Long Bone Panoramas from Fluoroscopic X-ray Images

Ziv Yaniv, *Student Member, IEEE,* Leo Joskowicz, *Senior Member, IEEE,*

Abstract— **This paper presents a new method for creating a single panoramic image of a long bone from several individual fluoroscopic X-ray images. Panoramic images are useful preoperatively for diagnosis, and intraoperatively for long bone fragment alignment, for making anatomical measurements, and for documenting surgical outcomes. Our method composes individual overlapping images into an undistorted panoramic view that is the equivalent of a single X-ray image with a wide field of view. The correlations between the images are established from the graduations of a radiolucent ruler imaged alongside the long bone. Unlike existing methods, ours uses readily available hardware, requires a simple image acquisition protocol with minimal user input, and works with existing fluoroscopic C-arm units without modifications. It is robust and accurate, producing panoramas whose quality and spatial resolution is comparable to that of the individual images. The method has been successfully tested on in-vitro and clinical cases.**

Current orthopedic practice heavily relies on fluoroscopic X-ray images to perform a variety of surgical procedures such as fracture reduction, joint replacement, osteotomies, and pedicle screw insertion, to name a few. Surgeons use fluoroscopic X-ray images acquired during surgery with a mobile fluoroscopic C-arm to determine the relative position and orientation of bones, implants, and surgical instruments. While inexpensive and readily available, X-ray fluoroscopy has several important limitations, including a narrow field of view, limited resolution and contrast, and geometric distortion. These limitations require the surgeon to frequently acquire images of the surgical situation and to mentally correlate them. They preclude precise evaluation and measurements across images and complicate the placement of long implants. This leads to positioning errors, cumulative radiation exposure to the surgeon, and sub-optimal results in a small but non-negligible number of cases.

While modern X-ray units incorporate geometric distortion correction and contrast enhancement, only a handful address the narrow field of view issue, and only Siemens' is designed for intraoperative use [1], [2], [3]. In this paper, we describe a novel, simple, and inexpensive method for creating a single panoramic image from a few individual fluoroscopic X-ray images. Our method produces an undistorted panoramic view, which is the equivalent of a single X-ray image with a field of view several times wider than that of the individual images by finding correlations between individual overlapping images and composing them (Fig. 1). It uses standard, readily available, and inexpensive hardware: a video frame grabber, a standard PC, a dewarp grid, and a radiolucent X-ray ruler
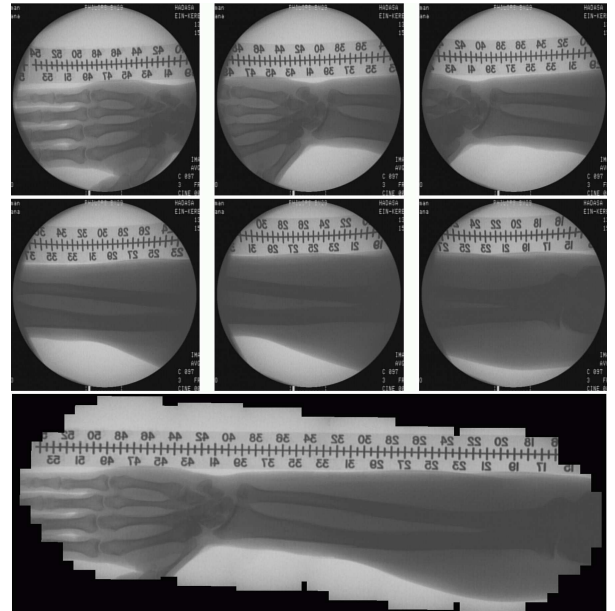
Fig. 1. Panorama construction of a humerus from six fluoroscopic X-ray images. The top two rows show the original images with the orthopedic ruler on the top. The bottom image shows the resulting panorama.

with graduations commonly used in orthopaedics. Our method can be used with existing fluoroscopic C-arm units without any modification, involves a simple imaging protocol, and requires minimal user input. Unlike existing methods, ours produces panoramic images which can be used during surgery.

Distortionless fluoroscopic X-ray panoramic images can be very useful in many preoperative and intraoperative situations. Preoperatively, they can be used for diagnosis and measurements, replacing expensive special purpose film-based X-ray machines with custom cassettes. They are readily available and require less radiation than conventional film X-ray images. Intraoperatively, they can be used in long bone surgery, to determine the mechanical axis of the bone, to align bone fragments, and to measure extremity length and anteversion. In joint replacement surgery, they can be used to assess the position of long implants, such as hip implants, and to determine the length of intramedullary nails. All these require the presence, in the same image, of relevant anatomical features, such as the condyles, the femur head, and the femur neck. Panoramas are also useful to document surgery outcomes. These measurements and images are difficult or impossible to obtain with existing methods and can help to improve diagnosis, shorten surgery time, and improve outcomes. A detailed analysis of orthopaedic applications shows that an accuracy of $\pm 2mm$ over $500mm$ is sufficient in most cases.

I. PREVIOUS WORK

The creation of image panoramas, also called image mosaicing, has a long history and is an active area of research in computer graphics and computer vision [4], [5], [6], [7], [8]. Panoramic images are created by correcting individual images for distortion (when necessary), aligning them, and composing them. The most challenging step is image alignment. The main technical issues are the number of images and the amount of overlap between successive images, the geometric constraints on camera poses, the type of mapping between images, and the identification of common image features. Most existing methods assume many closely related images, usually obtained from a video movie in which consecutive images are nearly identical. However, for medical images, identifying common features in two consecutive images is difficult, as anatomical features are hard to locate reliably and accurately.

Two systems for creating panoramas from X-ray images have been developed. The first uses a special purpose digital X-ray machine that acquires 30–40 partially overlapping images by simultaneously translating the X-ray source and the image intensifier over the patient [9], [10]. The 40-80x512 pixel images are processed by an EasyVision workstation, which merges them into a 2400x512 pixel panoramic image. The system was developed to measure spine scoliosis [11] and is also useful in imaging the colon and the limbs. A quantitative study of X-ray dose and image parallax is described in [10]. The advantages of this method are that it produces high quality undistorted panoramas with little parallax, that it has local exposure correction, and that it is suitable for any part of the body. However, it requires costly special hardware in its own suite and cannot be used intraoperatively.

In the second system, images are acquired using overlapping standard film cassettes [12], which are digitized and then input to a computer that composes them. The setup requires placing a radiolucent plaque with a metal grid pattern parallel, close to the imaged anatomy, and on the X-ray path. The grid placement requirements are necessary to minimize the difference between scale factors of the imaged grid and the patient anatomy. The advantages of this method are that it uses only a few images and that it can image any part of the body. Its disadvantages are that it requires film digitization, which is time-consuming and intraoperatively impractical, and that it yields erroneous measurements when the metal grid is not close to the imaged anatomy.

