Fluoroscopic Image Processing for Computer-Aided Orthopaedic Surgery

Z. Yaniv[1], L. Joskowicz[1], A. Simkin[2,3], M. Garza-Jinich[4], and C. Milgrom MD[3]

[1] Institute of Computer Science, The Hebrew Univ., Jerusalem 91904 Israel.
[2] Lab. of Experimental Surgery, Hadassah Univ. Hospital, Jerusalem 91120 Israel.
[3] Dept. of Orthopaedic Surgery, Hadassah Univ. Hospital, Jerusalem 91120 Israel.
[4] IIMAS – Univ. Nacional Autonoma de Mexico, AP 20-726, Mexico 01000 DF.
`{zivy,josko}@cs.huji.ac.il, ruskin@vms.huji.ac.il, milgrom@md2.huji.ac.il`

Abstract. This paper describes the fluoroscopic X-ray image processing techniques of FRACAS, a computer-integrated orthopaedic system for bone fracture reduction. Fluoroscopic image processing consists of image dewarping, camera calibration, and bone contour extraction. Our approach focuses on bone imaging and emphasizes integration, full automation, simplicity, robustness, and practicality. We describe the experimental setup and report results quantifying the accuracy of our methods. We show that after dewarping and calibration, submillimetric spatial positioning accuracy is achievable with standard equipment. We present a new bone contour segmentation algorithm based on robust image region statistics computation which yields good results on clinical images.

1 Introduction

Current orthopaedic practice heavily relies on fluoroscopic images to perform surgical procedures. Fluoroscopic X-ray images are captured by an image intensifier mounted on a C-arm and viewed on a monitor (Fig. 1). Surgeons rely on the images to determine the relative position and orientation of bones, implants, and surgical instruments. While inexpensive and readily available, fluoroscopy has limitations. The images have a narrow field of view, have poor resolution and contrast, and show significant geometric distortion. Because they are uncorrelated, two-dimensional static views, the surgeon must mentally recreate the spatio-temporal intraoperative situation. Significant skill, time, and frequent use of the fluoroscope are required, leading to positioning errors and complications and to significant cumulative radiation exposure to the surgeon [12].

Recent research shows that computer-aided systems can significantly improve the accuracy of orthopaedic procedures by replacing fluoroscopic guidance with interactive display of 3D bone models created from preoperative CT studies and tracked in real time. Examples include systems for acetabular cup placement [13], for total knee replacement, and for pedicle screw insertion [9, 10].

Flurosocopy can still play an important role in computer-aided surgery systems. By correcting, calibrating, and correlating them, a limited number of enhanced fluoroscopic images can be used for accurate navigation. For example,

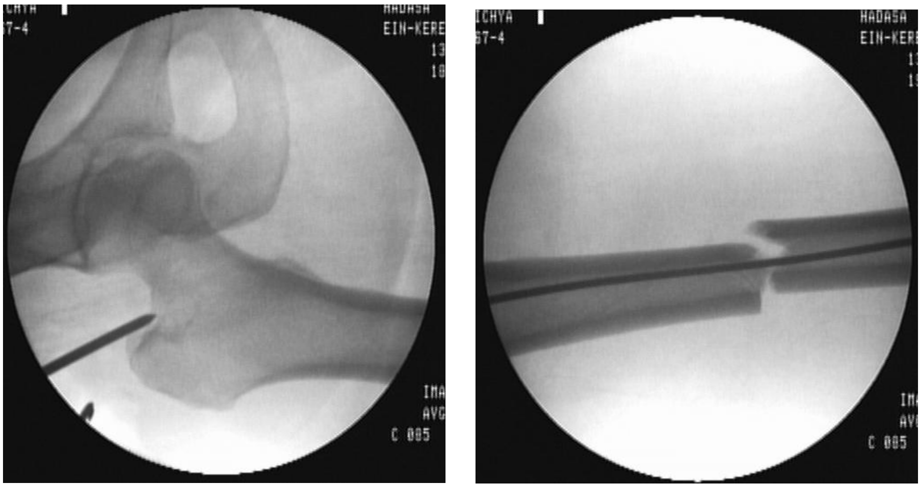

Fig. 1. Fluoroscopic images of a femur fracture with drill bit on the proximal femoral tronchanteric fossa (left), and medial linear fracture with guide wire (right).

