniques and semi-automatic implementations have successfully and efficiently been able to perform this step. In addition, we found that initializing the directions so that the translational parameters are searched along the three axes of the projection space was advantageous. This was implemented using the initial estimate to compute the projection space axes.

An important parameter of Powell’s method is the stopping criteria which signals that the algorithm has converged. In this case, failure to decrease the gradient correlation by some fractional tolerance, is used to determine whether to stop. For this reason, the absolute correlation coefficient is used, so that the magnitude of the fractional tolerance is meaningful.

| | Perturbation 1 | | | Perturbation 2 | | | Perturbation 4 | | |
|-------------|----------------|-------|--------|----------------|-------|-------|----------------|-------|-------|
| | Initial | Final | | Initial | Final | | Initial | Final | |
| | | .1 | .0001 | | .1 | .0001 | | .1 | .0001 |
| Mean | 13.0 | 2.6 | .96 | 25.9 | 3.00 | 1.91 | 51.5 | 111.4 | 5.2 |
| x | 9.3 | 0.17 | .27 | 18.3 | 0.34 | 0.54 | 35.6 | 2.9 | 1.0 |
| y | 9.0 | 0.70 | .83 | 18.5 | 2.48 | 1.71 | 37.3 | 20.9 | 5.1 |
| z | 0.8 | 2.49 | .40 | 1.8 | 1.67 | 0.67 | 4.2 | 109.4 | 0.7 |
| Tx | .48 | 6.0 | .89 | 1.2 | 14.4 | 1.8 | 2.4 | 64.5 | 44.7 |
| Ty | .52 | 8.8 | .35 | 1.0 | 2.5 | 0.6 | 1.6 | 8.5 | 2.1 |
| Tz | .65 | 1.2 | .25 | 1.1 | 2.0 | 0.4 | 2.9 | 111.9 | 4.2 |
| θ_x | .018 | .001 | .0003 | .041 | .002 | .0005 | .07 | .004 | .004 |
| θ_y | .013 | .005 | .0006 | .055 | .015 | .0001 | .06 | .048 | .048 |
| θ_z | .019 | .008 | .0006 | .034 | .004 | .0012 | .07 | .041 | .001 |
| Correlation | | .9985 | 1.0000 | | .9983 | .9999 | | .8597 | .9922 |

Table 1: 3D and Parameter Error for Low Resolution No Noise Case

6 Results

To evaluate the error in the registration, we computed the 3D positional error for random points in the 3D coordinate space viewed by the radiograph. For each configuration, we used a cubic region in CT or world space, centered at the femur at the vertical position in the femur which was in the vertical center in the radiograph. The assumption is that the radiograph used for 3D localization views the object of interest. The cubic region was 10cm on each side. We took 100 random 3D points in the cube and compared these with their locations as estimated by the optimization. Finally, we evaluated the components of these errors with respect to the coordinate axes of the projection space.

We performed two types of tests. To test our implementation, its accuracy, robustness to initial estimate, and tolerance to the stopping condition, we performed tests using a simulated projection instead of the actual radiograph. This is the case with no noise except resolution limitations. The results are shown in Table 1. Three studies of the no noise case were conducted. In each study, we reduced the CT data set to (79x64x64) and the radiograph size to (47x39). This corresponds to a CT resolution of (6.04mm,3.17mm,3.17mm) and a radiograph resolution of (3.8mm,6.07mm). For each study, we performed 5 trials with random initial estimates in the range $\pm 1mm$ for translation parameters and $\pm 1degree$ for rotation parameters, for Perturbation 1, ± 2 mm or degree for Perturbation 2, and ± 3 mm or degree for Perturbation 3. For each study we also computed the results for two different stopping conditions, a fractional tolerance of 0.1 and 0.0001 in the correlation. The table shows the resulting 3D and parameter errors.

First, we point out that the method can tolerate significantly more rotational error than translational error. Although not shown here, the results for the same perturbation

cases with reduced error in the initial estimate of the rotational parameters are very similar, except the initial 3D positional errors are much smaller. Notice also, since we have relatively large rotational error, the initial error in X and Y is significantly greater than the error in Z, the optic axis.

