1-1-2006

An Unsupervised Fluoroscopic Analysis of Knee Joint Kinematics

Charles Scott
Boise State University

Elisa H. Barney Smith
Boise State University

This document was originally published by IEEE in *19th IEEE International Symposium on Computer-Based Medical Systems, 2006. CBMS 2006.* Copyright restrictions may apply. DOI: 10.1109/CBMS.2006.42

An Unsupervised Fluoroscopic Analysis of Knee Joint Kinematics

Charles Scott and Elisa H. Barney Smith
Boise State University, Boise, ID 83725-2075
EBarneySmith@BoiseState.edu

Abstract

Knowledge of the three dimensional positions of bones at a joint as a function of time is required to accurately model joint kinematics. 3-D bone geometry data from a static computer tomography (CT) images can be combined with time sequence information from 2-D video fluoroscopy images to produce 3-D position data over time. The process involves creating virtual X-rays from the CT image through digitally reconstructed radiograph (DRR) projections. Historically, the process of matching the 3-D and 2-D data has required human interaction. We have eliminated the need for manual initialization using a Monte Carlo technique with a variable search range. The search range decreases as the matching improves, avoiding the inherent problems of local minima in the optimization search. Experiments demonstrate that image positions can be matched to within 1 degree rotation, azimuth and elevation without human intervention.

1. Introduction

The position of bones in a joint under stationary conditions is straightforward to determine. The position and kinematics of joint bones during motion is more difficult to attain using standard non-invasive motion capture techniques, and is the subject of research for many researchers. Knowledge of the bone positions over time will quantify the kinematics and allow modeling and measurement of the forces present on tendons, ligaments and muscles during activities such as jump landings and deep knee bends. This can allow researchers to analyze sources of anterior cruciate ligament (ACL) injuries, wear on prosthetic implants and tracking of surgical instruments during surgery relative to preoperative Computed Tomography (CT) scans. The goal of this research is to develop a method for collecting accurate, three-dimensional kinematic data of bones and joints in vivo using a video fluoroscopy technique combined with CT data.

Research to aid clinical assessments of knee joint motion has been performed by placing markers on the skin to measure the dynamic movement in a joint [2]. This approach does not necessarily describe what the bones are doing because the sensors are attached to a soft tissue instead of the bone. To obtain greater accuracy of bone movement, markers have been placed into the bones of canines and tracked as the joint moves [7]. This is not a suitable approach for use on humans.

Three dimensional position over time can be determined by matching CT data to 2-D fluoroscopy images. The CT data can be manipulated to create a virtual x-ray image of the joint in various positions, Figure 1a. These images can be matched to a true x-ray, Figure 1b, by adjusting the camera position. Once the images are matched, the spatial locations of the bones in the joint are known. This has been performed using joints with prosthetic implants. The fluoroscopic data is compared to a CAD model of the prosthetic joint [1, 5]. The CAD model is rendered in the desired position and from that 2-D rendering a silhouette of the prosthetic implant is produced. The silhouette image is then matched to the x-ray. Accurate three-dimensional data describing the orientation of the implant in the joint can be obtained using both of these methods because it is significantly easier to locate the edges of a metallic object in an x-ray as opposed to the edges of bone. These methods are also precise because the shape of the prosthetic joint is known in advance. Due to occlusion in the opaque projections of the implants, multiple equivalent projected images can result. This has been overcome by having a human make an initial positioning of the 3-D joints.

These methods have been extended to tracking natural bones by segmenting the bones in the CT model and using this information to create silhouettes. Contours of bone surface folds are also included in the projection to offer more detail to help reduce the effects of occlusion [4]. This projected image is then compared to a single plane radiograph for matching.

The full CT data can be used to obtain a virtual 2-dimensional rendering of the joint called a Digitally Reconstructed Radiograph (DRR). You et. al created DRRs from a 3-D CT model but utilized a biplane radio-

Proceedings of the 19th IEEE Symposium on Computer-Based Medical Systems (CBMS'06)
0-7695-2517-1/06 $20.00 © 2006 IEEE

