# Robust Patella Motion Tracking using Intensity-based 2D-3D Registration on Dynamic Bi-plane Fluoroscopy: Towards Quantitative Assessment in MPFL reconstruction surgery

Yoshito Otake[1], Matthieu Esnault[1], Robert Grupp[2], Shinichi Kosugi[3], Yoshinobu Sato[1]

[1] Graduate School of Information Science, Nara Institute of Science and Technology, Japan
[2] Department of Computer Science, Johns Hopkins University, USA
[3] Department of Orthopedic Surgery, Nara Prefecture Western Medical Center

## ABSTRACT

The determination of in vivo motion of multiple-bones using dynamic fluoroscopic images and computed tomography (CT) is useful for post-operative assessment of orthopaedic surgeries such as medial patellofemoral ligament reconstruction. We propose a robust method to measure the 3D motion of multiple rigid objects with high accuracy using a series of bi-plane fluoroscopic images and a multi-resolution, intensity-based, 2D-3D registration. A Covariance Matrix Adaptation Evolution Strategy (CMA-ES) optimizer was used with a gradient correlation similarity metric. Four approaches to register three rigid objects (femur, tibia-fibula and patella) were implemented: 1) an individual bone approach registering one bone at a time, each with optimization of a six degrees of freedom (6DOF) parameter, 2) a sequential approach registering one bone at a time but using the previous bone results as the background in DRR generation, 3) a simultaneous approach registering all the bones together (18DOF) and 4) a combination of the sequential and the simultaneous approaches. These approaches were compared in experiments using simulated images generated from the CT of a healthy volunteer and measured fluoroscopic images. Over the 120 simulated frames of motion, the simultaneous approach showed improved registration accuracy compared to the individual approach: with less than 0.68mm root-mean-square error (RMSE) for translation and less than 1.12° RMSE for rotation. A robustness evaluation was conducted with 45 trials of a randomly perturbed initialization showed that the sequential approach improved robustness significantly (74% success rate) compared to the individual bone approach (34% success) for patella registration (femur and tibia-fibula registration had a 100% success rate with each approach).

**Keywords**: Patella motion tracking, 2D-3D Registration, Dynamic bi-plane fluoroscopy, MPFL reconstruction.

## 1. INTRODUCTION

An *in vivo* measurement of knee kinematics is important to better understand the motion of the patella, as rupture of the medial patellofemoral ligament (MPFL), may cause recurrent patellar dislocation. MPFL reconstruction with an autologous tendon graft has often been performed clinically to repair patellar dislocation[1]. However, it has been difficult to evaluate the joint congruency during knee motion after reduction of the patella with MPFL reconstruction[1]. This study aims to provide a robust 2D-3D intensity-based registration method to extract the motion of the patella, femur, tibia and fibula using bi-plane fluoroscopic images and CT images by comparing several methods for multiple bone registration. Prior work has focused on low-resolution MRI[2], which does not provide a sufficient framerate to capture fluid motion of the knee, and computer assisted manual registration[3], which is limited by human inability to detect sub-millimeter alignments of complex anatomy reliably. Recent work has focused on GPU accelerated registration with a novel optimization hierarchy[4], however the coupling of two local optimization algorithms may not provide as much resistance to local optima as a search strategy designed for global optima[5]. We therefore propose to extend an existing 2D-3D registration framework which uses a robust, global, optimization component with various strategies for registering multiple bones.

## 2. METHODS

### 2.1 Proposed method

Bi-plane fluoroscopic images of a healthy knee were acquired using two C-arms forming a 90° angle. Ground truth patient pose was not available. Using Fujifilm Synapse Vincent, each bone was manually segmented in the CT image. Separate volumes for each bone were created and used, along with the fluoroscopic images, as input to the 2D-3D registration process.