[1], [6] and [11, 16] describe systems that use enhanced fluoroscopy and real-time tracking to assist surgeons in distal intramedulary nail locking. They can also be used for registration – establishing a common reference frame between the preoperative CT study and the intraoperative situation.

We are currently developing a computer-integrated orthopaedic system, called FRACAS [7], for closed medullary nailing for bone fracture reduction [2]. FRACAS' goals are to reduce the surgeon's cumulative exposure to radiation and improve the positioning accuracy by replacing uncorrelated static fluoroscopic images with a virtual reality display of spatial bone fragment models created from preoperative CT and tracked intraoperatively in real time. Fluoroscopic images are used to register the bone fragment models to the intraoperative situation and to verify that the registration is maintained.

This paper describes FRACAS' fluoroscopic image processing techniques and experimental results. To correlate the images and use them for accurate registration and navigation, we correct the distortion, determine the fluoroscopic camera parameters, and extract the bone contours. While many methods for these tasks are described in the literature, our method emphasizes integration, full automation, simplicity, robustness, and practicality. It focuses on fluoroscopic bone images and their use in 2D/3D anatomy-based registration. We report experimental results quantifying accuracy, distortion, and camera parameter estimation, and present segmentation results of several sets of clinical images.

2 Problem characteristics

Fluoroscopic images present substantial distortion due to three factors [1, 3, 6, 14]: (1) the image intensifier receptor screen is slightly curved, (2) the surrounding magnetic fields of the Earth and nearby instruments deflect the X-ray beam electrons, and (3) the C-arm armature deflects under the weight of the image intensifier, changing the focal length of the camera. The first effect can be modeled as radial pincushion distortion and is independent of the C-arm location. The second effect yields image translation and spatially variant rotation, and is C-arm orientation dependent. The third effect requires knowing the magnitude of the deflection. The distortion pattern resulting all the three factors is present in all units, including modern ones, and varies from unit to unit and from session to seses sion, with up to 10mm shift on the image edges [14].

Prior to each session, the amount of distortion and the fluoroscopic camera parameters must be determined for predefined C-arm orientations [3]. Usually, the dewarping and camera calibration steps are decoupled [8, 11, 14] and the parameters are obtained by imaging specially designed phantoms. To correct for distortion, a uniform grid of fiducials (e.g. steel balls or holes) is imaged. The location of the fiducial centers in the image and in the model are compared, and a *dewarp map* is computed specifying how to shift each pixel in the image to its real projected location. To obtain the camera characteristics, a parametric pinhole camera model defined by simultaneous equations relating the parameters is used. An object with known fiducial geometry and known location is imaged, and the positions of the image and the geometric points are matched. Solving the set of equations yields the parameter values [15].

Once the images have been corrected for distortion and the camera parameters are know, the next step is contour extraction of relevant anatomical structures. The main difficulties are that the images are noisy, have limited resolution, exhibit non-uniform exposure variation across the field of view, and have varying contrast and exposure from shot to shot. Common image processing techniques [17] yield poor results, with under and over segmentation, high sensitivity to threshold values, and require frequent threshold adjustments.

3 Materials and methods

We use Phillips BV 25 units with 9" field of view in all our experiments. The images were transfered from the fluoroscope's video output port to the computer using a frame grabber with a resolution of 720x560 and pixel size of 0.44mm. We built a custom dewarping grid and calibration object which is fitted via existing screw holes to the image intensifier plate. The design is inexpensive, simple to manufacture, and lightweight, to minimize additional C-arm deflection.

The dewarping grid is a 7mm thick coated aluminum alloy plate with 405 4mm diameter holes uniformly distributed at 10mm intervals machined to 0.02mm precision. It is simpler and cheaper to make than the commonly used steel balls mounted on a radiolucent plate and yields similar results. The calibration object (Fig. 2) is a hollow Delrin^{TM} three-step cylinder with eighteen 5mm diameter steel balls in three parallel planes angularly distributed to avoid overlap in the image. An additional ball in the top circular face marks the center of the object. A rectangular bar, affixed to the bottom of the cylinder, has holes that

