

International Journal of Computer Assisted Radiology and Surgery

Fluoroscopy-based tracking of femoral kinematics with statistical shape models

--Manuscript Draft--

Manuscript Number:	CARS-D-15-00079R1
Full Title:	Fluoroscopy-based tracking of femoral kinematics with statistical shape models
Article Type:	Review Article
Keywords:	Image processing; Statistical Shape Models; Digitally Reconstructed Radiographs; Computer Assisted Surgery
Corresponding Author:	Marta Valenti Politecnico di Milano Milano, ITALY
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Politecnico di Milano
Corresponding Author's Secondary Institution:	
First Author:	Marta Valenti
First Author Secondary Information:	
Order of Authors:	Marta Valenti Elena De Momi Weimin Yu Giancarlo Ferrigno Mohsen Akbari Shandiz Carolyn Anglin Guoyan Zheng
Order of Authors Secondary Information:	
Funding Information:	
Abstract:	<p>Purpose Precise knee kinematics assessment helps to diagnose knee pathologies and to improve the design of customized prosthetic components. The first step in identifying knee kinematics is to assess the femoral motion in the anatomical frame. However, no work has been done on pathological femurs, whose shape can be highly different from healthy ones.</p> <p>Method We propose a new femoral tracking technique based on Statistical Shape Models (SSMs) and two calibrated fluoroscopic images, taken at different flexion-extension angles. The cost function optimization is based on genetic algorithms, to avoid local minima. The proposed approach was evaluated on 3 sets of digitally reconstructed radiographic images of osteoarthritic patients.</p> <p>Results It is found that using the estimated shape, rather than that calculated from CT, significantly reduces the pose accuracy, but still has reasonably good results (angle errors around 2 degrees, translation around 1.5mm).</p>

Dr. Heinz U. Lemke,
International Journal of Computer
Assisted Radiology and Surgery

June 2015

Manuscript Reference Number: CARS-D-15-00079.

Cover Letter.

The authors thank the reviewers and the editor for their interest in this topic and for giving them the possibility to improve the paper with their comments. Hereunder we answer the questions arisen, with indications on the changes made to the manuscript

Reviewer #1:

Authors have addressed a very hot topic like the tracking of femoral kinematics from 2D fluoroscopy information. The manuscript is well organized and written.

However, a clear justification of the clinical application of the developments should be done. Since the acquisition of dynamics fluoroscopy can be translated into a high radiation dose, the use of a dynamic X-Ray acquisition to monitor articular function should be clearly indicated. A deep discussion on the added value of this technique over existing approaches to monitor femoral kinematics is missing.

According to the reviewer comment, we modified the text in the "State of the Art" section as follows:

"Knee kinematics assessment has great importance both to understand the problems associated with a large number of knee pathologies and to improve the design of prosthetic components. In case of severe osteoarthritis, that are eligible for joint implant surgery, in vivo pre-operative knee kinematics is fundamental to understand the relative motion between the three joint bones. The relative movement can give an insight of how the ligaments are stretched and their stability, and how could be the feeling for the patient. In this way, pre operative knee kinematics could help the surgeon to decide which prosthesis should be used and how to correct the misalignment of the bones [3].

Currently, the reconstruction of the pose of the knee can be done using 3D scans such as real-time Magnetic Resonance Imaging (MRI) or through a 2D/3D registration method that superimposes the shape extracted from MRI or Computed Tomography (CT) onto an image, usually X-ray or fluoroscopy. Real-time MRI is suitable to study joint kinematics, as it evidences the muscle structure during movements. However, it can only be used with relatively slow movements, and the accuracy obtained increases from 1mm to more than 3mm depending on the velocity of the movement. In addition, MRI scans are highly expensive [4]. Traditional CT and MRI provide an accurate evaluation of the morphology of the knee, but are limited to static positioning of the patient.

[...]

The use of dynamic fluoroscopy to detect knee kinematics is described in [12]. The authors use a fluoroscopic system flashing at 30Hz, obtaining continuous images of the knee flexion from 0° to 120°. A static CT is projected to reconstruct knee kinematics. A similar protocol is used in [13], where the authors describe a method to align CAD projections of knee implants to fluoroscopic images. They perform dynamic acquisitions at 8Hz moving a Sawbone model with an implant between 0° and 90°. In [14], the authors introduced Statistical Shape Models (SSMs) to reconstruct the shape of the femur. They performed analysis on the optimal number of X-ray scans that allow an accurate reconstruction of the shape of the knee. Their method was based on manual segmentation of the bone contour from X-ray images."

In the Discussion section, we emphasized the added value of our work over existing approaches

“Our approach shows the applicability of a semi-automatic algorithm for 2D/3D registration using SSM. With respect to dynamic MRI, our method is less expensive and more accurate [4]. Compared to actual clinical procedures, that are based on a previous static CT or MRI scan and a manual fitting of the shape, this method is less expensive (as CT and MRI can be avoided) and more accurate due to the automatic fitting. The ionizing radiation given to the patient in our procedure is just a few hundredths of a mSv (considering a total of 22 images and an average dose of 0.001mSv for each image, we have a total dose of 0.02mSv, much lower than a 2.2mSv average dose absorbed with a CT) [29, 30].

Compared to skin mounted markers, our system does not suffer from relative movements between the skin and the bone, and is thus suitable for accurate joint kinematics studies [6, 7, 8].

Our method proved to be stable without relying on implanted markers or joint implants, that will return more accurate results at the price of high invasiveness [5, 9, 11].”

3. Mohsen Akbari Shandiz. Component Placement in Hip and Knee Replacement Surgery: Device Development, Imaging and Biomechanics. PhD thesis, Biomedical Engineering, 2015.

4. CE Draper, JM Santos, LC Kourtis, TF Besier, M Fredericson, GS Beaupre, GE Gold, and SL Delp. Feasibility of using real-time MRI to measure joint kinematics in 1.5 T and open-bore 0.5 T systems. *Journal of magnetic resonance imaging* 28.1 (2008): 158-166.

13. S Acker, R Li, H Murray, P St John, S Banks, S Mu, U Wyss, and K Deluzio. Accuracy of single-plane fluoroscopy in determining relative position and orientation of total knee replacement components. *Journal of biomechanics*, 44(4):784-787, 2011.

29. Radiologyinfo.org. <http://www.radiologyinfo.org>. Accessed: 2015-06-05.

30. J Henckel, R Richards, K Lozhkin, S Harris, FM Rodriguez y Baena, ARW Barrett, and JP Cobb. Very low-dose computed tomography for planning and outcome measurement in knee replacement the imperial knee protocol. *Journal of Bone & Joint Surgery*, British Volume, 88(11):1513-1518, 2006.

Reviewer #2:

The paper describes a method of tracking the 3D position of bones in the knee from two flouroscopy views. The approach is to first fit the 3D shape model to the two 2D images at the first time point, using an optimisation. This estimates both the shape and 3D pose of the bone at that time. In subsequent frames the shape is fixed and the 3D pose optimised to best fit the 3D model to the two 2D flouroscope images. The system is validated on synthetic radiographs generated from real CT images of patients. The reconstructed shape and pose at each time was compared to the shape measured from the CT and the known pose. In addition, tracking was tested using the measured shape (to explore the effect of using estimated shape on the pose). It is found that using the estimated shape, rather than that calculated from CT, significantly reduces the pose accuracy, but still has reasonably good results (angle errors around 2 degrees, translation around 1.5mm).