Siemens AG (Munich, Germany) has developed the only system that produces panoramic images from fluoroscopic X-ray images [1], [2], [3]. It consists of a motorized C-arm which acquires overlapping images by precise simultaneous translation of the X-ray source and image intensifier. The images are aligned by semi-automatically detecting and matching semantically meaningful features on a reconstruction plane. Composed images of objects that are not on that plane will have parallax errors. The advantages of this method are that it requires only a few images, that it can image any part of the body, and that it does not require a phantom. Its disadvantages are that it only works with motorized C-arms, which are expensive and not very common. Also, no direct metric measurements are possible on the resulting panorama since no distance on the reconstruction plane is known a priori.

Panoramic views are widely used in other areas of medicine, such as in dentistry. They are usually tailored to the anatomy and obtained with special purpose cameras, hardware, and films. For example, two recent works describe how to generate digitally reconstructed panoramas of nailfold capillary patterns from video sequences [13] and from ultrasound image sequences [14]. Both methods assume that a sequence of many, largely overlapping images taken in a single plane, is available. These methods are not applicable to the current practice of X-ray fluoroscopy, where the undesirable "continuous mode" is seldom used.

Geometric distortion correction of individual fluoroscopic X-ray images is a necessary first step before composing them. Recent studies indicate distortion of up to $4mm$ on the image edges in older fluoroscopic C-arm units [15]. Fluoroscopic X-ray distortion correction is well understood and has been addressed in previous research, including our own [15], [16], [17]. Image alignment, also called image registration, consists of computing the transformation that aligns paired points in two data sets. Previous work shows how to establish and reduce the dissimilarity between the images [18], which can be based on geometric features or on pixel intensity values. Feature-based alignment assumes feature segmentation but requires less overlap between images [7]. Intensity-based matching does not require segmentation but only works for nearly identical images [19]. Neither feature-based nor intensity-based alignment are directly applicable to our problem. Accurate segmentation of anatomical features in fluoroscopic X-ray images is not, in general, sufficiently reliable [15], [20]. For anatomical structures such as long bones, there are not enough distinct features to perform the match. Intensity-based methods are impractical since several dozen images are required to achieve significant overlap between them.

II. EQUIPMENT AND IMAGE ACQUISITION PROTOCOL

We have developed a new method for creating a single panoramic image of a long bone from individual fluoroscopic X-ray images. In this section, we define an image acquisition protocol and describe customized techniques from image processing and computer vision to match overlapping images, determine their relative position, and compose them into a single panoramic view. The relative simplicity and robustness of our method derives from the combination of the right simplifying assumptions for image acquisition, the use of external markers to establish image correspondence, and the judicious adaptation of image processing algorithms. This protocol is designed to minimize the patient's radiation exposure, eliminates the radiation exposure to the surgeon, and can be performed by any X-ray technician with no additional training. It can be repeated at various stages during the surgery.

We first describe the equipment setup and image capture protocol. The equipment consists of a mobile fluoroscopic C-arm unit commonly used in the operating room, a standard PC with a video card and a monitor, a custom dewarp grid, and an off-the-shelve orthopaedic radiolucent X-ray ruler
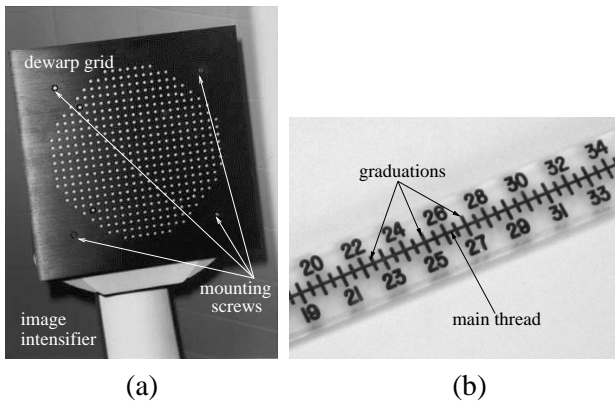
Fig. 2. Equipment and its setup: (a) dewarp grid mounted on the C-arm image intensifier; (b) orthopaedic X-ray ruler.
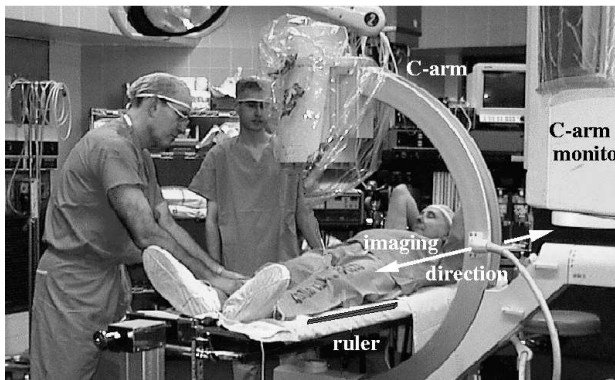

Fig. 3. Operating room setup: the ruler is placed next to the patient, parallel to the long bone axis. To acquire the overlapping images for the panorama, the fluoroscopic C-arm is moved parallel to the long bone axis.
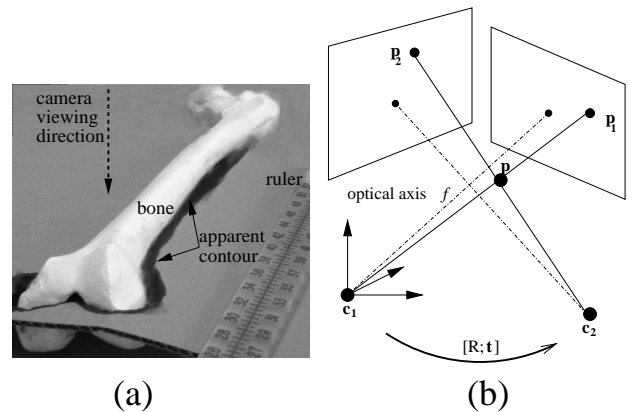

Fig. 4. (a) Fronto-parallel setup: the C-arm camera viewing direction is perpendicular to the ruler plane and the bone apparent contour plane, and; (b) Camera model: p is a point in space, p_1 and p_2 are its projections from camera poses c_1 and c_2. The coordinates origin coincides with c_1. The camera location c_2 is rotated by R and translated by t.

with graduations. Images are directly downloaded from the fluoroscopic C-arm unit to the computer via a digital port, or are captured from the video output port with an analog to digital frame grabber. The images are stored and processed in the computer and the resulting panoramic view is displayed on the fluoroscopic unit screen or on the computer screen.