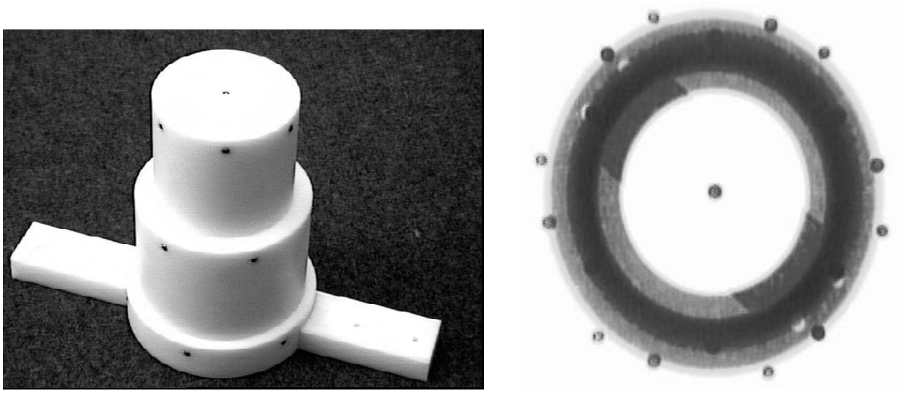

Fig. 2. Calibration object (left) and its fluoroscopic image (right).

allow mounting the object directly on the image intensifier plate. The balls are mounted at heights of 20, 100, and 180mm from the cylinder base, forming circles of 130, 115, and 90mm diameters respectively. The object weights 1.5kg.

To determine the intrinsic accuracy and repeatability error of the system, we acquired five series of images of the dewarping grid at a fixed C-arm orientation and exposure. We observed small relative rigid motion between shots introduced by the frame grabber. We correct for this motion in all our images by shifting the image pixels so that the center of the fluoroscope's circular field of view is always in the same position. Once this shift was corrected, we measured the distances between matching hole centers in pairs of images. For 1389 measurements, the mean error was $mean = 0.038$mm with standard deviation $\sigma = 0.032$mm, minimum $min = 0.001$mm, and maximum $max = 0.227$mm. Since the error is almost an order of magnitude smaller than other errors, we conclude that there is no need to take several exposures and average between them, as done in [14].

4 Image dewarping

Fluoroscopic image dewarping has received considerable attention [3]. It consists of computing a dewarp map from a reference image of a fiducials grid attached to the image intensifier plate and from the known fiducials centers geometric coordinates. The map is obtained in four steps: (1) identify the fiducials in the image from the background, (2) compute the coordinates of each fiducial center to sub-pixel accuracy, (3) pair the image and geometric fiducial centers, and (4) compute for each pairing the correction from the distances between the image and geometric fiducial center coordinates. New undistorted images are produced by computing for each pixel in the distorted image its new location and grayscale value in the undistorted image according to the dewarp map.

Global methods [3, 8] model the distortion across the entire image as a single function (e.g., a bivariate polynomial), whose coefficients are determined by least

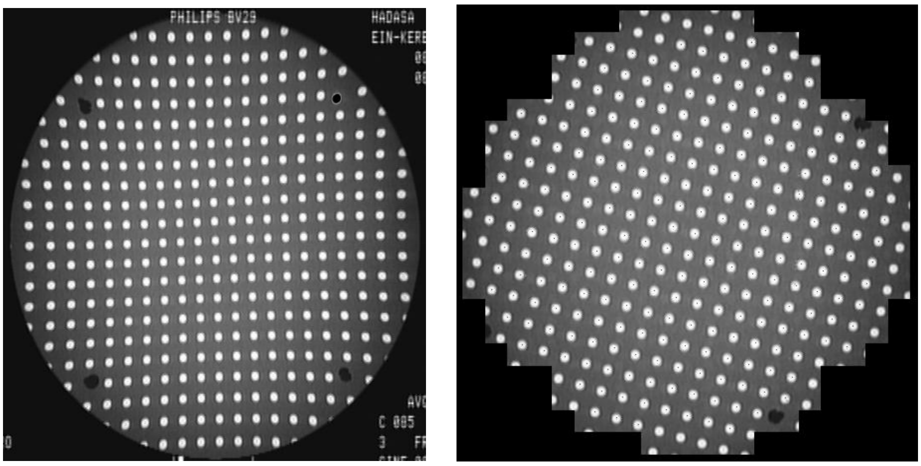

Fig. 3. Fluoroscopic images of the dewarping grid in its original position (left) and in its new position after dewarping (right). Black dots mark detected hole center points.