Medical Imaging 2016: Image-Guided Procedures, Robotic Interventions, and Modeling,
edited by Robert J. Webster, III, Ziv R. Yaniv, Proc. of SPIE Vol. 9786, 97860B
 · doi: 10.1117/12.2214699

For each bi-plane fluoroscopic frame pair, the 2D-3D registration process was used to determine the transformation which best aligned the fluoroscopic images and the CT bone volumes. The approach described by Otake et al.[6][7] is used as the foundation of the 2D-3D registration in this work. Fig. 1 illustrates an overview of the registration method. The method was modified to improve robustness and accuracy when tracking multiple rigid objects. The rotation of each bone is represented by three Euler angles about the standard image axes of each bone volume. The translation of each bone is represented by translation along the standard axes of each bone volume. The tibia and fibula were restricted to have identical rigid-body motions for this study, yielding three bone volumes and objects to register: femur, tibia-fibula, and patella.

For the first frame, the 2D-3D registration was initialized with a rough, manual, estimate of the transformation of each bone. (The error in this manual initialization was simulated with random perturbation in the simulation study as described below.) Subsequent frames used the previous frame's registration result as an initial guess and also used a smaller search range than the first frame's search range.

The Covariance Matrix Adaptation Evolution Strategy (CMA-ES) optimization algorithm[5] was utilized. The objective function computes Digitally Reconstructed Radiographs (DRRs) of the bone volumes previously extracted from CT. Gradient correlation (GC) was computed between each DRR and the true fluoroscopic frames and the sum of GC for the two views defined the output of the objective function. The GPU implementation of the objective function allowed for a large population size and enhanced the robust, global search, properties of CMA-ES. Two resolution levels were used; a coarse resolution and the full resolution.

Four registration strategies were compared: Individual bone registration, sequential bone registration, simultaneous bone registration, and the combination of sequential and simultaneous registrations. Individual bone registration consisted of three separate registrations, one for each bone. Sequential registration also consisted of three separate registrations, however the final registration of each bone was retained, kept fixed, and used in each DRR when registering subsequent bones. The following order was used: femur, tibia-fibula, and patella. The simultaneous registration performs a registration of the three bones simultaneously. The combination of sequential and simultaneous registrations first performed the sequential registration, and used the result as an initial guess for a final simultaneous registration. Table 1 outlines the approaches.

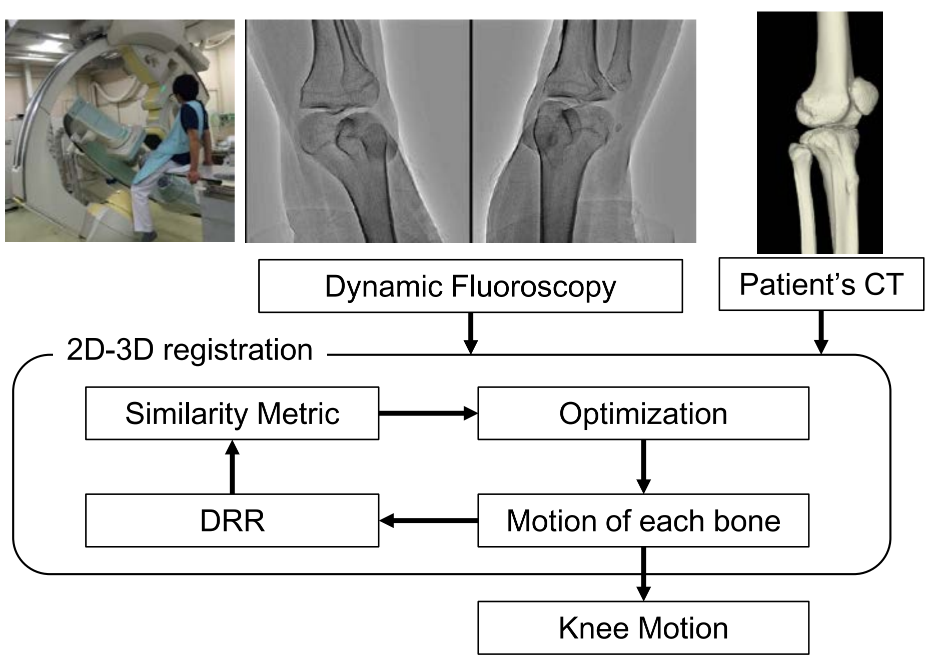


**Figure 1.** Overview of the proposed method

**Table 1.** Outline of the multi-object registration approaches that were evaluated in this paper ($I_{fix}$: fixed image, $P(\cdot)$: projection operator, $T_k$: transformation of the $k^{th}$ bone, $V_k$: volume of the $k^{th}$ bone, $GC(a,b)$: gradient correlation between image $a$ and $b$)