The custom dewarp grid is used to correct the images for geometric distortion (Fig. 2a). It is a $7mm$ thick coated aluminum alloy plate with 405, $4mm$ diameter holes uniformly distributed at $10mm$ intervals machined to $\pm 0.02mm$ tolerance. It attaches to the C-arm image intensifier on existing screw holes. This grid is simpler and cheaper to make than the commonly used steel balls or cross-hairs mounted on a radiolucent plate, and yields excellent results. The radiolucent X-ray ruler (Fig. 2b) is 1 meter long, has graduations at $5mm$ intervals, and is sterilizable.

The image acquisition protocol is as follows. Shortly before surgery, the C-arm is oriented at the pose which will be used in acquiring the panoramic images (usually anterior-posterior or lateral). The dewarp grid is attached to the image intensifier and an image of it is acquired and transferred into the computer, which computes the distortion correction map. The dewarp grid is then detached from the C-arm and the surgery proceeds as usual. When the panoramic view is required during the surgery, the ruler is placed next to the

patient, roughly parallel to the imaged long bone (Fig. 3).

The camera, ruler, and anatomy of interest are placed such that they form a fronto-parallel setup, in which the C-arm viewing direction is perpendicular to the ruler plane and the bone apparent contour plane (Fig. 4a). A sequence of overlapping images is then acquired by translating the C-arm parallel to the bone axis. Subsequent images should overlap in 20–60% of their area. The images are automatically downloaded to the computer as they are acquired, and the resulting undistorted panoramic image is displayed on the computer and on the C-arm screens.

The method just described relies on four assumptions: 1) images are acquired in a fronto-parallel setup, 2) the C-arm orientation does not change during the calibration and the image acquisition, 3) there is sufficient overlap between the individual fluoroscopic X-ray images, and 4) the user selects the reconstruction plane. The assumptions are both practical and realistic.

First, the fronto-parallel C-arm setup is commonly used in long bone surgery since it corresponds to the anterior-posterior and the lateral viewpoints which are very familiar to surgeons and X-ray technicians. It is the only viewing setup which enables metric measurements on the resulting panorama, since it induces a transformation that preserves angles and distance ratios. Other viewing setups require a priori knowledge of angular and distance relationships between viewed objects that are seldom available.

Second, the C-arm orientation remains the same during image acquisition, but may differ from the one at the time it was calibrated. The C-arm can be re-oriented to the calibrated orientation using the gantry graduations to an accuracy of five degrees or less. Larger angular variations will introduce more error but will not cause the method to fail. Our previous study shows that the distortion correction discrepancies for these small angle variations are sub-millimetric [15].

Third, sufficient ($\geq 20\%$) overlap between consecutive images can be easily achieved by comparing the ruler markings that appear in the acquired images. The X-ray technician can determine if the overlap is sufficient, and if not, discard the

acquired image, adjust the C-arm position, and acquire another image. Also, the entire image set can be discarded if, upon visual inspection, the result is incorrect. More overlap than 60% is neither necessary nor desirable, since the patient's radiation exposure should be minimized. Typically, 4–10 images are sufficient for long bone panoramic images.

Fourth, the user interaction to select the desired reconstruction plane is essential, since there is one such plane for each anatomical structure and implant. For example, in the intramedullary tibial fracture reduction surgery (described in detail Section IV.C), there are two planes: the tibial plane and the nail plane. One of them should be chosen, depending on where the visualization and the measurements will be done. To define the plane, the system prompts the user to mark the contour of interest. Inputing the contour requires minimal user effort and can be completed in several seconds of user interaction.

III. PANORAMA CONSTRUCTION ALGORITHM

The algorithm creates a single rectangular panoramic view from the individual images in three steps: (1) distortion correction, (2) image alignment, and (3) image composition (Table I). For distortion correction, we compute the distortion map from the previously acquired grid image and the geometric model of the grid using local bi-linear interpolation [15] and apply it to each image. For image alignment, we use the ruler's graduations to compute the planar rigid transformation between successive pairs of images by finding correspondences in their overlapping regions. To extract the graduations, we first identify the ruler by extracting its main thread with a modified Hough transform. This defines a region of interest where the graduations can be isolated by performing edge detection. We then compute the planar transformation that maximizes the Normalized Cross Correlation similarity measure between the original images. To adjust the transformation to the bone's apparent contour, the user specifies a contour of interest (there might be more than one) and the images are realigned according to it. For image composition, we select the first image as the reference and align all other images to it. We compute the panoramic image size and apply the computed transformations to the undistorted individual images. The resulting panoramic image has overlapping pixels coming from two or more images. Their values are computed by taking either the average, median, maximum, or minimum of the individual pixel values. From visual inspection, we concluded that the maximum yields the most uniform image. All the panoramas in this paper were computed using this rule. Fig. 5 illustrates the steps of the algorithm on a dry femur. Note that the method handles partial ruler occlusions.

Next, we describe in detail the camera projection model and the image alignment method, which is the most technically challenging step.

A. Camera and projection model

We model the fluoroscopic camera as a pin-hole camera, as this has been shown to be an appropriate approximation of the X-ray imaging process [15]. We assume, based on our

---

Input: dewarp grid image, a sequence of partially overlapping images
1. *Distortion correction:* compute an image dewarp map
2. *Image alignment:* align subsequent pairs of images
    1. *Ruler region identification:* find the ruler in each image by locating its horizontal thread with a modified line Hough transform.
    2. *Graduations segmentation:* segment the ruler graduations with a one-dimensional edge detection method
    3. *Transformation computation:* compute the transformations that aligns on the ruler's plane the ruler graduations in successive images using the Normalized Cross Correlation similarity measure.
    4. *Image realignment:* compute the transformations that align the images on the bone apparent contour plane.
3. *Image composition:* compose the overlapping pixels of the aligned images.

Output: panoramic X-ray image

TABLE I

PANORAMA CONSTRUCTION ALGORITHM.

previous studies and that of others, that the camera internal parameters do not change when the camera poses are only translations on a plane [16], [17], [15].

The intrinsic camera parameters are:

f – camera focal length.
u_0, v_0 – image origin coordinates at the intersection of the optical axis and the image plane.
k_u, k_v – horizontal and vertical pixel scale factors in the image plane.
θ – angle between image plane axis.