squares fitting of the image and geometric center coordinates. Local methods [6, 14] model the distortion by tesselating the image field of view into triangles or quadrilaterals for which individual distortion functions are computed. The functions are determined by the distances between the image and geometric fiducial center coordinates, usually by bilinear interpolation. Global methods produce compact maps, but assume that the distortion in the image is smooth and continuous. Local methods make no assumptions on the nature of the distortion and model it more accurately when it varies considerably across the field of view. Recently [3] reported comparable results when using local bilinear interpolation and global 4th-order polynomials.

We chose local bilinear interpolation because of its simplicity, computational efficiency, and generality in modeling unknown distortions. The procedure is simple to use and, unlike some others, does not require any user input for hole segmentation and center identification. The gray-scale value used in hole segmentation is automatically determined by finding the saddle point of the grid image histogram. Empirically, this proved to be an adequate threshold which yielded correct segmentation regardless of exposure setting and C-arm orientation. The center hole coordinates are computed to sub-pixel accuracy by weighted pixel gray scale average. To compute the dewarp map, the program tessellates the field of view into quadrilaterals whose endpoints are the hole center points. It uses the bilinear radial function to compute the undistorted coordinates of each image pixel. The coefficients for each region are obtained by solving a set of eight linear equations expressing the distances from the quadrilateral endpoints. The gray scale value of each new undistorted pixel is also obtained by pixel gray scale value bilinear interpolation. Since the C-arm orientation – its pitch and yaw – influences the dewarp map, we acquire distortion maps for a set of predeter-

Pose angle	Mean	Std Dev	Min	Max
(0, -10)	0.381	0.201	0	0.890
(0, 10)	0.415	0.205	0	0.937
(0, -15)	0.390	0.203	0	0.991
(0, 15)	0.313	0.193	0	0.838
(10, 15)	0.489	0.211	0	0.946
(10,-15)	0.344	0.201	0	0.913
(-10,15)	0.310	0.193	0	0.870
(0, 90)	1.931	0.541	0	2.917
(80, 0)	2.708	0.861	0	4.219
(0, 180)	2.550	0.703	0	3.717

(a) dewarping results

Param	Mean	Std Dev	Min	Max
T_x (mm)	0	0.882	-1.300	1.185
T_y (mm)	0	0.251	-0.317	0.339
R_x (deg)	0	0.342	-0.393	0.348
R_y (deg)	0	0.145	-0.233	0.169
R_z (deg)	0	0.213	-0.407	0.170
T_z (mm)	915.756	15.129	891.825	929.508
f	48.598	0.772	47.433	49.402
C_x	257.544	0.182	257.289	257.760
C_y	203.815	0.085	203.699	203.960
κ	0.00013	0.00001	0.00012	0.00015
s	1.00032	1.00283	1.00165	0.0009

(b) calibration results

Table 1. (a) Distance variation of 348 pairs of furthest center points for different C-arm orientations. All distance measurements are in millimeters relative to the pose angle yaw=0^o, pitch=0^o. (b) Calibration parameters nominal values and sensitivity to C-arm orientation. The extrinsic parameters T and R are with respect to a coordinate frame on the center of the image intensifier.

mined orientations. We compute dewarp maps for the most common poses, e.g., anterior-posterior and lateral views, and small angular neighborhoods around them, instead of acquiring many maps for different angular segments and interpolating between them [1, 3].

To quantify how sensitive the dewarp map is to changes in C-arm orientations, and thus to determine how many predetermined orientations must be captured, we computed distortion maps at different orientations. Table 1(a) summarizes the results. We observe a significant point center shift of up to 4mm between extreme C-arm orientations and of almost 1mm for orientation of 15^o apart. To determine the accuracy of the dewarp map function on new images, we acquired an image of the grid attached to the image intensifier cover at a fixed C-arm orientation and computed the dewarp map. Then, we detached the grid, placed it at an arbitrary angle on the cover, acquired a new image, and corrected it with the dewarp map (Fig. 3). We located the image hole centers in the new dewarped image with the hole segmentation routine, and computed a worst case error bound by taking the relative distances between pairs of points that are furthest apart. For 30 measurements, the mean error was 0.104mm, with σ = 0.060mm, min = 0.007mm, and max = 0.198mm. Previous studies report similar residual errors after correction.