(a) Individual bone method

for each bone k=1:N

compute $\hat{T}_k = \underset{T_k \in \mathbb{R}^6}{\operatorname{argmax}} GC\left(I_{fix}, P(V_k, T_k)\right)$

end

(b) Sequential method

for each bone k=1:N

$I_{bg}$ ←projection of 1 to k-1th bone

compute $\hat{T}_k = \underset{T_k \in \mathbb{R}^6}{\operatorname{argmax}} GC\left(I_{fix}, I_{bg} + P(V_k, T_k)\right)$

end

(c) Simultaneous method

$$\{\hat{T}_1, \hat{T}_2, \dots, \hat{T}_N\} = \underset{T_1, T_2, \dots, T_N \in \mathbb{R}^6}{\operatorname{argmax}} GC\left(I_{fix}, P(V_1, \dots, V_N, T_1, \dots, T_N)\right)$$

(d) Sequential + Simultaneous method

Perform simultaneous method with initialization by sequential method

### 2.2. Experiments with Simulated Images

To evaluate each registration method, fluoroscopic images were simulated using DRRs from a CT of a healthy knee. Fig 2. shows a synthetic frame. A registration was performed on a subset of the healthy knee fluoroscopic dataset. An idealized fluid knee motion was generated by smoothly interpolating the measured motion, resulting in one flexion-extension cycle of 120 frames.

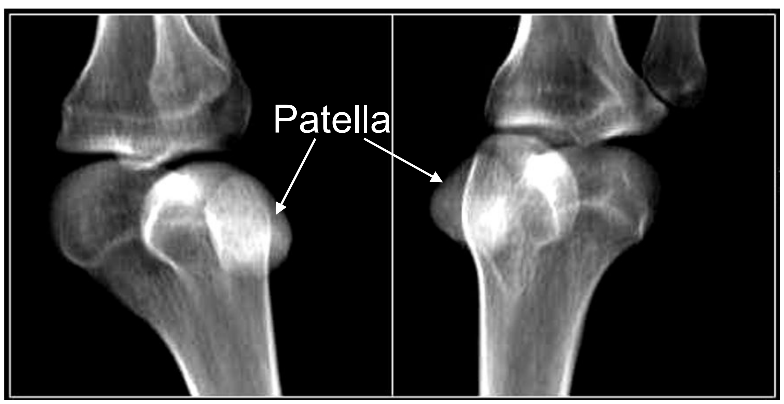


**Figure 2.** An example simulated image used in the experiment

These transformations served as ground truth data for evaluating the individual bone and simultaneous methods. For each trial, the initial guess was perturbed from ground truth to simulate error in the manual initialization as mentioned above. In our current method, the first frame registration error depends on the quality of the initial guess; with registration errors typically propagating to subsequent frames. Thus, robustness was evaluated on only the first frame. To take into account variation of the first frame, three frames from the simulated motion were tested: the minimum of flexion (frame 1 in the simulated data), the maximum of flexion (frame 60 in the simulated data) and an intermediate frame (frame 30 in the simulated data). Fifteen randomized perturbations along each axis in the range of -20 mm to +20 mm for translation and -8° to +8° for rotation were added to the ground truth transformation to create the initial guess, resulting in 45 registrations for each method. Each registration method was executed with the same optimization parameters (i.e., search range, convergence criteria and other parameters used in CMA-ES optimization).

### 2.3. Experiments with Measured Fluoroscopic Images

Feasibility of the proposed approach was tested using measured fluoroscopy images (shown in Figure 1). A “ground truth” registration for each bone was estimated using the simultaneous approach starting from an initial transformation carefully set to the visually interpreted solution (i.e., the “best guess” that we could achieve with our current tools). Fifty registration trials for all four registration approaches were conducted with an initialization randomly perturbed from the ground truth (±30 mm translation, ±10 degrees rotation); error from ground truth was computed.