Following [21], the camera imaging parameters are modelled by a 3×3 camera projection matrix K defined by

$$K = \begin{bmatrix} -fk_u & \frac{fk_u}{\tan\theta} & u_0 \\ 0 & -\frac{fk_v}{\sin\theta} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (1)$$

Without loss of generality, we align the world coordinate system with the coordinate system of the first image camera pose. The second image is acquired after the camera has undergone a rotation, represented by a 3×3 rotation matrix R and a translation, represented by a vector $\mathbf{t}$. Let $\mathbf{p}$ be a point in space, and let $\mathbf{p_1}$ and $\mathbf{p_2}$ be its two projections in homogenous coordinates from two different camera viewpoints (Fig. 4(b)).

We seek the projective mapping, homography matrix H_π, which maps $\mathbf{p_1}$ onto $\mathbf{p_2}$:

$$\mathbf{p_2} = H_\pi \mathbf{p_1}$$

The relations between the point $\mathbf{p}$ and its projections $\mathbf{p_1}$ and $\mathbf{p_2}$ are:

$$\mathbf{p_1} = K\mathbf{p} \quad \text{and} \quad \mathbf{p_2} = K[R\mathbf{p} + \mathbf{t}]$$

In the first camera coordinate system we have:

$$\mathbf{p} = zK^{-1}\mathbf{p_1} \quad (2)$$

where z is the third coordinate of point $\mathbf{p} = [x, y, z]$.

Similarly, in the second camera coordinate system we have:

$$\mathbf{p}' = z'K^{-1}\mathbf{p_2} = R\mathbf{p} + \mathbf{t} \quad (3)$$

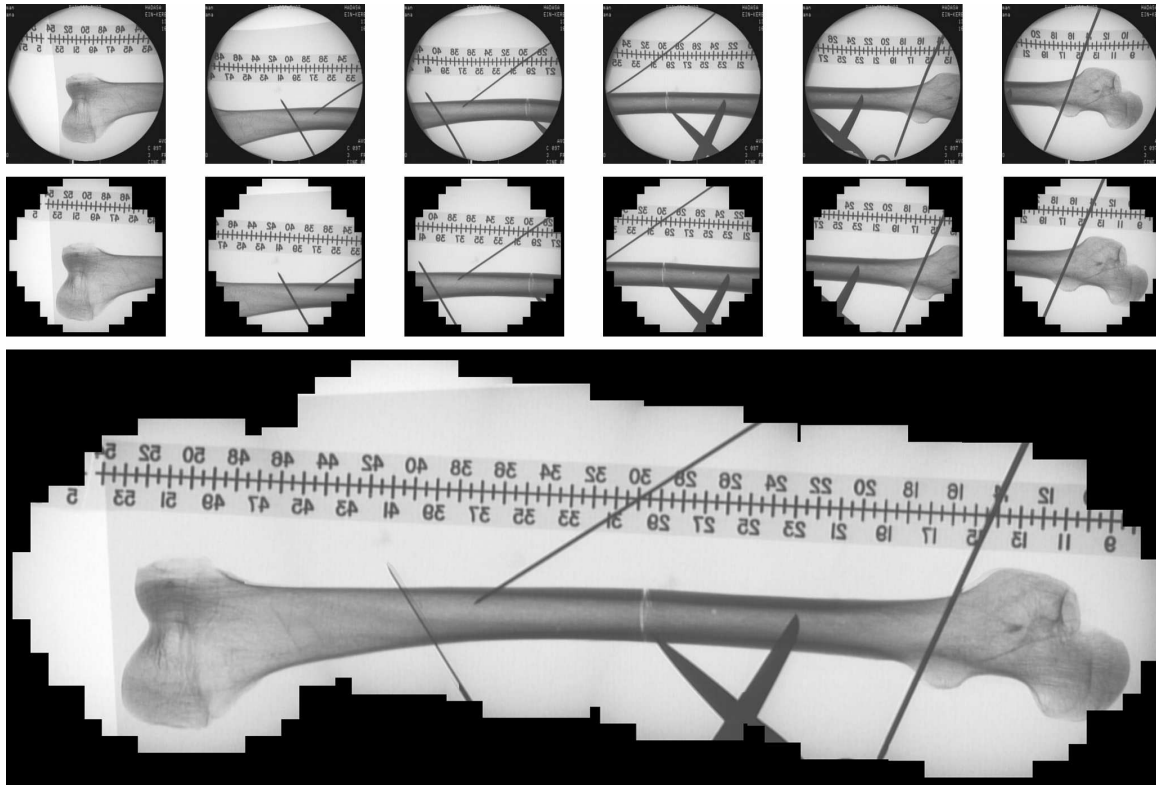where z' is the third coordinate of point $\mathbf{p}' = [x', y', z']$.

Fig. 5. Panorama of a dry femur. The top row shows the original images, the middle row shows the images after distortion correction, and bottom row shows the resulting panorama. Scissors and k-wires were placed below and above the femur.

Substituting Eq. 2 into Eq. 3 we obtain:

$$z' K^{-1} \mathbf{p_2} = z R K^{-1} \mathbf{p_1} + \mathbf{t}$$
$$z' \mathbf{p_2} = z K R K^{-1} \mathbf{p_1} + K \mathbf{t} \qquad (4)$$
$$\mathbf{p_2} = \frac{z}{z'} K R K^{-1} \mathbf{p_1} + \frac{1}{z'} K \mathbf{t}$$

The homography H_π relates the two projections $\mathbf{p_1}$ and $\mathbf{p_2}$ and is computed without actually knowing the spatial coordinates of $\mathbf{p}$. Let $\Pi_{\mathbf{n},d}$ be a plane containing point $\mathbf{p}$ defined by normal vector $\mathbf{n}$ and distance offset d in world coordinates. Then:

$$\mathbf{n}^T \mathbf{p} = d$$
$$\frac{1}{d} \mathbf{n}^T \mathbf{p} = 1 \qquad (5)$$

Substituting Eq. 2 into Eq. 5 we obtain:

$$\frac{1}{d} \mathbf{n}^T (z K^{-1} \mathbf{p_1}) = 1 \qquad (6)$$

Substituting Eq. 6 into Eq. 4 we obtain:

$$\mathbf{p_2} = (\frac{z}{z'} K R K^{-1}) \mathbf{p_1} + (\frac{z}{z'} \frac{1}{d} K \mathbf{t} \mathbf{n}^T K^{-1}) \mathbf{p_1}$$

as $\mathbf{p_1}$ is in homogenous coordinates we can drop the scale factor $\frac{z}{z'}$:

$$\mathbf{p_2} = (K R K^{-1}) \mathbf{p_1} + (\frac{1}{d} K \mathbf{t} \mathbf{n}^T K^{-1}) \mathbf{p_1}$$
$$\mathbf{p_2} = (K R K^{-1} + \frac{1}{d} K \mathbf{t} \mathbf{n}^T K^{-1}) \mathbf{p_1} \qquad (7)$$
$$\mathbf{p_2} = H_\pi \mathbf{p_1}$$