5 Camera calibration

We use Tsai's [15] 11 parameter pinhole camera model and solution method to model the fluoroscopic camera. Since the parameters are pose dependent, we compute them for the same C-arm orientations as for dewarping. The parameters are the relative position $T = T_x, T_y, T_z$ and orientation $R = R_x, R_y, R_z$ of the pinhole with respect to the imaging plane, the focal length f, the image center location C_x, C_y, and the image scaling and radial distortion coefficients, s and κ. Because the images have been previously corrected, the radial image distortion assumption holds. We could set the parameters $s = 1$ and $\kappa = 0$, but we compute them anyway to further verify the dewarping procedure.

The set of equations relating the parameters are obtained by formulating the transformations from the world coordinate to the camera coordinates, transforming the 3D camera coordinates into 2D coordinates in an ideal undistorted image, and adding radial distortion, shifting, and scale. The equations can be solved in two steps, based on the radial alignment constraint. Following the camera calibration procedure for single view non-coplanar points, the extrinsic parameters R and T, with the exception of T_z, are found by solving a set of linear equations. Based on these values, the remaining parameters are derived. While this method requires at least seven points, we use the least squares method to incorporate more points.

To determine the variation for the different C-arm orientations, we conducted measurements for six extreme orientations. Table 1(b) summarizes the results. Note that the variation in T_z, which measures the distance between the camera pinhole and the image plane, is significant and confirms the deflection of the C-arm [6]. The small radial distortion and scaling deviations show that the dewarping procedure is very accurate. To quantitatively validate the accuracy of the calibration, we imaged the calibration object and computed the calibration parameters. We then constructed the projection matrix and used it to compute the geometric coordinates of the ball centers. For each ball, we computed the distance between the geometric and the image coordinate centers. The mean distance error for 78 measurements is 0.201mm with $\sigma = 0.089$mm, $min = 0.033$mm, and $max = 0.449$mm.

6 Contour extraction

Reliably extracting bone contours in fluoroscopic images is difficult because the images are noisy, have limited resolution, exhibit non-uniform exposure variation across the field of view, and have varying contrast and exposure from shot to shot. The bone structures are surrounded by tissue, contain overlapping contours, and have internal contours. Since our ultimate goal is to register the images with 3D contour models, we aim at obtaining a sufficiently dense set of points, possibly disconnected, on the bone contour with the fewest possible number of outliers.

We ruled out top-down model-based segmentation methods because they are difficult to use for anatomical structures. We considered the three main bottom-up approaches to contour segmentation: edge detection, active contours, and region growing [17]. Our experiments with standard edge detection techniques on actual fluoroscopic images showed that the Marr-Hilderth edge operator is overly sensitive to noise and non-uniform image exposure, producing too many false contours. The Canny edge detector yielded better results but required extensive threshold adjustments for every image, with frequent over and under segmentation. We ruled out the active contours techniques because they cannot detect overlapping contours require an initial guess near the target, and are computationally expensive [1]. The region growing methods yielded better results but created many spurious boundaries because of the non-uniform exposure across the field of view.

We developed a new bone contour segmentation algorithm based on robust image region statistics computation [4]. Its main advantage is that it adaptively sets local segmentation thresholds from a robust statistical analysis of image content. Working on the gradient image, it starts from global threshold setting and performs region growing based on adaptive local thresholds and zero-crossings filtering. Because the algorithm uses both global and local thresholds, it is less sensitive to the exposure variations across the field of view. Pixels are classified into one of three categories, *bone*, *candidate*, or *background*, according to the number of pixels above a predefined percentile, and not according to a prespecified absolute value. The percentile indicates the number of pixels in the gradient image histogram with gray values below (background) or above (bone), with candidate pixels in between. Initial region classification is obtained with global percentile thresholds. To overcome the non-uniform exposure, the classification is adaptively updated with local percentile thresholds over a fixed size window. Filtering the result with the original image zero crossings localizes the contour inside the region. The contour segmentation inputs global and local, upper and lower percentile thresholds, and a window size. It finds edge pixels in four steps:

1. Initial global classification
Compute the gradient image and its histogram. Set the global threshold values according to the given global image percentiles. The gradient image pixels are classified according to the global thresholds as *background* (below the lower threshold), *bone* (above the upper threshold), or *candidate* (between the lower and upper thresholds).