Issues: - The method requires manual initialisation (p4) of the landmarks on the first frame. Given that fully automatic methods exist to locate points, why not use them (or at least mention them). For instance, Active Shape Models, Active Appearance Models, or more recent methods such as: T.F.Cootes, M.Ionita, C.Lindner and P.Sauer, "Robust and Accurate Shape Model Fitting using Random Forest Regression Voting", ECCV 2012

We thank the reviewer for the comment and for the suggestion. In our paper, we must differentiate between two types of landmarks: the first type are 3D landmarks on the shapes as are usually called in literature [18,19]. These landmarks are corresponding point on different shapes, that are found to construct the mean shape. The second type of landmark, used for the initialization, are points on the mean shape that are easily recognizable in the fluoroscopic images. As suggested by the reviewer, these landmarks could be automatically extracted using a proper algorithm. However, we chose to do it manually to ensure a low computational time and no requirement of organizing enough training data when the 2D images that are used are not acquired from standard orientation such as antero-posterior (AP) or lateral-medial (LM) direction.

We modified the Methods section, paragraph 2.1 as following:

Every instance of this SSM is then expressed as:

$$M_{SSM} = \bar{M} + \sum_{o=1}^{O'} \alpha_o \vec{M}_o$$

Where α_o is the weight corresponding to the o^{th} eigenvector and O' is the number of significant eigenvalues ($O' < O$). In our dataset, $O = 24$, $O' = 23$.

For this study, we used 24 MRI datasets of healthy knees, from which we semi-automatically segmented the femur contours using Amira R (VSG|FEI, Mèrignac Cedex, France). From each volume, we extracted the femur's triangulated surface with 2562 points and 5120 facets using the marching cubes technique [19]. To create the SSM, corresponding points on different shapes have to be determined [20] [21]. In our case, corresponding points between the O shapes were selected using the automatic algorithm described in [18]. Within the corresponding models, we computed the mean model and the covariance matrix. All the patients signed an informed consent and the institutional review board approved the study.

In paragraph 2.2 we modified the text as follows

The femur tracking dataset is composed of two fluoroscopic images (frames) acquired during knee flexion-extension at each time step t with $t = 0, \dots, T$.

The contour of the distal femur is extracted on each fluoroscopic image.

We defined a set of I ($I = 7$) landmarks on the mean model ($\mathcal{L}_{i(\bar{M})}$) that were used to initialize the pose of the SSM in the world Reference Frame (RF) RF_w . The landmarks were identified as the most identifiable anatomical landmarks in the images.

Recent studies have addressed the topic of automatically segment contours from X-Ray images using random forests [22, 23]. However, this methods needs to be trained over a consistent number of images to be used with non-standard projections. In order to save time and to give the possibility to use any angle of projection the operator is asked to manually identify the I corresponding landmarks on the two fluoroscopic images

20. Timothy F Cootes, Christopher J Taylor, David H Cooper, and Jim Graham. Active shape models-their training and application. Computer vision and image understanding, 61(1):38-59, 1995.

21. Timothy F Cootes, Gareth J Edwards, and Christopher J Taylor. Active appearance models. IEEE Transactions on pattern analysis and machine intelligence, 23(6):681-685, 2001.

22. Tim F Cootes, Mircea C Ionita, Claudia Lindner, and Patrick Sauer. Robust and accurate shape model fitting using random forest regression voting. In Computer Vision{ECCV 2012, pages 278-291. Springer, 2012.

23. Claudia Lindner, S Thiagarajah, J Wilkinson, The Consortium, G Wallis, and Timothy F Cootes. Fully automatic segmentation of the proximal femur using random forest regression voting. Medical Imaging, IEEE Transactions on, 32(8):1462-1472, 2013.

- Eq.3 measures the error in the 3D world frame. Is it not better to measure the error in the image frame? Although it is a harder optimisation problem it is a better metric.

Eq. 3 represents the minimization of the distance between the 7 landmark points used for the initialization on the shape and the 14 points identified by the user on the images. The fastest method to get a rough initialization of the shape's pose was to identify the corresponding points and implement a non-linear least squares algorithm in the world reference frame to get the transformation.

To use the error in the image frame would have meant to project the landmark of the 3D model on both image planes at each iteration, and would have enlarged the computational times. Since the initialization is not the definitive pose for the shape, the authors thought that it was simpler to keep the 3D minimization, even if not as accurate as the 2D one.

- Data. It is not clearly stated how many patient images were used in the evaluation, or what their disease status is.

According to the reviewer comment, we modified the text in paragraph 2.4.1 section as follows:

"In order to validate our tracking algorithm we used the CT scans of three different patients, which have respectively severe, mild and moderate osteoarthritis of the knee."

One of the main claims of novelty over other work such as [1] is that this paper demonstrates work on knees of people with disease ([1] did not). However, there is no description of how many knees were tested, or how diseased they were, or whether there is actually a difference in performance between normal and diseased knees.

According to the reviewer comment, we modified the text in paragraph 2.4.1 section as follows:

“In order to validate our tracking algorithm we used the datasets of three different patients, which have respectively severe, mild and moderate osteoarthritis of the knee. For each patient, we have the CT scan and we generate the DRRs.

The CT dataset was composed of DICOM images acquired with a SIEMENS Sensation 64 CT machine. Each slice is 512x512 pixel (0.3516 mm/pixel) with a slice thickness of 0.6 mm and a spacing between slices of 0.4 mm. In order to assess the model reconstruction performances we segmented the bone shape using Amira (VSG|FEI, France)..

For each patient, a virtual environment around the CT was created, in order to simulate fluoroscopic scans from three different point of views. 11 angles of flexion were simulated for a total of 11 x 3 patients x 3 views = 99 images. The center of the CT dataset was taken as the F_w .”

Overall the paper describes a method which seems to work (on the synthetic radiographs it was tested on), but does not seem to be a particularly significant advance over existing published methods (such as [1]). Significant information about the test set appears to be missing.

According to the reviewer comment, we modified the text in the Discussion section as follows:

The advances of the proposed new methodology with respect to the current state of the art resides in the fact that our method has been applied to pathological patients resulting in accuracy comparable to the current state of the art. Moreover, we managed to achieve limited decrease in performance with a lower angle (up to 10°) between the fluoroscopic projections. In this way, the range of movement for the flexion-extension of the knee is enlarged, extending the possibilities for the tracking with different movements.

Minor points:

Abstract: pathological -> pathological

p5 The "affine transform" is actually a similarity transformation (scale, rotation, translation)

p6 "pixel on each" -> "pixels on each"

p6 The variable $c_{k_n}(n)$ is undefined. What is it?

We modified the description of $C_{k_n(n)}$ in order to make it clearer:

$C_{k_n(n)}$, the closest point on the line between the line and the corresponding silhouette point

Table 1: What are the uncertainties - are the standard deviations?