This yields the desired projective mapping H_π:

$$H_\pi = K R K^{-1} + \frac{1}{d} K \mathbf{t} \mathbf{n}^T K^{-1} \qquad (8)$$

Note that the first term of this expression, $K R K^{-1}$, is the same for all points and is independent of the plane parameters. The second term, $\frac{1}{d} K \mathbf{t} \mathbf{n}^T K^{-1}$, establishes the dependency with the plane parameters $\mathbf{n}$ and d. The homography holds for all imaged points when the camera transformation is purely rotational (the second term equals zero) or when all imaged points are coplanar (both terms are identical for all points). We use this equation later on to estimate the parallax error and compute a scale factor for the translation between pixels and millimeters.

Based on our image acquisition protocol, we can make the following simplifications:

- Since the individual images have been corrected for geometric distortion, the angle between image plane axis, θ, is $\frac{\pi}{2}$ and the horizontal and vertical pixel scale factors k_u, k_v, are both 1. The camera projection matrix then becomes:

$$K = \begin{bmatrix} -f & 0 & u_0 \\ 0 & -f & v_0 \\ 0 & 0 & 1 \end{bmatrix} \qquad (9)$$

- Since the images are acquired by translating the C-arm, camera poses lie on a plane with a rotational component between images only around the optical axis. The apparent contour plane of the bone is nearly planar (the shape of long bones is nearly cylindrical) and roughly parallel to the image plane, the images are in fronto-parallel position (Fig. 4a). Thus, the viewed plane is parallel to the image plane and the viewed points $\mathbf{p}$ are on a plane $\Pi_{\mathbf{n},d}$ defined by $\mathbf{n}^T = [0, 0, -1]$ and $d = z$. This setup restricts the

relationships between ruler, anatomy, and camera and is the only one that preserves angles and scales distances.

- Since the images are acquired by translating the C-arm parallel to the viewed plane, the translations between images is $\mathbf{t} = [x, y, 0]$.

These conditions yield the well known expression for the rigid planar mapping:

$$H_\pi = \begin{bmatrix} \cos\theta & -\sin\theta & x \\ \sin\theta & \cos\theta & y \\ 0 & 0 & 1 \end{bmatrix}$$

The mapping in a fronto-parallel acquisition is a special case of the general perspective projection in which the distance ratios between points and the angles between lines are preserved. Once a known distance on the image is identified (e.g., the distance between ruler graduations), metric measurements can be made directly on the image.

B. Image alignment

Pairs of consecutive images are aligned using a feature-based alignment method that uses the graduations of the ruler to compute the planar mapping transformation. Since the images are in fronto-parallel position, the transformation relating the images is planar, and thus the problem reduces to estimating three parameters x, y, θ between subsequent images. Image alignment consists of four steps (Table I): (1) find the ruler region, (2) segment the graduations, (3) compute the transformation, and (4) realign the images to compensate for parallax. We describe each step in detail next.

1. Ruler region identification

The main ruler thread is located using a modified two dimensional line Hough transform [22]. The transform computes the ruler's angle and distance from the origin. Since we are looking for a pair of parallel lines (the boundaries of the main thread), we add a constraint to the line Hough transform voting scheme. This constraint specifies that pixels vote for a certain line only if there is another pixel in the image which is on a line parallel to the current one. The difference in angle and translation between pairs of images are the first two alignment parameters, y and θ. This method is robust to partial occlusions of the ruler (Fig. 5).

2. Graduations segmentation

Next, we examine the region near the ruler's main thread (Fig. 6a). We sum the columns in the region of the image and perform a one dimensional edge detection on the resulting signal (Fig. 6b). This one dimensional signal has pronounced local minima at the graduation edges. Performing edge detection on this signal yields the right and left edges of the ruler graduations. Ruler graduations are now identified as pairs of these right and left edges (Fig. 6c-e).

3. Transformation computation

Next, we compute the missing translation parameter x, which corresponds to translation along the ruler's main thread. The visible graduations define the possible relative image translations along the axis parallel to the ruler. Valid translations are only those that align the graduations between images. Since there are at most a few dozen such graduations, the
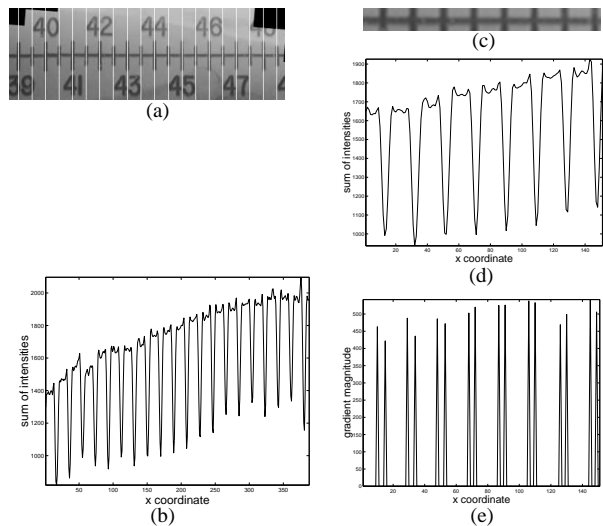

Fig. 6. Ruler's graduations segmentation: (a) ruler image after graduations segmentation; (b) graph of the sum of gray level pixel values (vertical axis) as a function of its horizontal position; (c) detail of the ruler graduations; (d) its corresponding sum of intensities graph, and; (e) corresponding gradient magnitude graph.

search space for the translation is both discrete and small. We use Normalized Cross Correlation (NCC) [23], also known as the Pearson correlation coefficient, as our similarity measure:

$$NCC = \frac{\sum_{i,j}(I_1 - \overline{I_1})(I_2 - \overline{I_2})}{\sqrt{\sum_{i,j}(I_1 - \overline{I_1})^2 \sum_{i,j}(I_2 - \overline{I_2})^2}}$$

where I_1 and I_2 are the image pixel intensities at location i, j, and $\overline{I_1}$ and $\overline{I_2}$ are the average image pixel intensities of the overlapping areas of the first and second images, respectively. The NCC measure is invariant to the overlap between images and can be computed efficiently. Because the registration is unimodal, the linear dependence measured by NCC is appropriate.