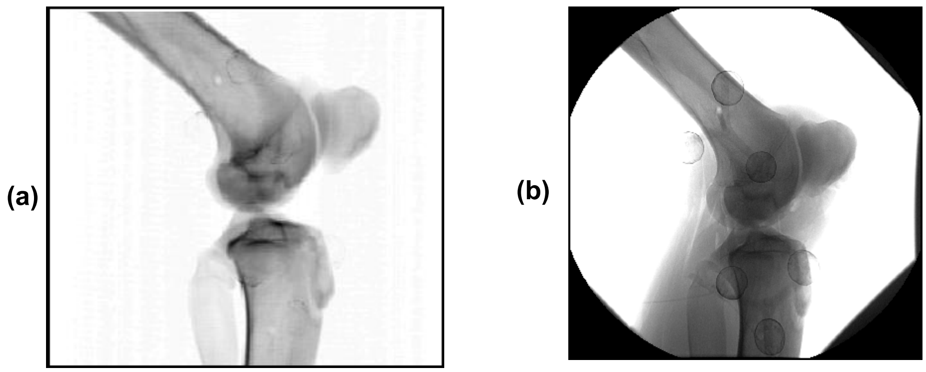

Figure 1: (a) Digitally Reconstructed Radiograph (DRR) (b) Fluoroscopy frame.

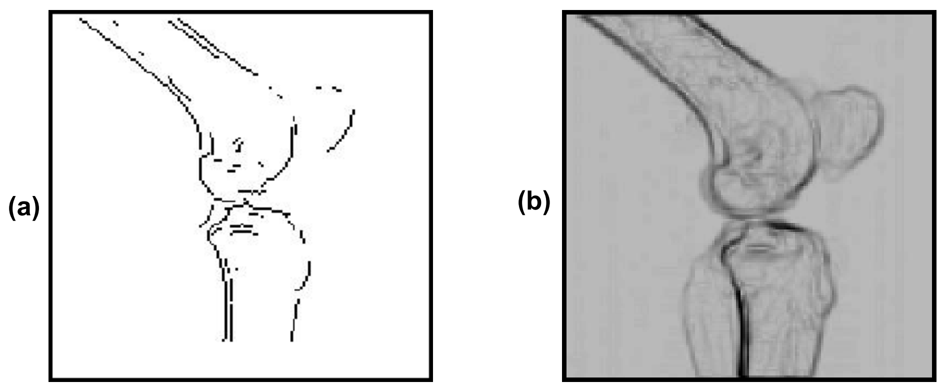

Figure 2: (a) bilevel DRR edge image (b) multi-level DRR edge image.

graph for the matching image [9]. This allows for greater accuracy but the system requires two x-ray sources to produce the biplane image, increasing x-ray exposure and system complexity.

Penny et. al [3, 8] used a single DRR compared to an X-ray image to track instrumentation during surgery relative to a CT taken prior to surgery. They manually restricted their field of view to a single vertebra to be sure the matching algorithm would select the vertebra of interest.

The proposed method is minimally invasive, does not require the subject to have a prosthetic joint, and uses relatively little equipment. It also removes the requirement of having a human intervene to initialize the position search. Section 2 will describe the proposed method in full. The experiments that have been conducted and the results attained are described in Section 3. Section 4 describes areas where these results can be expanded.

2. Method

The process begins by taking a static CT scan of the joint as well as fluoroscopic video of the joint while in motion. The CT data is used to create a DRR of the joint, Figure 1a, using a process outlined by Siddon [6]. If the simulated x-ray source and film are positioned correctly the resulting DRR should match the target fluoroscopic image frame. A coordinate system consisting of an x-ray source and film positioned at equal radial distances on opposite sides of the origin was created. The CT cube is nominally in the center of this sphere and the x-ray source and film move with an azimuth and elevation around this center. The CT cube can translate from the center position of this sphere. The film can also rotate its position. The diameter of this sphere (the distance from the x-ray source to the film) is set using apriori information about the fluoroscopy unit configuration during image capture. This model was created in such a way that it allows selection of the virtual camera positions with respect to the CT data in six degrees of freedom.

Once a DRR is obtained for a given pose, the edges of the target and search images are found. The edges in the two images are used to identify the match. Two different methods of extracting the edges are used. The first sums the square of the vertical and horizontal edges found using a Sobel operator and thresholds it at a twice the RMS value of the resulting image, Figure 2a. The second edge image is found by summing the absolute value of the images found with the horizontal, vertical and two diagonal Sobel operators, Figure 2b. The first of these methods produces a bilevel image, the second has a range of values.