We modified the caption of the table, in order to make it clearer:

Table 1 Mesh comparison statistical parameters: distances from point to surface represented as mean and standard deviation

Fig.8 : "exted" -> "extend"

Noname manuscript No. (will be inserted by the editor)
--

Fluoroscopy-based tracking of femoral kinematics with statistical shape models

Marta Valenti · Elena De Momi ·
Weimin Yu · Giancarlo Ferrigno ·
Mohsen Akbari Shandiz · Carolyn
Anglin · Guoyan Zheng

Received: date / Accepted: date

Abstract *Purpose* Precise knee kinematics assessment helps to diagnose knee pathologies and to improve the design of customized prosthetic components. The first step in identifying knee kinematics is to assess the femoral motion in the anatomical frame. However, no work has been done on pathological femurs, whose shape can be highly different from healthy ones.

Method We propose a new femoral tracking technique based on Statistical Shape Models (SSMs) and two calibrated fluoroscopic images, taken at different flexion-extension angles. The cost function optimization is based on genetic algorithms, to avoid local minima. The proposed approach was evaluated on 3 sets of digitally reconstructed radiographic images of osteoarthritic patients.

Results It is found that using the estimated shape, rather than that calculated from CT, significantly reduces the pose accuracy, but still has reasonably good results (angle errors around 2 degrees, translation around 1.5mm).

Keywords Image processing · Statistical Shape Models · Digitally Reconstructed Radiographs · Computer Assisted Surgery

1 State of the art

Knee kinematics assessment has great importance both to understand the problems associated with a large number of knee pathologies and to improve

M. Valenti · E. De Momi · G. Ferrigno
Department of Electronics, Information and Bioengineering
Politecnico di Milano, via Colombo 40, 20133 Milano, Italy
Tel.: +39 02 2399 9502
E-mail: marta.valenti@polimi.it

W. Yu · G. Zheng
Universität Bern, Staffaucherstr. 78, 3014 Bern, CH

M. Akbari Shandiz · C. Anglin
University of Calgary, 2500 University Drive NW Calgary, AB, Canada, T2N 1N4

the design of prosthetic components [1, 2]. In case of severe osteoarthritis, that are eligible for joint implant surgery, in vivo pre-operative knee kinematics is fundamental to understand the relative motion between the three joint bones. The relative movement can give an insight of how the ligaments are stretched and their stability, and how could be the feeling for the patient. In this way, pre-operative knee kinematics could help the surgeon to decide which prosthesis should be used and how to correct the misalignment of the bones [3].

Currently, the reconstruction of the pose of the knee can be done using 3D scan such as real time Magnetic Resonance Imaging (MRI) or through a 2D/3D registration method that superimposes the shape extracted from MRI or Computed Tomography (CT) onto an image, usually X-ray or fluoroscopy. Real-time MRI is suitable to study joint kinematics, as it evidences the muscle structure during movements. However, it can only be used with relatively slow movements, and the accuracy obtained increases from 1mm to more than 3mm depending from the velocity of the movement. In addition, MRI scans are highly expensive [4]. Traditional CT and MRI provide an accurate evaluation of the morphology of the knee, but are limited to static positioning of the patient. A great number of pathologies, such as patellofemoral pain and osteoarthritis require a dynamic evaluation of the knee motion [5, 1, 3].

Three main methods are used in research centers to in vivo assess knee kinematics: using skin-mounted markers, implanted markers and 2D image based methods.

Skin-mounted retro-reflective markers are used in gait analysis [6, 7, 8] using optical localization techniques. Since this type of measurement suffers from relative motion between the skin and the bone, results are not reliable to properly investigate joint kinematics.

The second method is based on tantalum markers implanted in the bone; even if this method is much more precise than the previous one, it is not commonly used in clinics due to the high invasiveness of the intervention [9, 5, 10]. In [11], the authors used mono and biplane fluoroscopic images to accurately reconstruct the pose of the patella, implanting markers in cadaver specimens and artificially flexing the knee joint.

The use of dynamic fluoroscopy to detect knee kinematics is described in [12]. The authors use a fluoroscopic system flashing at 30Hz, obtaining continuous images of the knee flexion from 0° to 120° . A static CT is projected to reconstruct knee kinematics. A similar protocol is used in [13]; the authors describe a method to align CAD projections of knee implants to fluoroscopic images. They perform dynamic acquisitions at 8Hz moving a Sawbone model with an implant between 0° and 90° . In [14], the authors introduced Statistical Shape Models (SSMs) to reconstruct the shape of the femur. They performed a deep analysis on the optimal number of X-ray scans that allow an accurate reconstruction of the shape of the knee. Their method was based on manual segmentation of the bone contour from X-ray images. In [1], the authors introduced a fully automatic technique to extract the contour from the fluoroscopic images, based on a Canny edge detector [15]. They used two fluoroscopic sequences of drop-landing motion with intra-fluoroscopic distance

between $58^\circ - 82^\circ$. They used a cadaveric knee as gold standard, implanting tantalum markers and simulating motion of the knee while doing fluoroscopic imaging of the joint. Their mean errors are less than 1 millimeter, spanning to 4 mm both with the CT-extracted shape and the SSM. However, the method proposed by Baka was only evaluated on healthy subjects, without evidence that the same accuracy could be obtained with osteoarthritic patients.

The aim of this work is to propose a method able to reconstruct the 3D pose of the femur using biplane fluoroscopies and a SSM. We used osteoarthritic femurs in order to expand the validity of the method not only to healthy subjects. The method was evaluated generating Digitally Reconstructed Radiographs (DRRs) from static CT acquisitions of osteoarthritic patients and simulating femur flexion-extensions. We evaluate the accuracy of the pose reconstruction of the distal femur both with the CT segmented shape and SSM.

2 Methods

Our femoral kinematics tracking method is based on SSM whose pose is obtained using a biplane fluoroscopy. We applied a Genetic Algorithms (GA) optimization technique, in order to improve the accuracy.

The tracking algorithm is divided into three main phases:

- Statistical Shape Model creation (2.1)
- Shape pose initialization (2.2)
- Femur tracking (2.3)

2.1 Statistical Shape Model (SSM) creation

SSM give an effective parameterization of the shape variations found in a collection of sample models of a given population [16, 17]. For our study we used the distal part of the femur only. Each bone model (\mathbf{M}_o with $o = 1, \dots, O$) is represented as an ordered set of $p = 1, \dots, P$ vertices $\mathbf{M}_o = [x_1, y_1, z_1, \dots, x_p, y_p, z_p, \dots, x_P, y_P, z_P]$ and a list of triangular facets connecting the vertices. Applying the algorithm described in [18], we found the correspondence between corresponding vertices on different shapes. The SSM is defined as the mean model $\bar{\mathbf{M}}$ and a set of eigenvectors obtained applying Principal Component Analysis (PCA) to the model vectors \mathbf{M}_o .