4. Image realignment

The image alignment described above aligns the images on the ruler's plane. However, the correct alignment should be on the plane defined by the bone apparent contour. The conversion scale factor from pixels to millimeters should also be corrected to the bone apparent contour plane.

We compute a translational correction for the pairwise transformations as follows. Let d be the ruler plane distance and $d + \delta d$ the apparent contour plane distance. From Eq. 8, the homography defined by the ruler plane is:

$$H_{ruler} = KRK^{-1} + \frac{1}{d}K\mathbf{t}\mathbf{n}^T K^{-1}$$

The homography defined by the apparent contour plane is:

$$\begin{aligned} H_{contour} &= KRK^{-1} + \frac{1}{d + \delta d}K\mathbf{t}\mathbf{n}^T K^{-1} \\ &= KRK^{-1} + \frac{d}{d + \delta d}\left(\frac{1}{d}K\mathbf{t}\mathbf{n}^T K^{-1}\right) \end{aligned}$$

The homographies H_{ruler} and $H_{contour}$ differ only by a scaling of the translational component. To find the translational correction ($\frac{d}{d + \delta d}$) the user marks a contour of interest on two

consecutive images using the Livewire [24] segmentation tool. It allows the user, with a few coarsely-placed input points, to accurately segment the contour edges and thus define the reconstruction plane. For each contour point on the first image, we rotate it by the rotational component of H_{ruler} and then search for the closest contour point on the second image in the direction of the translational component of H_{ruler}. The scale factor is the mean scale factor computed for all matched contour points.

The scaling of the conversion from pixels to millimeters is computed as follows. Let f be the camera focal length, δ_1 and δ_2 be the image distance of an object in the ruler and apparent contour plane, Δ_1 a known distance in the ruler plane, and Δ_2 the corresponding unknown distance in apparent contour plane. Then, we have:

$$\delta_1 = \frac{f\Delta_1}{d} \qquad \text{and} \qquad \delta_2 = \frac{f\Delta_2}{d+\delta d}$$

Thus,

$$\Delta_2 = \delta_2 \left(\frac{d+\delta d}{d}\right)\left(\frac{\Delta_1}{\delta_1}\right)$$

To convert from pixels to mm we scale the known translation factor $\frac{\Delta_1}{\delta_1}$ by $\frac{d+\delta d}{d}$, which is the inverse of the translational scale factor.

IV. Experimental results

We conducted experiments to validate our method, quantify its accuracy, and determine its clinical value. Specifically, we performed a geometric error analysis study to determine the accuracy of the distortion correction and alignment steps in the panorama construction algorithm. We created panoramic images of several in-vitro and in-vivo long bones and showed them to doctors for evaluation. Finally, we created panoramas during a tibial intramedullary fracture reduction surgery which were used by the surgeon to evaluate the mechanical axis, choose the nail length, and document the surgery outcome. We describe each set of experiments in detail next.

All images were acquired with a 9" BV29 C-arm (Phillips, The Netherlands). They were 800x600 pixels, 8-bit gray-scale with pixel size of about $0.5mm \times 0.5mm$ in the imaging setup described above. The distortion correction and panorama construction algorithms were implemented in C++ and ran on a 600 MHz, 256MB RAM PC running Windows NT. Processing times were 5–10 seconds per data set consisting of 4–10 individual images.

A. Error quantification

Geometric errors in the panorama arise from three sources: image distortion correction, Hough transform discretization, and distance difference between the ruler and the apparent contour plane. Our experiments show that the geometric distortion can be corrected across the entire field of view to an average accuracy of $0.25mm$ ($0.4mm$ in the worst case) [15]. The standard Hough transform discretizes the parameter space, both the angle and translation. To minimize the errors introduced by the discretization, we use a parameter space
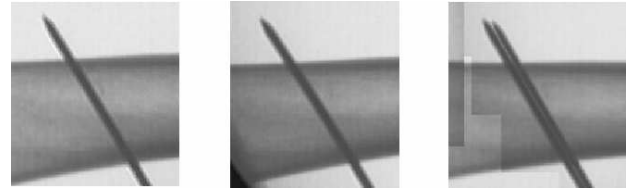

Fig. 7. Illustration of the parallax effect on the leftmost K-wire in Fig. 5. The right image is the result of composing the two left images using the transformation computed for the bone reconstruction plane. The K-wire tip is about 50mm above the ruler.
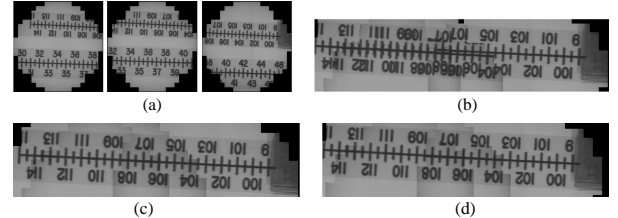

Fig. 8. Rescaling accuracy experiment:(a) original dewarped images (b) without correction (c) own plane (d) with correction.

with increments of 0.5 pixels for the translation and 0.5^o for the rotation.

The main cause of errors is the distance difference between the radiation source and the planes defined by the apparent contour and the ruler. When this distance difference is not taken into account, the resulting panorama has undesired scaling errors and parallax effects (Fig. 7).

To evaluate the accuracy of the rescaling, we imaged two rulers (top and bottom) side by side at different heights (Fig. 8a). We split each image into two, so that only one ruler appears in each, computed the transformations, and created panoramic images on the ruler's own plane (Fig. 8c). We then applied the transformation of the top ruler panorama to the bottom ruler panorama, and created two new panoramas, without and with realignment correction (Fig. 8b and d, respectively). The ruler's own plane and realignment corrected panoramas are nearly identical (Fig. 8c and d), while the non-realigned panorama (Fig. 8b) shows parallax. Table II quantifies these discrepancies as a function of the height difference. The average image measurement error is estimated at ± 1 pixel. The results show that realignment is necessary, and that the realignment error is at most $1.1mm$, which is clinically acceptable.

B. Preoperative experiments

We acquired six sets of fluoroscopic X-ray images of dry long bones (humerus and femur) and two sets of in-vivo long bones (humerus and tibia). We acquired 4–8 images for each data set, placing a ruler with $5mm$ graduations next to the anatomy following the protocol described above. The overlap between consecutive images was about 50%, which was visually verified with the aid of the ruler's graduations.