Since we did not perform calibration to obtain relative pose between the two fluoroscopic views, separate single view registrations of the femur using each view were performed to estimate the relative poses of each C-Arm prior to the registration evaluation..

### 2.4. Visualization tool

To allow for analysis of knee motion trajectories obtained from each 2D-3D registration, a module for 3D Slicer (Brigham and Women’s Hospital, Boston, MA) was implemented. The module visualizes a series of transformations applied to the CT bone models. Fig.3 shows a screenshot with knee kinematics loaded. The module provides a user-friendly interface to load and visualize kinematic data in a 3D scene. An anatomical coordinate system may be defined for each bone. The relative motion of bones may also be analyzed quantitatively with 2D plots as shown in Fig.3.

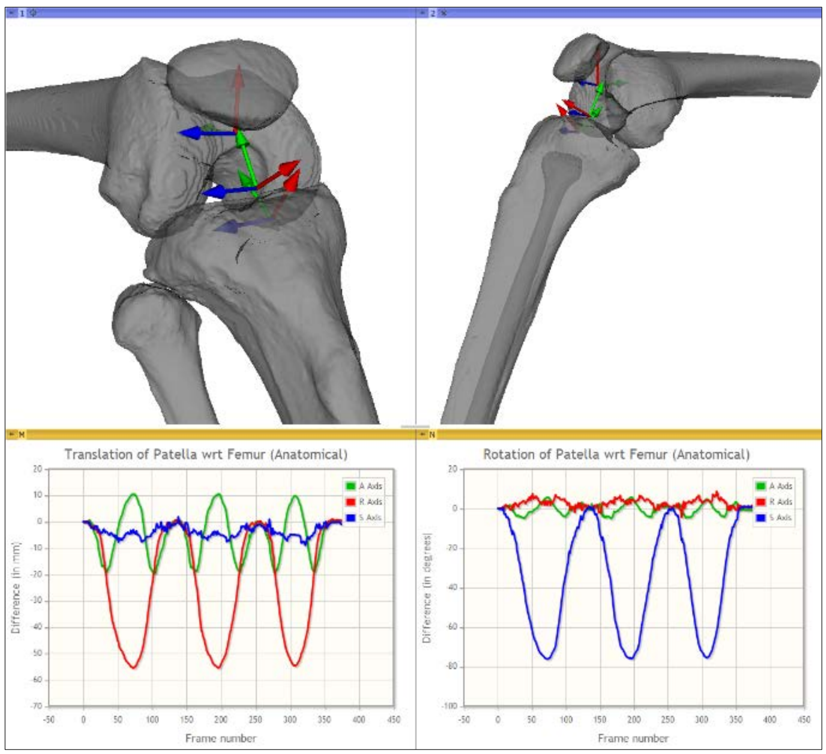


**Figure 3.** The 3D Slicer module implemented for the visualization of multi-bone kinematics

## 3. RESULTS

### 3.1. Experiments with Simulated Images

We describe the results of the experiments with simulated images in this section. For the femur and tibia-fibula, registration error did not show a significant difference for all approaches, however, for the patella which is smaller than the other bones and contrast of its edges often reduced by overlapped other bones, the simultaneous approach demonstrated higher accuracy than the individual bone approach, as seen in Fig 4. The simultaneous approach reduced the maximum translation RMSE from 1.37mm to 0.68mm and the maximum rotation RMSE from 2.44° to 1.12°. In terms of robustness, the sequential approach was more robust than the other approaches, as seen in Fig 5. The simultaneous method generally performed better than the single bone registration for small perturbations but became less robust for larger perturbations.

The combination of the sequential and the simultaneous approach ranked between the simultaneous and the sequential approach.