$$\begin{aligned} \mathbf{D} &= \frac{1}{O-1} \sum_{o=1}^O (\mathbf{M}_o - \bar{\mathbf{M}})(\mathbf{M}_o - \bar{\mathbf{M}})^T \\ \mathbf{D} \cdot \vec{\mathbf{M}}_o &= \sigma_o^2 \cdot \vec{\mathbf{M}}_o \\ \sigma_1^2 &\geq \sigma_2^2 \geq \dots \geq \sigma_{O-1}^2 \end{aligned} \tag{1}$$

where σ_o^2 are the descending-order eigenvalues of the covariance matrix \mathbf{D} and $\vec{\mathbf{M}}_o$ are the corresponding eigenvectors. Every instance of this SSM is then

expressed as:

$$\mathbf{M}_{SSM} = \bar{\mathbf{M}} + \sum_{o=1}^{O'} \alpha_o \vec{\mathbf{M}}_o \quad (2)$$

where α_o is the weight corresponding to the o^{th} eigenvector $\vec{\mathbf{M}}_o$ and O' is the number of significant eigenvalues ($O' < O$). In our dataset, $O = 24, O' = 23$. For this study, we used 24 MRI datasets of healthy knees, from which we semi-automatically segmented the femur contours using Amira[®] (VSG | FEI, Mérignac Cedex, France). From each volume, we extracted the femur's triangulated surface with 2562 points and 5120 facets using the marching cubes technique [19]. To create the SSM, corresponding points on different shapes have to be determined [20] [21]. In our case, corresponding points between the O shapes were selected using the automatic algorithm described in [18]. Within the corresponding models, we computed the mean model and the covariance matrix.

All the patients signed an informed consent and the institutional review board approved the study.

2.2 Shape pose initialization

The femur tracking dataset is composed of two fluoroscopic images (frames) acquired during knee flexion-extension at each time step t with $t = 0, \dots, T$. The contour of the distal femur is extracted on each fluoroscopic image.

We defined a set of I ($I = 7$) landmarks on the mean model ($\mathcal{L}_{i(\bar{\mathbf{M}})}$, with $i = 1, \dots, I$) that were used to initialize the pose of the SSM in the world Reference Frame (RF) ${}^w\text{RF}$. The landmarks were identified as the most identifiable anatomical landmarks in the images. Recent studies have addressed the topic of automatically segment contours from X-Ray images using random forests [22, 23]. However, these methods need to be trained over a consistent number of images to be used with non-standard projections. In order to save time and to give the possibility to use any angle of projection the operator is asked to manually identify the I corresponding landmarks on the two fluoroscopic images at $t = 0$ ($\ell_{i(n)}$ with $i = 1, \dots, I$ and $n = 1, 2$, where i indicates the number of the landmark and n indicates the view).

Each fluoroscopic image has its own calibration parameters, so that we know the position of the source and the position of the image plane in the ${}^w\text{RF}$, as well as its normal vector. The initialization is used to estimate the initial position of the SSM with respect to the 2 fluoroscopic images, i.e. ${}^w\mathbf{T}_m^0$. We can back-project each landmark pixel ($\ell_{i(n)}$) to the corresponding source \mathbb{S}_n . In this way, for each landmark we find 2 skew lines (one for each projection) that should (ideally) intersect in one point (the 3D position of the landmark) but, due to errors and noise, they actually do not intersect. Thus to define the position of the landmark we take the middle point of the shortest line connecting the rays (see Fig. 1).

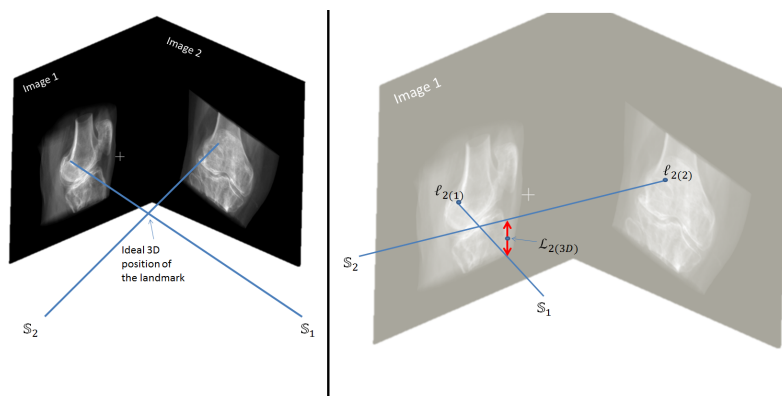


Fig. 1 Definition of the landmark pose in the 3D space. In the right part of the figure the position of the real landmark is shown ($\mathcal{L}_{i(3D)}$)

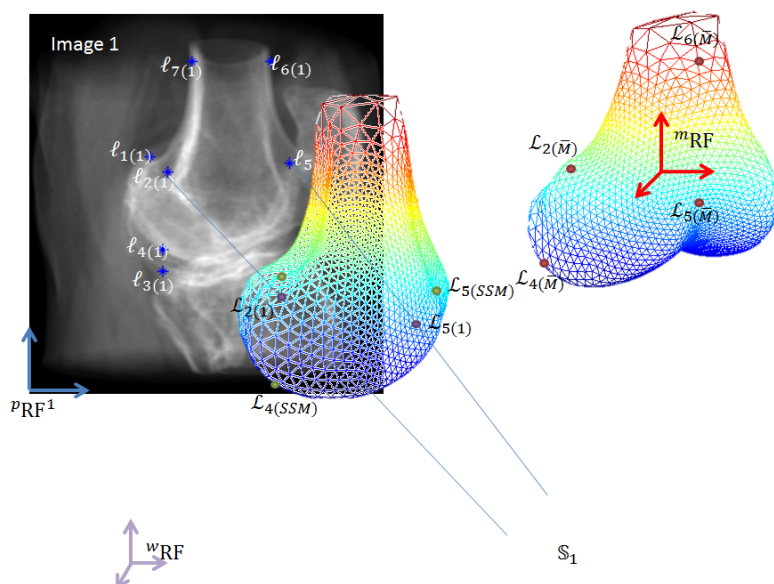


Fig. 2 Schematic of the reference frames involved in this algorithm. Points on the image plane, expressed in pixel index (i, j) , are in the PRF^n , where n indicates the number of the image plane. mRF is the reference frame of the SSM, and wRF is the world reference frame. Points in red on the model are the landmarks $\mathcal{L}_{i(\bar{M})}$, points in blue on the image are the selected landmarks on the first image $l_{i(n)}$, points in purple are the landmarks in the wRF extracted from source-image rays ($\mathcal{L}_{i(n)}$) and points in green are the model landmarks in the wRF after the initialization ($\mathcal{L}_{i(SSM)}$).