For qualitative evaluation, we showed the panoramas to an orthopedic surgeon and got very satisfactory results: the bone structures and the ruler showed continuous boundaries, with very small "jumps" (one or two pixels) at locations where

Height	Ruler	Actual size	Without correction	Own plane	With correction
10	Top	150	148.9 (1.0)	150.2 (0.2)	150.6 (0.6)
	Bottom	150	151.1 (1.1)	150.0 (0.0)	150.1 (0.1)
50	Top	120	111.5 (8.4)	119.7 (0.2)	119.4 (0.5)
	Bottom	130	136.2 (6.2)	129.6 (0.3)	129.0 (0.9)
90	Top	150	140.2 (9.7)	150.0 (0.0)	151.0 (1.0)
	Bottom	160	173.0 (13.0)	160.8 (0.8)	158.9 (1.1)

TABLE II

REALIGNMENT MEASUREMENTS IN MILLIMETERS (ERROR IN PARENTHESIS). THE FIRST COLUMN SHOWS THE HEIGHT DIFFERENCE BETWEEN THE TWO RULERS. THE SECOND AND THIRD COLUMN SHOW WHICH RULER WAS USED AND ITS LENGTH. THE FOURTH, FIFTH, AND SIXTH COLUMN SHOW THE RULER LENGTH MEASURED IN THE PANORAMA WITHOUT CORRECTION, IN THE RULER'S OWN PLANE, AND WITH RESCALING CORRECTION IN THE OTHER RULER'S PLANE, RESPECTIVELY.

Data Set	Number of Images	Actual	Measured	Error
1. humerus head	1	42	42.5	0.5
2. humerus length	4	287	285.9	1.1
3. femur head	1	47	49.8	2.8
4. femur proximal	3	195	193.5	1.5
5. femur distal	3	223	220.8	2.2

TABLE III

ACTUAL AND MEASURED DISTANCES ON DRY LONG BONES, IN MILLIMETERS. THE FIRST COLUMN INDICATES THE BONE REGION IN WHICH THE MEASUREMENTS WERE PERFORMED. THE SECOND COLUMN INDICATES THE NUMBER OF IMAGES IN WHICH THE ANATOMY APPEARS. THE THIRD AND FOURTH COLUMNS SHOW THE ACTUAL AND MEASURED DISTANCE VALUES. THE FIFTH COLUMN IS THE IMAGE MEASUREMENT ERROR.

images were composed. Figs. 1 and 5 show in-vivo and in-vitro panoramas. The images were deemed accurate and of clinical value.

For quantitative evaluation, we performed physical distance measurements on the dry humerus and femur and compared them with distance measurements on the images. The physical distance measurements were performed with a desk ruler with $1mm$ graduations, so the average measurement error is estimated at $\pm 1mm$. Table III summarizes these results. We observe that the error is below $2.8mm$ (6%) in all cases. The relatively large error in case 3 is due to the fact that the measurement on the femoral head was relatively far from the apparent contour plane.

C. Intraoperative experiment

To test the method and protocol in an intraoperative setting, we participated in a tibial intramedullary fracture reduction surgery. The surgery restores the integrity of the fractured bone by means of a nail inserted in the medullary canal. The nail is placed without surgically exposing the fracture through an opening on the tibial fossa, right below the knee joint. The surgeon manually aligns the bone fragments by manipulating them through the skin, drives the nail in, and inserts lateral proximal and distal interlocking screws to prevent fragment rotation and bone shortening. The procedure is performed under X-ray fluoroscopy, which is used to view the position of bone fragments, surgical tools, and implants. The mechanical axis position and nail length are determined empirically, and sometimes require several trials.

Right before the surgery, the C-arm was fitted with the de-warp plate and an anterior-posterior image was acquired. The dewarp plate was then removed and the image intensifier was draped. This step required four minutes and did not interfere with the surgery preparations. Before the fracture reduction, the surgeon placed the sterilized ruler on the operating table, right next to the patient tibia. We acquired nine fluoroscopic X-ray images and created a panorama (Fig. 9a). The overlap was visually verified using the ruler's graduations. The surgeon used the resulting panorama to determine the extent of the fracture compression, to determine the nail length, and to assess the mechanical axis. The surgery proceeded as usual. Once the reduction and nail locking were completed, we acquired an additional set of six fluoroscopic X-ray images and created a panorama (Fig. 9b). The surgeon used it to verify that the reduction was properly performed, that the leg length was as desired, and that the nail axis was properly aligned with the bone mechanical axis. The image acquisition for the panorama required only a few minutes in each case. The surgeon concluded that the panoramas helped him make the right nail length decision, and reduced the need for additional validation X-ray images. It also eliminated the need for a postoperative film X-ray.

V. CONCLUSION

We have presented a simple and robust method for creating a single panoramic image of a long bone from individual fluoroscopic X-ray images. Panoramic images are useful pre-operatively for diagnosis, and intraoperatively for tasks such as long bone fragment alignment, for making anatomical measurements, and for documenting surgical outcomes. Unlike existing methods, ours uses readily available hardware, requires a simple image acquisition protocol with minimal user input, and works with existing fluoroscopic C-arm units without modifications. The method has been successfully tested on in-vitro and clinical cases. Our experiments indicate that the method is practical and produces images and measurements which are clinically acceptable.

We are planning to conduct a quantitative study to compare the accuracy of measurements with the panoramas to those obtained from CT models, which constitute the golden standard. We are also planning to further evaluate the clinical value of the panoramic images on actual patient cases in a variety of routine preoperative and intraoperative procedures.