The next observation we would like to make is that the parameter error is not a useful measure in this case. This makes sense, since the origin of the projection space coordinate system lies at the X-ray source, whose position can be moved by a relatively large factor without the same effect on localization error. Lastly, we observe, that the method is very sensitive to the range in the initial estimate and this can be partially overcome by decreasing the fractional tolerance used as a stopping condition. However, this implies that the method is vulnerable to noise in the data, which will make it difficult to obtain the optimal solution as the size of the perturbation of the initial estimate increases. Notice, the error in the 3D components of the localization error for the larger perturbations reflects the resolution limitations in the data - in this case, the y-axis is roughly aligned with the CT slice or vertical radiograph spacing which was on the order of 6mm, and the z-axis is the optical axis for which a single radiograph contains little information.

Table 2 show the results of the registration method applied to the real radiographs and the full CT data set. As we discussed in the study using simulated data, i.e. data with no noise, only small perturbations in the initial estimate are acceptable. Furthermore, high accuracy in the correlation is necessary; for these tests we set the fractional tolerance to 0.001, and the initial estimate perturbation size to ± 2 mm or degrees. For initial estimates greater than ± 4 mm or degrees, the method did not typically reach the optimal solution. We assume here that a good initial estimate is given, either by re-initializing Powell's method, by implementing hierarchical

| | Configuration 1 | | Configuration 2 | |
|-------------|-----------------|-------|-----------------|-------|
| | Initial | Final | Initial | Final |
| Mean | 25.3 | 2.82 | 25.9 | .78 |
| x | 20.6 | .58 | 18.3 | .61 |
| y | 14.1 | 1.66 | 18.2 | .44 |
| z | 3.5 | 2.20 | 1.8 | .22 |
| Max | 26.4 | 3.2 | 27.3 | .93 |
| x | 22.7 | .89 | 19.7 | .67 |
| y | 14.1 | 1.91 | 20.1 | .58 |
| z | 4.5 | 2.64 | 2.6 | .27 |
| Correlation | 0.7545 | | 0.6115 | |

Table 2: 3D Error for Full Resolution Radiograph/CT Registration

optimization, and/or using contour/surface matching techniques. For this research, we are primarily interested in the accuracy that can be achieved.

The results are mixed. For Configuration 2, the 3D error is less than 1mm. The optimal correlation was within the tolerance of the correlation for the “true” solution. For Configuration 1, however, the correlation for the “true” solution was only 0.7396, while the optimal correlation was 0.7543, indicating that a simulated radiograph was generated which was more highly correlated than the calibrated solution. This has two potential explanations. First, our calibration results may include error in the intrinsic parameters, thereby corrupting the results here. In fact, our calibration results of the intrinsic parameters from Configuration 1, were indeed questionable. This is corroborated by the error components which are greatest along the optic axis. Second, the error along the optic axis and in the vertical direction, are reasonable given the limitations of a single radiograph with a focal length of 1000mm and vertical spacing of 3mm. This was confirmed by visual inspection of the objective function over translation space; a nearly constant-value ridge occurred along the optical axis. This is shown in Figure 4 in which the axes coming out of the page is T_z . In the no-noise case (upper surface), a peak is barely discernible at the appropriate location (0,0), while for the real radiograph (lower surface), the peak is off center, in the fore-ground, i.e. along a different position in the optical axis.

7 Conclusion

In summary, we have verified that by using intensity-corrected simulated radiographs, a single planar film radiograph can be registered to CT data at high accuracy. For our two test radiographs whose vertical axes are nearly aligned with the CT slice dimension and CT data with slice thickness

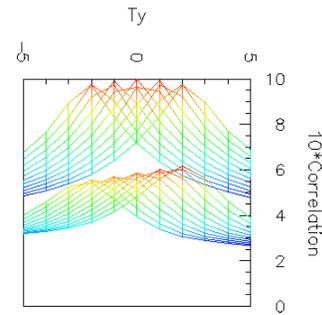


Figure 4: Surface plot of the objective function for Config. 1 over translation space (T_y, T_x). Plot shows the ridge along the optic axis peaking at the calibration solution for the no-noise case (upper surface where absolute correlation values are higher), and peaking at a different position along the optic axis for the real data.

of 3mm, a maximum error of less than 1mm in the horizontal direction, less than 2mm in the vertical direction, and less than 3mm along the optic axis was achieved. We assume intrinsic parameters can be accurately measured and other methods can be used to efficiently find a good initial estimate. Previous methods have relied on multiple radiographs, have been tested only by simulations, and have not exploited the full sensor model.