As the landmarks are defined both in the 3D space and on the model, we can use a simple registration algorithm for corresponding points to find the initial pose of the model in the 3D space [24]. $\mathcal{L}_{i(\bar{\mathbf{M}})}$ is the landmark on the model surface, the similarity transformation ${}^w\hat{\mathbf{T}}_m^0$ (rotation \mathbf{R} , translation \mathbf{t} and scaling factor s) transforms the model landmarks into $\mathcal{L}_{i(SSM)}$ (Fig. 2). To obtain the transformation ${}^w\hat{\mathbf{T}}_m^0$ we define a cost function F based on the Euclidean distance between the transformed landmarks and the position of the points defined by the backprojection of the landmarks on the images $\mathcal{L}_{i(3D)}$.

$$F = \sum_{i=1}^7 \|\mathcal{L}_{i(3D)} - \mathcal{L}_{i(SSM)}\| \quad (3)$$

$$\mathcal{L}_{i(SSM)} = s \cdot \mathbf{R} \cdot \mathcal{L}_{i(\bar{\mathbf{M}})} + \mathbf{t}$$

We thus apply the same transformation matrix ${}^w\hat{\mathbf{T}}_m^0$ to all the points of the mean model ($\bar{\mathbf{M}}$).

$$\mathbf{M}_{(SSM)} = {}^w\hat{\mathbf{T}}_m^0 \cdot \bar{\mathbf{M}} \quad (4)$$

2.3 Femur tracking

After the manual initialization of the pose in the first frames, the algorithm returns the pose of the shape as homogeneous matrix ${}^w\mathbf{T}_m^t$. To track the pose of the femur in the ${}^w\text{RF}$ during time, we repeat the steps described in 2.3.1 for each sample time $t = 0, \dots, T$. In each frame the initialization is given by the homogeneous matrix of the previous frame (${}^w\mathbf{T}_m^{t-1}$).

The shape of the femur is optimized at $t = 0$, as described in 2.3.2, and then kept constant for all the tracking times.

2.3.1 Pose optimization

Using the initialization of the SSM's pose in the ${}^w\text{RF}$, for each given image plane ${}^i\text{RF}^n$ we can extract the silhouette of the model, identifying the points that share a "contour" edge. This edge is in common between two facets that have the normals pointing in opposite directions from the corresponding source $\mathbb{S}_{(n)}$ (see Fig. 3). We define the silhouette of the model as $\mathbf{S}_{j_n(n)}$, with $j_n = 1, \dots, J_n; n = 1, 2$, where each point is shared between two "contour" edges. We then project the silhouette extracted on the corresponding image plane, defining the corresponding set of pixels on each image plane as $\mathbf{s}_{j_n(n)}$.

For each contour pixel in $\mathbf{c}_{k_n(n)}$ we find the nearest pixel of $\mathbf{s}_{j_n(n)}$, and associate it to the corresponding silhouette point in $\mathbf{S}_{j_n(n)}$. We call $\mathbf{S}_{k_n(n)}$ the associated silhouette point.

For each point on the image plane we define the projection line from the source ($\mathbb{S}_{(n)}$) and find $\mathbf{C}_{k_n(n)}$, the closest point on the line between the line and the corresponding silhouette point $\mathbf{S}_{k_n(n)}$, as shown in Fig. 4.

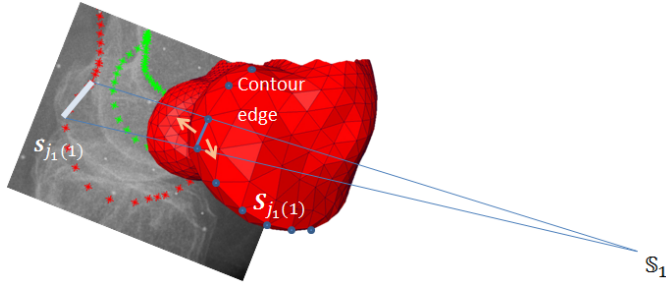


Fig. 3 Definition of the contour edge describing the silhouette of the model. The red points are the projection of the silhouette points. The green points are the extracted contour.

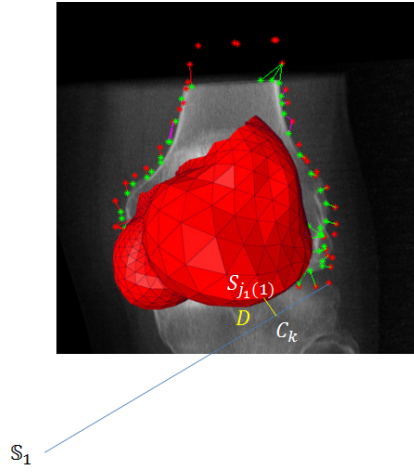


Fig. 4 Definition of the objective function: minimization of the distance D

The cost function F_{pose} to minimize can thus be defined as:

$$\begin{aligned}
 F_{pose} &= \min_{{}^w\mathbf{T}_m^t} \left(\sum_{k_n=1}^{K_n} D_{k_n} \right) \\
 D_{k_n} &= \| \mathbf{C}_{k_n(n)} - \mathbf{S}_{k_n(n)} \| \\
 {}^w\mathbf{T}_m &= \begin{bmatrix} {}^w\mathbf{R}_m & {}^w\mathbf{t}_m \\ 0 & \frac{1}{s} \end{bmatrix}
 \end{aligned} \tag{5}$$

where ${}^w\mathbf{T}_m^t(s, {}^w\mathbf{R}_m, {}^w\mathbf{t}_m)$ indicates the transformation matrix from the ${}^m\text{RF}$ in the ${}^w\text{RF}$, composed of rotation, translation and scaling factor, $\|\dots\|$ is the Euclidean distance, $\mathbf{C}_{k_n(n)}$ and $\mathbf{S}_{k_n(n)}$ are respectively the 3D point on the contour line and the 3D point of the silhouette of the model.

A GA optimization process is performed independently on each frame, to avoid local minima [25]. The population is composed of the parameters that define the pose of the femur in the world reference frame, i.e. Euler angles,

translation vector and scaling. The initial population (40 samples) of each minimization is extracted from uniform parameter distributions defined as small deviations ($\pm 10^\circ$, $\pm 10\text{mm}$, $\pm 0.1\text{scale factor}$) from an initial pose, i.e. the initialization pose for the first frame or the result of the previous minimization otherwise. The maximum number of iterations was equal to 100.

2.3.2 Shape optimization

After the first optimization of the pose, we performed a shape optimization, based on the algorithm described in [16]. The shape optimization process implemented a closed form solution for the best approximation of the original shape with a deformable model. The optimization is inserted in a minimization procedure F_{shape} , based only on the $O - 1 = 23$ weights of the SSM (α_o).

$$F_{shape} = \min_{\alpha_o} \left(\sum_{k_n=1}^{K_n} D_{k_n} \right) \quad (6)$$

$$D_{k_n} = \|\mathbf{C}_{k_n(n)} - \mathbf{S}_{k_n(n)}\|$$

2.4 Validation

2.4.1 Datasets

In order to validate our tracking algorithm we used the datasets of three different patients, which have respectively severe, mild and moderate osteoarthritis of the knee. For each patient, we have the CT scan and we generate the DRRs.

The CT dataset was composed of DICOM images acquired with a SIEMENS Sensation 64 CT machine. Each slice is 512x512 pixel (0.3516 mm/pixel) with a slice thickness of 0.6 mm and a spacing between slices of 0.4 mm. In order to assess the model reconstruction performances we segmented the bone shape using Amira[®] (VSG—FEI, France).