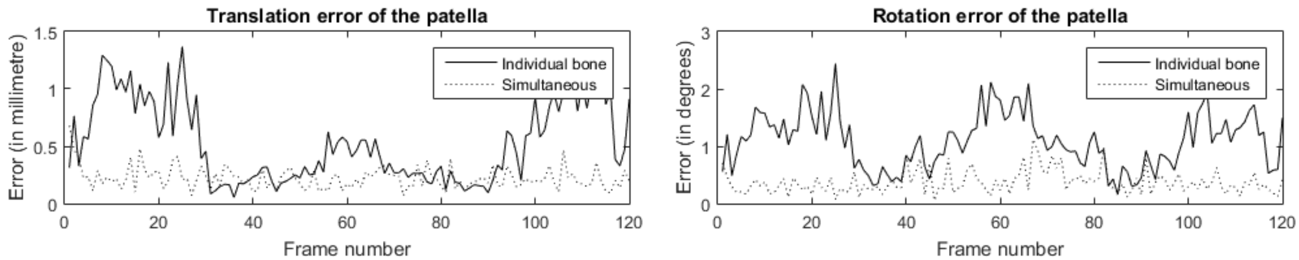


**Figure 4.** RMSE for each frame in the simulation study. Individual and simultaneous registration methods are shown.

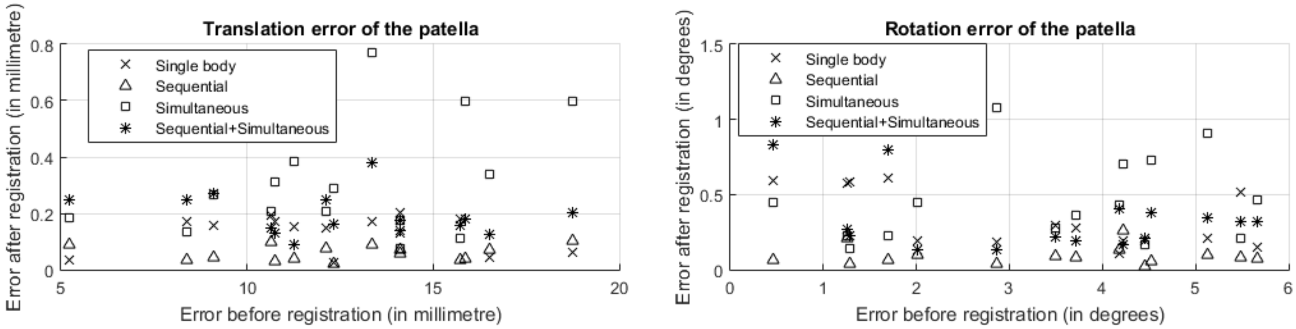


**Figure 5.** First frame registration method comparison with 5 perturbations for each of the 3 selected frames. Failed registrations do not appear due to the suppression of outliers in this plot.

A threshold of 2 mm for translation error and 2° for rotation error was used to define a successful registration. The success rate of the femur and the tibia-fibula was 100% in all methods. However, the individual bone method for the patella was 87%, compared to 100% with the other three methods.

Execution speed and the number of objective function evaluations for registration of the first fluoroscopic frame were recorded for each registration method. The tests were conducted using a workstation with an Intel Xeon X5690 (3.47GHz) CPU and 4 GPUs (Tesla C2075). The mean execution times of the sequential, individual, combined sequential and simultaneous, and simultaneous methods to register all three bones were 133.7 seconds, 143.7 seconds, 162.5 seconds, and 305.4 seconds, respectively. The mean number of objective function evaluations of the sequential, individual, combined sequential and simultaneous, and simultaneous methods to register all three bones were ~1.1 million, ~1.1 million, ~1.1 million, and ~1.7 million function evaluations, respectively. An analysis was not performed for registration of frames after the first, as the registration was much faster due to the reduced search range.

### 3.2. Experiments with Measured Fluoroscopic Images

With measured fluoroscopic images, all femur and tibia-fibula registrations were successful. For patella registration, the individual bone, sequential, simultaneous, and combined sequential and simultaneous methods exhibited 38%, 74%, 70%, and 68% success rates, respectively. The individual method demonstrated significantly poorer performance compared to the other three methods, which showed similar success rates

## 4. DISCUSSION

A robust, intensity-based, 2D-3D registration for tracking of multiple rigid bodies was proposed and demonstrated in an application of tracking of the knee kinematics. Four approaches to register a group of bones were proposed and compared using a simulation study and an experiment with measured images. According to these results, the sequential registration method showed the best performance in the case of the patella tracking. One limitation in the study with measured images was a lack of "true" ground truth. As is always the case with research of registration on real data, we used the best possible guess as our ground truth in this study. A visualization tool for kinematics was developed as a 3D-Slicer module, and was used to simulate realistic movements of the knee joint.