2. Revised local classification
For each *candidate* pixel in the gradient image, place a local window of prespecified size centered at the pixel and compute the local thresholds from its histogram. The pixel label is modified according to the local threshold values.

3. Region growing and small components elimination
Recursively relabel as *bone* all pixels labeled *candidate* with one or more neighboring *bone* pixels (either the four or eight neighboring scheme can be used). Next, remove all connected *bone* pixel components with too few pixels (e.g., less than 50) by relabeling them as *background*. They are most likely noise.

4. Filtering with zero-crossings image
Compute the binary zero-crossings image of the original image and perform an AND operation with the binary labeled gradient image. The labeled gradient image is converted to a binary image by setting *bone* pixels to 1 and *background* pixels to 0. The result are the pixels on the bone contours.

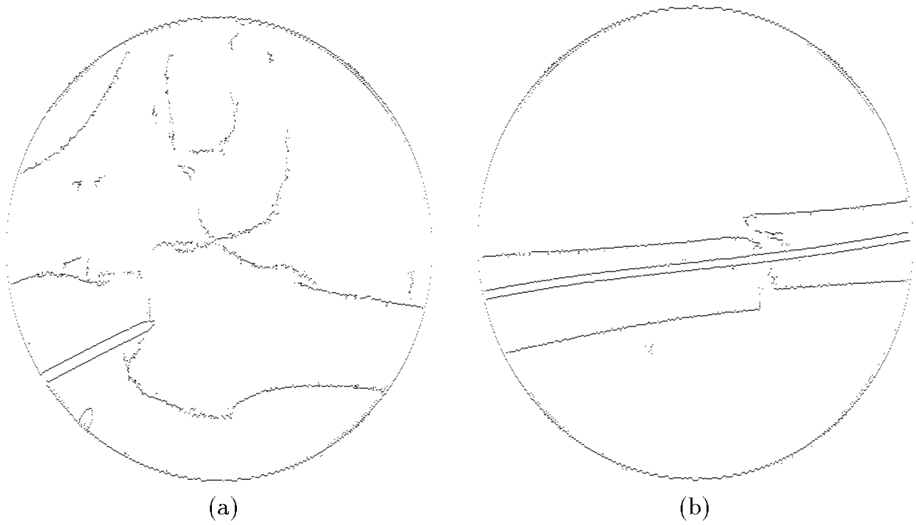

Fig. 4. Extracted pixel contours of the fluorosocopic images in Fig. 1. The contours of the metal tools were also segmented because metal density is higher than bone density.

We conducted a preliminary evaluation of the contour extraction algorithm on three sets of fluoroscopic images taken from actual surgeries. The global and local gradient image threshold percentiles (lower=60%, upper=94.7%, lower=60%, upper=99%), window size (13x13 pixels^2), and number of neighbors ($n = 4$) were kept constant for all images in a session. Fig. 4 shows typical results. Note that there are very few outliers, which can be removed with a simple model-based scheme or by combining segmentation and registration, as in [5].

7 Conclusion

We have presented a practical approach to fluoroscopic image processing consisting of image dewarping, camera calibration, and bone contour extraction. We report experimental results quantifying the accuracy of our setup and methods. We found an intrinsic system accuracy of 0.04mm, average dewarping accuracy of 0.1mm and always below 0.2mm, and average calibration error of 0.2mm and always below 0.45mm. We found significant dependence on the C-arm orientation, with dewarp variations of up to 1mm for angles of 10^o or more, and as much as 4mm between extreme orientations. Changes in the C-arm deflection have also a significant effect on the calibration parameters. These results match those of previous studies and suggest that submillimetric spatial positioning accuracy is achievable with standard equipment. Preliminary results of the contour segmentation algorithm show good contour tracing and very few outliers. Our current work focuses on registering the extracted bone contours with surface bone models obtained from preoperative CT images and on designing experiments to establish the overall accuracy of our method.