For each patient, a virtual environment around the CT was created, in order to simulate fluoroscopic scans from three different point of views. 11 angles of flexion were simulated for a total of 11 x 3 patients x 3 views = 99 images. The center of the CT dataset was taken as the ^wRF. The setup of the acquisition was virtually created in order to have the ground truth for the pose of the femur. The center of the CT dataset was taken as the ^wRF. The DRR is built integrating the density of each voxel of the CT along the direction of each ray as in [26]. Three different sources and image planes were simulated for each patient, resulting in three sets for each patient: the first image shows a lateral view of the femur (L_0), the second and third images show a view rotated on the sagittal plane of 10 (L_{10}) and 90 degrees (F) (see Fig. 5). To represent the knee flexion-extension, we rotated the images T times ($t = 0, \dots, T$, with $T = 10$). The pose of the CT shape is fixed and the source and the image plane rotate in order to simulate a rotation of the femur.

To build the images on the lateral plane, we rotated the image plane of 8° clockwise. The same rotation is made on the frontal plane, rotating both the image and the source, to get a consistent motion of the femur in the 3D space (see Fig. 6).

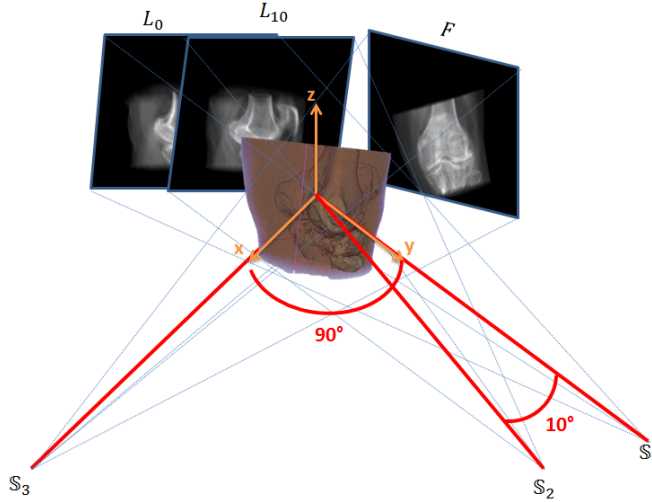


Fig. 5 Description of the virtual environment setup for the DRR creation. The CT model reference frame was used as world reference frame (orange reference frame). The source of L_0 was established on the y -axis, as well as the center of L_0 . L_{10} is obtained rotating the source-plane axis of 10 degrees on the x - y plane. F is obtained rotating the source-plane axis of 90 degrees, having it correspondent to the x -axis [27]

2.4.2 Experimental protocol

To reduce the computational effort of the minimization function, we used 3D models with a reduced number of points (1:8). In order to evaluate the performances of the shape reconstruction and the tracking separately, we used the CT-based shape and the SSM shape. An open source software called Cloud Compare (www.danielgm.net/cc) was used to assess the differences between the SSM and the CT shape. The software returns both the point to shape distances for the meshes (calculated as the distance between points and facets) and a graphical view of them. Referring to Fig. 5, we call L_0 the lateral view (Image 1 in the figure), L_{10} the projection at 10 degrees (Image 2 in the figure) and F the frontal view (Image 3 in the figure). The tracking algorithm was then computed on the following images sets:

- $CT(L_0 - L_{10})$: Tracking is done using the CT shape and images L_0 and L_{10} .
- $CT(L_0 - F)$: Tracking is done using the CT shape and images L_0 and F .

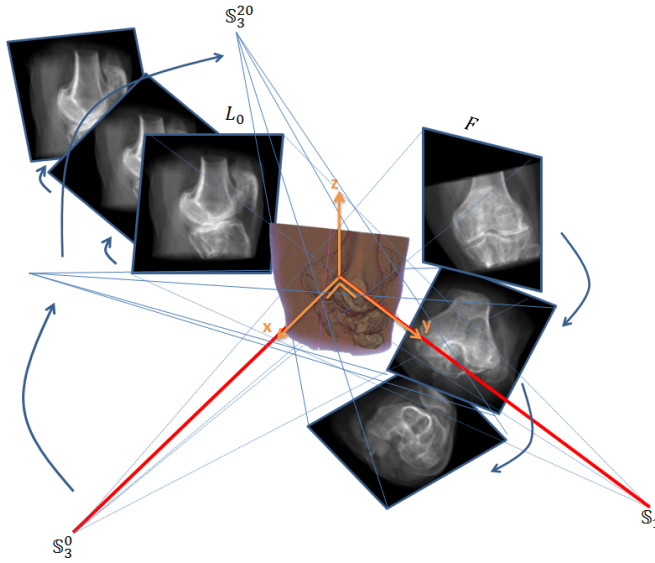


Fig. 6 Description of the virtual environment setup for the DRR creation. To represent the knee flexion-extension, we rotated the images T times of 8 degrees each ($T = 10$). \mathbf{S}_3^0 indicates the position of the source of image 3 at time 0, \mathbf{S}_3^5 indicates the position of the source of image 3 at time 5 and \mathbf{S}_3^{10} indicates the position of the source at time 10, i.e. after 80 degrees of rotation from time 0

- $SSM(L_0 - L_{10})$: Tracking is done using the SSM shape and images L_0 and L_{10} .
- $SSM(L_0 - F)$: Tracking is done using the SSM shape and images L_0 and F .

These four setups are tracked for each one of the three models from which the DRR were generated. Kruskal-Wallis test $p < 0.05$ was performed to verify the significative difference between the median of each test.

3 Results

We tested the capability of the algorithm to deform the SSM in order to better approximate the CT shape and the ability to reconstruct the pose of the femur with two images per sample time.

In most cases, the distance between the SSM and CT was less than 3mm (Fig. 7 and Table 1)

For each CT shape we identified the anatomical axes (as defined in [28]) and used these axes to report the errors.

As in [29], we defined the **precision** as the standard deviation of the re-meaning error after removing the mean error of all frames for that specific sequence. Difference in coordinate system definition is thus minimized, as it is mainly formed by mean error. The three sequences (one for each patient) are

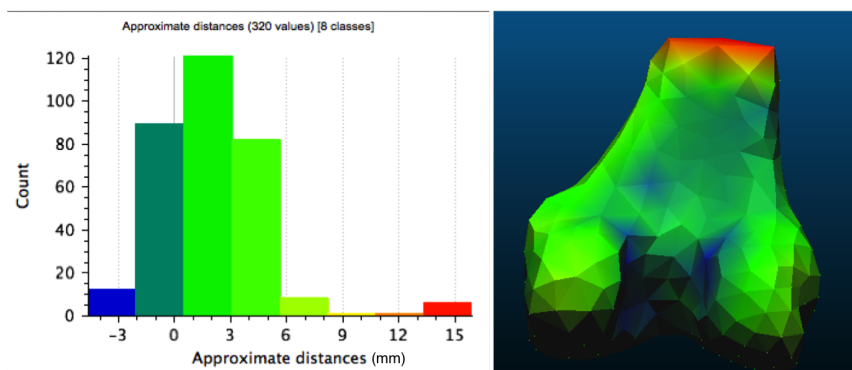


Fig. 7 Similarity between the original CT mesh and the reconstructed SSM mesh

Table 1 Mesh comparison statistical parameters: distances from point to surface represented as mean and standard deviation

	DRR 1	DRR 2	DRR 3
$SSM(L_0 - L_{10})$	$0.516 \pm 2.438mm$	$-0.607 \pm 5.104mm$	$2.188 \pm 2.933mm$
$SSM(L_0 - F)$	$0.545 \pm 2.391mm$	$-0.418 \pm 4.955mm$	$2.066 \pm 2.835mm$

than considered as a single population. In Fig. 8 the boxplots of the angular precision for each anatomical axes are shown, while in Fig. 9 the translation precision on each axis are shown. Kruskal-Wallis test was performed to verify the significant difference between the median of each test.