The edge image produced by each DRR during the search process is compared to the edge image of the target image by evaluating the overlap of the edges using a contour match score

$$\text{contour match score} = \frac{\sum_{(x,y)} J(x,y)K(x,y)}{\sum_{(x,y)} K(x,y)} \quad (1)$$

where $J(x,y)$ is the target image, and $K(x,y)$ is the DRR search image [5]. For edge images that contain only 0's and 1's, the resulting edge score is between zero and one with a perfect match indicated by a one. Edge images with a range of intensity values will not have a match score of 1 even if there is a perfect fit. To produce a comparable metric, the edges of the multivalued target image are also compared to itself. Contour match scores are then calculated as in Equation 1 and are divided by this value to obtain a score between zero and one.

The process of determining the unknown variables is divided into two parts: an initial search and a fine search. The initial search is to get a set of position parameters in the approximate orientation of the bones. This replaces a manual approach to initialization used in [1]. Our method uses a Monte Carlo approach to randomly select values for the camera orientation of each DRR. This was chosen to avoid the problems of local minima that were cited in other work [5]. Initially an azimuth, elevation and rotation of 0 degrees is set. Likewise an initial position of (0, 0, 0) (x, y, z) translational mm is set. A DRR is created, then the azimuth, elevation and rotation angles as well as the three translational distances are all randomly perturbed. For the coarse search the variables are all perturbed by some amount. They are also peturbed by an equal amount in the opposite direction. A match score is calculated for the seed position, as well as the positive and negative perturbation positions. The position with the maximum match score of these three values is set as the new seed point. The size of this perturbation is gradually decreased as the match score increases. Then for the fine search, the variables are perturbed independently and the six resulting match scores are compared with the global maximum. By continually decreasing the search step-size the edge match score approaches 1.

For a fluoroscopic video sequence, only small changes in bone position are expected between frames. The search in subsequent frames will begin with the optimal position of the previous frame and only the fine search will be used for determining the position on these frames.

Once the camera position relative to the bones in the CT image has been established for each frame, the bone positions are also known. This data can then be used to construct a three-dimensional rendering of the joint in motion.

3. Experiments and Results

Data for our experiments was gathered from CT scans of a porcine knee. The CT volume was reduced by a factor of four in each dimension to reduce the computational complexity. This still retains the essential shape information of the joint bones and a significant amount of detail. The process was implemented in Matlab, with the DRR routine written in C and called via CMEX.

Two categories of experiments were run. Both experiments used DRR images at arbitrary poses as the target image rather than using actual Fluoroscopic image so the abilities of the edge features and search strategies could be evaluated. The first experiment evaluates the error between search and target positions and the value of the contour match score corresponding to the difference. Two target poses are chosen for this experiment. For each pose the edge images corresponding to that DRR are calculated. The search pose is then generated at elevation, azimuth and rotation values in 0.1° increments away from that pose. The corresponding contour match scores are calculated. From this experiment it was seen that the elevation would be the first to converge in a search. When the match score for bilevel edges was greater than 0.68 the error in elevation was less than 1°, when the match score was greater than 0.86 the rotation error was less than 1°, and when the contour match score was greater than 0.98, the azimuth error was less than 1°. For gray scale edges when the match score was greater than 0.70 the error in elevation was less than 1°, when the match score was greater than 0.89 the rotation error was less than 1°, and when the contour match score was greater than 0.98, the azimuth error was less than 1°. It was also observed that the contour match metrics, while not identical nor having perfect correspondence, were highly correlated; higher match scores for gray edges were also found when the bilevel images had higher match scores.

The second set of experiments evaluated the search strategy. For these experiments the translational position was fixed. The coarse search strategy started with a perturbation between 0 and 2 degrees until an edge score of 0.2 or higher is found. Then the perturbation reduces to a random value between 0 and 1 degree until the match score of 0.3 is reached. Then finally the search step is reduced from 1 degree to 0.6 degrees until a match score of 0.7 is attained. At that time the fine search begins. For the fine search the azimuth, elevation and rotation are each independently varied by a positive and negative measure between 0 and 0.6 degrees until a final match of 0.8 is attained.

To be able to know with certainty the closeness of the best fit, matching experiments were done by creating a target DRR. For gray level edges this was done at each of 28 poses. For each pose the search was run five times. Searching was done until a match score of 0.98 was attained. The aggregate results are shown in Table 1. In all experiments a match to within half a degree was found, and in most cases better for all three angles. Better matches could be found if the search was continued until a higher match score was attained. For bilevel edges the search was run for 15 poses and each continued until a match threshold of 0.85 was found. The results are shown in Table 2. The amount of error was less, but the time to converge was more. This is likely because for gray scale edges the edge information is spread over a wider range of pixels and small improvements in registration will make a noticeable difference in the match score, whereas when the match is close, the spread of edge information makes identifying the optimum more challenging.