As real-time tracking systems become ubiquitous in the operating room, we can envision extending our method to take advantage of the data they provide (however, we do not believe that the use of a tracker for the purpose of constructing panoramas alone is justified). By fitting a tracking device onto the C-arm and calibrating it, we can obtain accurately in real time the position and orientation of its imaging plane. Because the relative positions of the individual C-arm images are known, we no longer need to restrict the viewing to a fronto parallel setup or require a planar apparent contour.
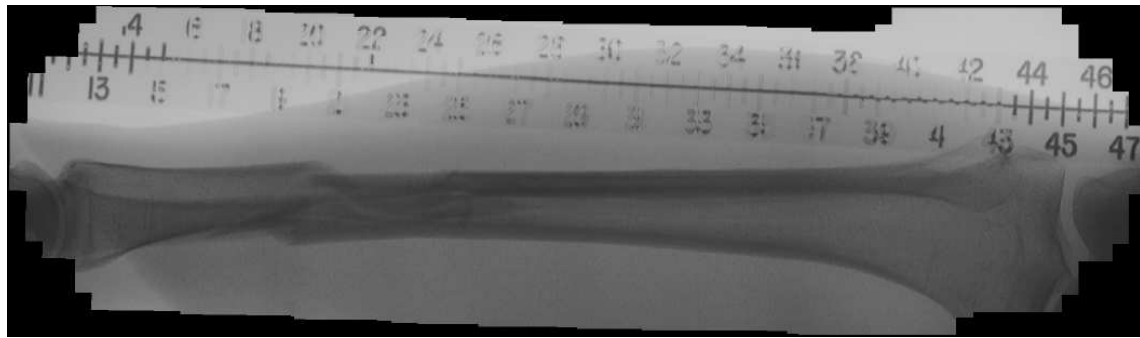
(a)
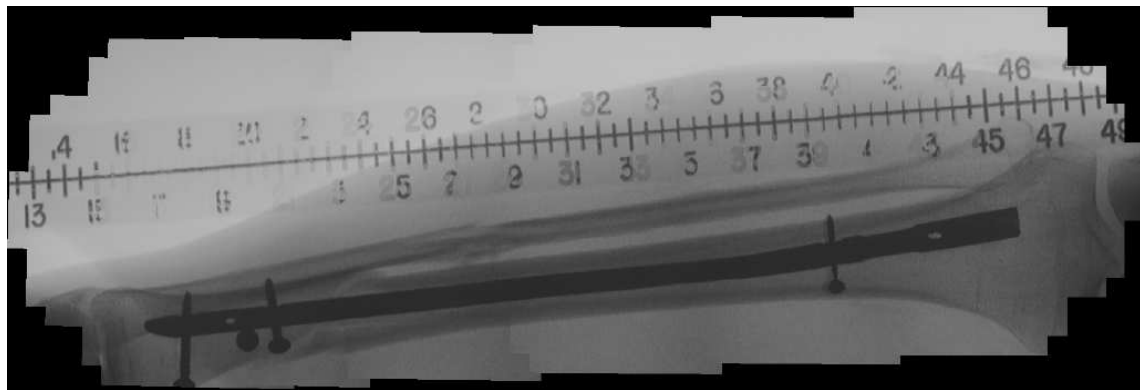
(b)

Fig. 9. Intraoperative panoramas of a fractured tibia (a) before fracture reduction (b) after intramedulary nailing and locking.

REFERENCES

[1] J. S. Chou and J. Qian, "Automatic full-leg mosaic and display for peripheral angiography," U.S. Patent 5 833 607, Nov. 10, 1998.

[2] S. K. Murthy, C. L. Novak, J. Qian, and Z. Wu, "System for generating a compound x-ray image for diagnosis," U.S. Patent 6 101 238, Aug. 8, 2000.

[3] D. L. Wilson, "Whole-leg x-ray image processing and display techniques," U.S. Patent 5 123 056, June 16, 1992.

[4] D. Capel and A. Zisserman, "Automated mosaicing with super-resolution zoom," in *Proc. Computer Vision and Pattern Recognition*, 1998.

[5] M. Irani and P. Anandan, "Video indexing based on mosaic representations," *Proc. IEEE*, vol. 86, no. 5, 1998.

[6] B. Rousso, S. Peleg, and I. Finci, "Mosaicing with generalized strips," DARPA Image Understanding Workshop, 1997.

[7] I. Zoghlami, O. Faugeras, and R. Deriche, "Using geometric corners to build a 2d mosaic from a set of images," in *Proc. Computer Vision and Pattern Recognition*, 1997.

[8] A. Zomet and S. Peleg, "Efficient super-resolution and applications to mosaics," in *Int. Conf. on Pattern Recognition*, 2000.

[9] H. W. van Eeuwijk, S. Lobregt, and F. A. Gerristen, "A novel method for digital x-ray imaging of the complete spine," in *Proc. Computer Vision, Virtual Reality and Robotics in Medicine*, 1997.

[10] B. Verdonck, R. Nijlunsing, N. Melman, and H. Geiger, "Image quality and x-ray dose for translation reconstruction overview imaging of the spine, colon, and legs," in *Proc. Computer Assisted Radiology and Surgery*, 2001.

[11] H. Geijer, K. Beckman, B. Jonsson, *et al.*, "Digital radiography of scoliosis with a scanning method: initial evaluation," *Radiology*, 2001.

[12] P. Dewaele, P. Vuylsteke, S. V. de Velde, and E. Schoeters, "Full-leg/full-spine image stitching, a new and accurate cr-based imaging technique," in *Proc. SPIE Medical Imaging: Image Processing*, 1999.

[13] P. D. Allen, C. J. Taylor, A. L. Herrick, and T. Moore, "Image analysis of nailfold capillary patterns from video sequences," in *Proc. Medical Computing and Computer-Assisted Intervention*, 1999.

[14] L. Weng and A. P. Tirumali, "Method and apparatus for generating large compound ultrasound image," U.S. Patent 5 575 286, Nov. 19, 1996.

[15] Z. Yaniv, L. Joskowicz, A. Simkin, *et al.*, "Fluoroscopic image processing for computer-aided orthopedic surgery," in *Proc. Medical Computing and Computer-Assisted Intervention*, 1998.

[16] C. Brack, R. Burghart, A. Czpof, *et al.*, "Accurate x-ray navigation in computer-assisted surgery," in *Proc. Computer Assisted Radiology and Surgery*, 1998.

[17] R. Hofstetter, M. Slomczykowski, M. Sati, and L. P. Nolte, "Fluoroscopy as an imaging means for computer-assisted surgical navigation," *Computer Aided Surgery*, vol. 4, no. 2, 1999.

[18] S. Lavallée, *Computer-integrated surgery, Technology and clinical applications*. Cambridge Ma.: MIT Press, 1995, ch. 5.

[19] M. J. Murphy, "An automatic six-degree of freedom image registration algorithm for image-guided frameless stereotactic radiosurgery," *Medical Physics*, vol. 24, no. 6, 1997.

[20] T. O. Ozanian and R. Phillips, "Image analysis for computer-assisted internal fixation of hip fractures," *Medical Image Analysis*, vol. 4, no. 2, 2000.

[21] O. Faugeras, *Three-Dimensional Computer Vision : A Geometric Viewpoint*. MIT Press, 1993.

[22] R. Jain, R. Kasturi, and B. G. Schunk, *Machine Vision*. McGraw-Hill, 1995.

[23] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 2nd ed. Prentice Hall, 2002, ch. 12.

[24] E. N. Mortensen and W. A. Barrett, "Interactive segmentation with intelligent scissors," *Graphical Models and Image Processing*, vol. 60, no. 5, 1998.