We have also detailed the relationship between CT and digitized planar radiograph film measurements. We have described a model for simulating radiographs from CT data for registration with real radiographs. To optimize the intensity relationship between simulated and real radiographs the linearity of the measurements, film characteristics, and relations between initial and final intensities are utilized to derive a formal mathematical relationship between measured attenuation and digitized planar film. Other factors influencing the discrepancy between simulated and real radiographs were addressed including the differences in domain, limitations in CT resolution and partial volume effects, and variations in X-ray source spectrum and the Heel Effect.

For future work, we propose to confirm our model and registration method using more test cases including different and more realistic patient data, and other radiographic systems including bi-planar radiographic devices and different film/screen combinations. It would also be useful to corroborate our model with calibration information and possibly pre-calibrate to remove known error sources such as domain and source intensity differences and intensity variations due to the Heel Effect. One interesting potential experiment is to register CT data with scout data from the same scanner.

This eliminates several factors influencing the simulation error such as X-ray source spectrum and some calibration errors but would allow us to evaluate the potential accuracy of the radiograph simulation.

Our ultimate objective is to combine this type of highly accurate but inefficient registration method with a classical contour/surface approach. In this way, we would be able to achieve both the computational speed and accuracy necessary for clinical applications. Lastly, we would like to use the relationship derived for CT and radiographic film to derive the more complex relationship between CT and fluoroscopy.

8 Acknowledgements

We would like to thank Professor Gerald Q. Maguire of the Royal Institute of Technology(KTH)/Teleinformatik, Stockholm, Sweden for significant contributions to this work.

References

- [1] J. R. Adler, "Image-Based Frameless Stereotaxy," in *Interactive Image-Guided Neurosurgery*, ed. R. J. Maciunas, 1993, 81-98.
- [1a] Correspondence with J. R. Adler.
- [2] L. G. Brown, *Multi-Modal Medical Image Registration - Exploiting Sensor Relationships*, Ph.D. Dissertation, Columbia University, 1996.
- [3] J. Feldmar, N. Ayache, F. Betting, "3D-2D Projective Registration of Free-Form Curves and Surfaces," Institut National de Recherche en Informatique et en Automatique, Sophia-Antipolis, France, Rapport de Recherche, No. 2434, December 1994.
- [4] G. T. Herman, *Image Reconstruction from Projections*, Academic Press, Orlando, FL, 1980.
- [5] D. L. G. Hill and D. J. Hawkes, "Medical Image Registration using Voxel Similarity Measures," *AAAI Applications of Computer Vision in Medical Image Processing*, March 21-23, 1994, 34-37.
- [6] S. Lavallée, J. Troccaz, P. Sautot, L. Brunie, P. Cinquin, P. Merloz, and J. Chirossel, "Robot Assisted Spine Surgery," *IEEE Trans. on Robotics and Automation*.
- [7] S. Lavallée, J. Troccaz, P. Sautot, P. Sautot, B. Mazier, P. Cinquin, P. Merloz, and J. Chirossel, "Computer Assisted Spine Surgery," *Proc. of 1st Int'l Symp. on Medical Robotics in Computer Assisted Surgery*, Vol I, Pittsburgh, PA, 1994.
- [8] B. Lee, *Stereo Matching of Skull Landmarks* Ph.D. Dissertation, Stanford University, 1991.
- [9] L. Lemieux, R. Jagoe, D. R. Fish, N. D. Kitchen, and D. G. T. Thomas, "A Patient-to-Computed-Tomography Image Registration Method based on Digitally Reconstructed Radiographs," *Medical Physics*, **21**, No. 11, November 1994, 1749-1760.
- [10] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flanner, *Numerical Recipes in C*, Cambridge University Press, 1992.
- [11] A. Schweikard, R. Tombropoulos, J. R. Adler, and J-C. Latombe, "Planning for image-guided radiosurgery," *Applications of Computer Vision in Medical Image Processing*, March 21-23, 1994, 96-101.
- [12] P. A. van den Elsen, "Retrospective Fusion of CT and MR Brain Images Using Mathematical Operators," *AAAI Applications of Computer Vision in Medical Image Processing*, March 21-23, 1994, 30-33.