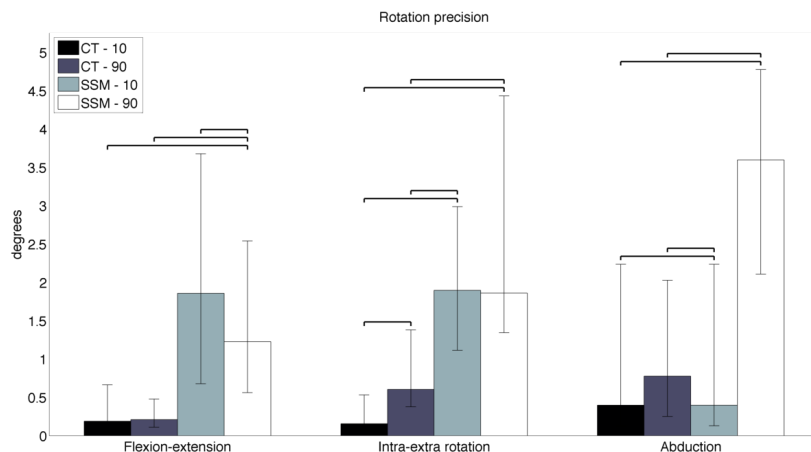


Fig. 8 Rotation precision with respect to the ground truth (known as artificially generated). The height of the box indicates the median value and whiskers extend from the 25th to the 75th percentile. Square parenthesis above two boxes indicates that the two populations are statistically different.

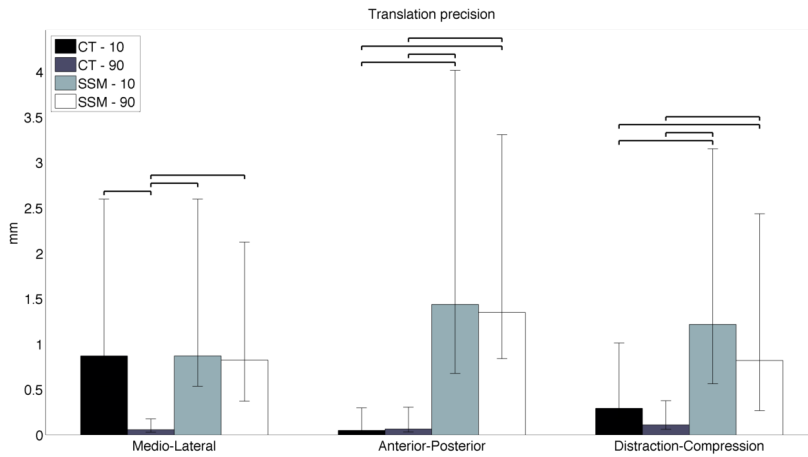


Fig. 9 Translation precision with respect to the ground truth (known as artificially generated).

4 Discussion

In this paper, we present a femoral kinematics reconstruction technique based on biplane fluoroscopic images taken during knee flexion-extension movements using SSMs.

The tracking capabilities of the proposed method were separately evaluated using the exact model, semi-automatically reconstructed from the CT dataset, and the implemented SSM [18].

Our approach shows the applicability of a semi-automatic algorithm for 2D/3D registration using SSM. With respect to dynamic MRI, our method is less expensive and more accurate [4]. Compared to actual clinical procedures, that are based on a previous static CT or MRI scan and a manual fitting of the shape, this method is less expensive (as CT and MRI can be avoided) and more accurate due to the automatic fitting. The ionizing radiation given to the patient in our procedure is just a few hundredths of a mSv (considering a total of 22 images and an average dose of 0.001mSv for each image, we have a total dose of 0.02mSv, much lower than a 2.2mSv average dose absorbed with a CT) [30, 31]. Compared to skin mounted markers, our system does not suffer from relative movements between the skin and the bone, and is thus suitable for accurate joint kinematics studies [6, 7, 8]. Our method proved to be stable without relying on implanted markers or joint implants, that will return more accurate results at the price of high invasiveness [5, 9, 11].

In order to have the ground truth pose of the model, we artificially simulated the fluoroscopic projections using DRR technique. Our SSM was derived using 24 MRI datasets of healthy patients. Such a number of samples is quite limited compared to other dataset in the literature e.g. [2] used 43 CT images

of the knee and [16] used 30 CT-based models of the hip. When we evaluated the exact model reconstruction ability of our SSM and the implemented deformation technique we found residual errors on the order of 2.2 mm as median Euclidean distance in the worst case. Such error is due to the different levels of cut of the femoral diaphysis and to the limited number of femurs used as SSM. Also, the subject used to obtain the DRRs were osteoarthritic, thus with evident deformity of the femur's shape. The result obtained is in line with the ones by [2], where they state a point to surface distance of 2 mm and maximum errors of 3 mm. Having a larger dataset, including greater variations, would allow increasing the morphing possibilities, reducing reconstruction errors.

In the case of the CT shape, all the median translation errors are lower than 1 mm, in line with the results presented by [2]. We used two different angles for the DRR definition, and the results showed that the higher translation error in medio-lateral precision at 10° is due to a lower capability to resolve the depth information. The residual error at 90° can be associated to rounding error and to a limited number of generations for each GA running.

Using the SSM shape the results are statistically different from the one obtained with the CT shape. The translation median errors are lower than 1.5 mm, probably due to the low deformation possibilities given by the reduced number of shapes of the training set, whereas the rotation error spans up to 4° in almost all directions. In [29] the authors state errors up to 3° but they used a SSM including a longer femoral diaphysis, thus the informative content of the statistical model is bigger.

The advances of the proposed new methodology with respect to the current state of the art resides in the fact that our method has been applied to pathological patients resulting in accuracy comparable to the current state of the art. Moreover, we managed to achieve limited decrease in performance with a lower angle (up to 10°) between the fluoroscopic projections. In this way, the range of movement for the flexion-extension of the knee is enlarged, extending the possibilities for the tracking with different movements.

One of the limits of our study is the reduced testing dataset for tracking, nevertheless since we used 11 image projections, progressively extending the knee joint, the dataset is enough to assess the statistical power of the analysis. Fluoroscopic images could be much noisier than the virtually reconstructed projections, due to fast acquisition of the images that causes blurring, calibration plates and white noise due to instrumentation. Further analysis will be directed towards adding realistic noise to the images and testing the tracking algorithm performances. Acquired images as performed in [27] will also be used to test the algorithm.