In order to further improve the robustness and accuracy of the registration, several improvements could be developed. The definition of an anatomical coordinate system would allow for a more realistic search space [8] for the pose of each bone, and additionally serve as an implicit constraint to help avoid local maxima and accelerate convergence to the global optimum. Additionally, knowledge of the relative position between the two C-arms should improve two-view registration results for measured data. Registration of several consecutive frames simultaneously and also optimizing for the relative camera position could improve the registration when it is not possible to track the C-arm poses.

An additional simulation study to analyze the effect of the fluoroscopic field of view and C-arm pose on registration performance could assist a clinician to determine the appropriate bi-plane fluoroscopic acquisition parameters in order to maximize the likelihood of a successful registration.

Advancement of this technology has the potential to provide a reliable solution for clinicians to perform more accurate and consistent post-operative assessments of various MPFL reconstruction outcomes. Furthermore, it is feasible that the proposed methods be used for a computer-assisted MPFL reconstruction system providing real-time tracking of the patella with respect to the femur and tibia, and bio-mechanical quality indicators of a fixation. Intraoperative tracking of the patella via fluoroscopy is particularly lucrative as it does not require the surgeon to apply screws to affix an optical tracking fiducial onto the small patellar structure.

## ACKNOWLEDGEMENT

This research was supported in part by JST PRESTO and MEXT/JSPS Grant-in-Aid for Scientific Research on Innovative Areas (Multidisciplinary Computational Anatomy) KAKENHI No. 26108004.

## REFERENCES

[1] T. Neri, R. Philippot, O. Carnesecchi, B. Boyer, and F. Farizon, “Medial patellofemoral ligament reconstruction: Clinical and radiographic results in a series of 90 cases,” Orthopaedics & Traumatology: Surgery & Research 101(1), pp. 65-69, 2015.

[2] R. Fellows, N. Hill, H. Gill, N. MacIntyre, M. Harrison, R. Ellis, and D. Wilson, “Magnetic resonance imaging for in vivo assessment of three-dimensional patellar tracking,” Journal of biomechanics 38(8), pp. 1643-1652, 2005.

[3] K. W. Nha, R. Papannagari, T. J. Gill, S. K. Van de Velde, A. A. Freiberg, H. E. Rubash, and G. Li, “In vivo patellar tracking: clinical motions and patellofemoral indices,” Journal of Orthopaedic Research 26(8), pp. 1067-1074, 2008.

[4] T. Ohnishi, M. Suzuki, T. Kobayashi, S. Naomoto, T. Sukegawa, A. Nawata, and H. Haneishi, “Robust 2d/3d registration for fast-flexion motion of the knee joint using hybrid optimization,” Radiological physics and technology 6(1), pp. 170-179, 2013.

[5] N. Hansen, S. D. Müller, and P. Koumoutsakos, “Reducing the time complexity of the derandomized evolution strategy with covariance matrix adaptation (cma-es),” Evolutionary Computation 11(1), pp. 1-18, 2003.

[6] Otake, Y., M. Armand, R. S. Armiger, M. D. Kutzer, E. Basafa, P. Kazanzides and R. H. Taylor (2012). "Intraoperative image-based multiview 2D/3D registration for image-guided orthopaedic surgery: incorporation of fiducial-based C-arm tracking and GPU-acceleration." IEEE Trans Med Imaging 31(4): 948-962.

[7] Otake, Y., A. S. Wang, A. Uneri, G. Kleinszig, S. Vogt, N. Aygun, S.-f. L. Lo, J.-P. Wolinsky, Z. L. Gokaslan and J. H. Siewerdsen (2015). "3D–2D registration in mobile radiographs: algorithm development and preliminary clinical evaluation." Physics in Medicine and Biology 60(5): 2075.

[8] A. E. Kedgley, E. J. McWalter, and D. R. Wilson, “The effect of coordinate system variation on in vivo patellofemoral kinematic measures,” The Knee 22(2), pp. 88-94, 2015.