Table 1: Aggregate Results of Search Experiment

Gray Edges	AZ (degrees)	EL (degrees)	ROT (degrees)
Mean	0.40	0.46	0.11
Std	0.33	0.34	0.07
Median	0.32	0.40	0.11
Min	0.00	0.00	0.00
Max	1.97	1.55	0.33

Table 2: Aggregate Results of Search Experiment

B/W Edges	AZ (degrees)	EL (degrees)	ROT (degrees)
Mean	0.14	0.14	0.04
Std	0.10	0.11	0.03
Median	0.11	0.12	0.04
Min	0.01	0.00	0.00
Max	0.41	0.44	0.10

4. Future work

The results reported above give a good basis for justifying further work on this method. Searches including the translational variables will show how these variables add to the process. Since the search using gray level edges converged faster and the search with bilevel images was more accurate, a search that starts with gray level edges and then switches to using thresholded edges after a certain convergence may provide better overall results.

The work shown to this point was for matching to a simulated x-ray. The search for the tibia and femur matches must be done separately to allow matches to fluoroscopy frames acquired of the same porcine knee where the bones may be in different positions relative to each other.

The fluoroscopic images need to be calibrated for their true position. Markers were attached to the bones (drilled in) and used to calibrate position using a two-camera digital video system. These markers are visible in the CT images, DRR and fluoroscopy images. Their positions need to be extracted and used to calibrate real world coordinates to the azimuth, elevation and rotation coordinates used in this work.

5. References

[1] Banks, S. A., and W. A. Hodge, “Accurate Measurement of Three-Dimensional Knee Replacement Kinematics Using Single-Plane Fluoroscopy,” IEEE Trans. Biomed. Eng., Vol. 43 no. 6, June 1996, pp. 638-649.

[2] Moorehead, J. D. and S. C. Montgomery, “A Surface-Marker Imaging System to Measure a Moving Knee's Rotational Axis Pathway in the Sagittal Plane,” IEEE Trans. Biomed. Eng., Vol. 48 no. 3, 2001, pp. 384-393.

[3] Penny, G., J. Weese, J. Little, P. Desmedt, D. Hill, and D. Hawkes, “A Comparison of Similarity Measures for Use in 2-D--3-D Medical Image Registration,” IEEE Trans. Medical Imaging, Vol. 17, No. 4, August 1998, pp. 586-595.

[4] Rahman, H., B.J. Fregly, and S. A. Banks, “Accurate Measurement of Three-dimensional Natural Knee Kinematics Using Single-Plane Fluoroscopy,” 2003 Summer Bioengineering Conference. Key Biscayne, Florida, 25-29 June, 2003.

[5] Sarojak, M., W. Hoff, R. Komistek, and D. Dennis, “An Interactive System for Kinematic Analysis of Artificial Joint Implants,” Proc. of 36th Rocky Mountain Bioengineering Symposium, Copper Mountain, Colorado, 16-18 April 1999.

[6] Siddon, R. L., “Fast calculation of the exact radiological path for a three-dimensional CT array,” Medical Physics 12(2), Mar/Apr 1985, pp. 252-255.

[7] Tashman, S., W. Anderst, P. Kolowich, S. Havstad, and S. Arnoczky, “Kinematics of the ACL-deficient Canine Knee During Gait: Serial Changes Over Two Years,” Journal of Orthopaedic Research, Vol. 22, 2004, pp. 931-941.

[8] Weese, J., G. Penny, P. Dsmedt, T. Buzug, D. Hill, “Voxel-Based 2-D/3-D Registration of Fluoroscopy Images and CT Scans for Image-Guided Surgery,” IEEE Trans. Imformation Technology in Biomedicine, Vol. 1, No. 4, December 1997, pp. 284-293.

[9] You, B., P. Siy, W. Anderst, and S. Tashman, “In Vivo Measurement of 3-D Skeletal Kinematics from Sequences of Biplane Radiographs: Application to Knee Kinematics,” IEEE Trans. Medical Imaging, Vol. 20 no 6, 2001, pp. 514-525.

Proceedings of the 19th IEEE Symposium on Computer-Based Medical Systems (CBMS'06)
0-7695-2517-1/06 $20.00 © 2006 IEEE