References

1. Brack, C., Burghart, R., Czpof, A., *et. al* "Accurate X-ray Navigation in Computer-Assisted Surgery", *Proc. of the 12th Int. Symp. on Computer Assisted Radiology and Surgery*, H. Lemke *et al* eds., Springer, 1998.
2. Brumback R.J., "Regular and Special Features - The Rationales of Interlocking Nailing of the Femur, Tibia, and Humerus" *Clinical Orthopaedics and Related Research*, **324**, Lippincott-Raven, 1996.
3. Fahrig, R., Moreau, M., Holdsworth, D.W., "Three-Dimensional Computed Tomographic Reconstruction Using a C-arm Mounted XRII: Correction of Image Intensifier Distortion" *Med. Phys.* **24**(7), July 1997.
4. Garza-Jinich, M., Meer P., Medina, V., "Robust Retrieval of 3D Structures from Magnetic Resonance Images" *Proc. of the Conf. on Pattern Recognition*, 1996.
5. Hamadeh, A., Sautot, P., Lavallée, S., Cinquin, P. "Towards Automatic Registration Between CT and X-ray Images: Cooperation Between 3D/2D registration and 2D Edge Detection", *Proc. 2nd Int. Symp. on Medical Robotics and Computer Assisted Surgery*, Baltomore, 1995.
6. Hofstetter, R., Slomczykowski, M. Bourquin, I, Nolte, L.P., "Fluoroscopy Based Surgical Nagivation – Concept and Clinical Applications", *Proc. 11th Int. Symp. on Computer Assisted Radiology and Surgery*, H.U. Lemke *et al.* eds, 1997.
7. Joskowicz, L., Tockus, L., Yaniv, Z, Simkin, A., Milgrom, C. "Computer-Aided Image-Guided Bone Fracture Surgery – Concept and Implementation", *Proc. 12th Int. Symposium on Computer Assisted Radiology and Surgery*, H.U. Lemke *et. al.* eds, Springer 1998.
8. Koppe, R., Klotz, E., Op de Beek, J., Aerts, H., "3D Vessel Reconstruction Based on Rotational Angiography", *Proc. of 9th Int. Symp. on Computer Assisted Radiology*, Springer, 1995.
9. Lavallée, S., Sautot, P., Troccaz, J., Cinquin, P., Merloz, P. "Computer Assisted Spine Surgery: a Technique for Accurate Transpedicular Screw Fixation Using CT Data and a 3D Optical Localizer", *J. of Image-Guided Surgery* 1(1), 1995.
10. Li Q.H., Holdener H.J., Zamorano L., *et. al.* Nolte L.P., Visarius H., and Diaz F., "Computer-assisted transpedicular screw insertion" *Lecture Notes in Computer Science*, Springer, 1996.
11. Phillips R., Viant W.J., Mohsen A.M.M.A., Griffits J.G., Bell M.A., Cain T.J., Sherman K.P. and Karpinski M.R.K., "Image Guided Orthopaedic Surgery - Design and Analysis" *IEEE Transactions on Robotics and Control*, March 1996.
12. Sanders R. "Exposure of the Orthopaedic Surgeon to Radiation", *J. of Bone Joint Surgery* 75 A(3), 1993.
13. Simon, D.A, Jaramaz, B, Blackwell, *et al.* "Development and Validation of a Navigational Guidance System for Acetabular Implant Placement", *Proc. of CVRMed-MRCAS'97*, Lecture Notes in Computer Science 1205, Springer 1997.
14. Schreiner, S., Funda, J., Barnes, A.C., Anderson, J.H., "Accuracy Assessment of a Clinical Biplane Fluoroscope for Three-Dimensionial Measurements and Targeting", *Proc. of SPIE Medical Imaging*, 1995.
15. Tsai, R., "A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses" *IEEE Journal of Robotics and Automation*, Vol. RA-3, No. 4, August 1987.
16. Viant, W.J., Phillips, R., Griffiths, et al. "A Computer Assisted Orthopaedic System for Distal Locking of Intramedullary Nails", *Proc. of MediMEC'95*, 1995.
17. Woods R.E., and Gonzalez R.C., *Digital Image Processing*, Addison-Wesley, 1992.