In conclusion, we showed the clinical applicability of our method for femoral tracking using a biplane fluoroscopy and based on SSM, thus reducing costs and lowering the patient's radiation dose.

5 Conflict of interest

Marta Valenti, Elena De Momi, Weimin Yu, Giancarlo Ferrigno, Mohsen Akbari Shandiz, Carolyn Anglin and Guoyan Zheng declare that they have no conflict of interest.

References

1. N Baka, BL Kaptein, and M de Bruijne. 2d/3d shape reconstruction of the distal femur from stereo x-ray imaging using statistical shape models. *Medical Image Analysis*, 15(6):840–50, December 2011.
2. N Baka and M de Bruijne. Statistical shape model-based femur kinematics from biplane fluoroscopy. *IEEE transactions on Medical Imaging*, 31(8):1573–1583, 2012.
3. M Akbari Shandiz. *Component Placement in Hip and Knee Replacement Surgery: Device Development, Imaging and Biomechanics*. PhD thesis, Biomedical Engineering, 2015.
4. CE Draper, JM Santos, LC Kourtis, TF Besier, M Fredericson, GS Beaupre, GE Gold, and SL Delp. Feasibility of using real-time mri to measure joint kinematics in 1.5 t and open-bore 0.5 t systems. *Journal of Magnetic Resonance Imaging*, 28(1):158–166, 2008.
5. S Tashman and W Anderst. In-vivo measurement of dynamic joint motion using high speed biplane radiography and ct: Application to canine acl deficiency. *Journal of Biomechanical Engineering*, 125(2):238, 2003.
6. JP Holden, JA Orsini, KL Siegel, and TM Kepple. Surface movement errors in shank kinematics and knee kinetics during gait. *Gait & Posture*, 5:217–227, 1997.
7. A Leardini, L Chiari, U Della Croce, and A Cappozzo. Human movement analysis using stereophotogrammetry. part 3. soft tissue artifact assessment and compensation. *Gait & Posture*, 21(2):212–25, February 2005.
8. A Cappozzo, F Catani, U Della Croce, and A Leardini. Position and orientation in space of bones during movement: anatomical frame definition and determination. *Clinical Biomechanics (Bristol, Avon)*, 10(4):171–178, June 1995.
9. S. Tashman. Abnormal rotational knee motion during running after anterior cruciate ligament reconstruction. *American Journal of Sports Medicine*, 32(4):975–983, April 2004.
10. E De Momi, E Beretta, and G Ferrigno. Hip joint centre localisation with an unscented kalman filter. *Computer Methods in Biomechanics and Biomedical Engineering*, 16(12):1319–29, January 2013.
11. TSY Tang, NJ MacIntyre, and HS Gill. Accuracy of a fluoroscopy technique for assessing patellar tracking. *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2003*, 2003.

12. D Dennis, MR Mahfouz, RD Komistek, and W Hoff. In vivo determination of normal and anterior cruciate ligament-deficient knee kinematics. *Journal of Biomechanics*, 38(2):241–53, February 2005.
13. S Acker, R Li, H Murray, PS John, S Banks, S Mu, U Wyss, and K Deluzio. Accuracy of single-plane fluoroscopy in determining relative position and orientation of total knee replacement components. *Journal of Biomechanics*, 44(4):784–787, 2011.
14. M Fleute and S Lavallée. Nonrigid 3-d/2-d registration of images using statistical models. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI99*, pages 138–147. Springer, 1999.
15. J Canny. A computational approach to edge detection. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 679–698, 1986.
16. G Zheng, X Dong, KT Rajamani, X Zhang, M Styner, RU Thoranaghatte, LP Nolte, and MG Ballester. Accurate and robust reconstruction of a surface model of the proximal femur from sparse-point data and a dense-point distribution model for surgical navigation. *IEEE transactions on Bio-Medical Engineering*, 54(12):2109–22, December 2007.
17. TF Cootes, DH Cooper, CJ Taylor, and J Graham. Trainable method of parametric shape description. *Image and Vision Computing*, 10(5):289–294, 1992.
18. M Valenti, C Chen, E De Momi, G Ferrigno, and G Zheng. 3d shape landmark correspondence by minimum description length and local linear regularization. In *XIII Mediterranean Conference on Medical and Biological Engineering and Computing 2013*, pages 1837–1840. Springer, 2014.
19. WE Lorensen and HE Cline. Marching cubes: A high resolution 3d surface construction algorithm. In *ACM Siggraph Computer Graphics*, volume 21, pages 163–169. ACM, 1987.
20. TF Cootes, CJ Taylor, DH Cooper, and J Graham. Active shape models—their training and application. *Computer Vision and Image Understanding*, 61(1):38–59, 1995.
21. TF Cootes, GJ Edwards, and CJ Taylor. Active appearance models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(6):681–685, 2001.
22. TF Cootes, MC Ionita, C Lindner, and P Sauer. Robust and accurate shape model fitting using random forest regression voting. In *Computer Vision—ECCV 2012*, pages 278–291. Springer, 2012.
23. C Lindner, S Thiagarajah, J Wilkinson, The Consortium, G Wallis, and Timothy F Cootes. Fully automatic segmentation of the proximal femur using random forest regression voting. *IEEE Transactions on Medical Imaging*, 32(8):1462–1472, 2013.
24. CW Spoor and FE Veldpaus. Rigid body motion calculated from spatial co-ordinates of markers. *Journal of Biomechanics*, 13(4):391–393, 1980.
25. Lawrence Davis et al. *Handbook of genetic algorithms*, volume 115. Van Nostrand Reinhold New York, 1991.
26. CT Metz. Digitally reconstructed radiographs. *Utrecht: Utrecht University*, page 79, 2005.

27. GB Sharma, SK Saevarsson, S Amiri, S Montgomery, H Ramm, DD Lichti, R Lieck, S Zachow, and C Anglin. Radiological method for measuring patellofemoral tracking and tibiofemoral kinematics before and after total knee replacement. *Bone and Joint Research*, 1(10):263–271, 2012.
28. DL Miranda, MJ Rainbow, EL Leventhal, JJ Crisco, and BC Fleming. Automatic determination of anatomical coordinate systems for three-dimensional bone models of the isolated human knee. *Journal of Biomechanics*, 43(8):1623–1626, 2010.
29. N Baka, BL Kaptein, JE Giphart, M Staring, M de Bruijne, BPF Lelieveldt, and E Valstar. Evaluation of automated statistical shape model based knee kinematics from biplane fluoroscopy. *Journal of Biomechanics*, 47(1):122–129, 2014.
30. Radiologyinfo.org. <http://www.radiologyinfo.org>. Accessed: 2015-06-05.
31. J Henckel, R Richards, K Lozhkin, S Harris, FM Rodriguez y Baena, ARW Barrett, and JP Cobb. Very low-dose computed tomography for planning and outcome measurement in knee replacement the imperial knee protocol. *Journal of Bone & Joint Surgery, British Volume*, 88(11):1513–1518, 2006.