3D Bone Surface Image Reconstruction from Freehand Ultrasound using Visual-Inertial Odometry and Deep Inertial Odometry

Russell Buchanan
Institute of Perception, Action and Behaviour
University of Edinburgh
Edinburgh, UK
russell.buchanan@ed.ac.uk

S Jack Tu
Botnar Institute for Musculoskeletal Sciences
University of Oxford
Oxford, UK
jack.tu@ndorms.ox.ac.uk

Marco Camurri
Faculty of Engineering
Free University of Bozen-Bolzano
Bolzano, Italy
marco.camurri@unibz.it

Stephen J Mellon
Botnar Institute for Musculoskeletal Sciences
University of Oxford
Oxford, UK
stephen.mellon@ndorms.ox.ac.uk

Maurice Fallon
Oxford Robotics Institute
University of Oxford
Oxford, UK
mfallon@robots.ox.ac.uk

***Abstract*—One in four people can be affected by patellofemoral joint (PFJ) pain or instability, and one in five patients has chronic knee pain despite treatment for these problems. Patellar mal-tracking after arthroplasty is often suspected to cause poor outcomes and/or pain. However, conventional static imaging techniques like CT and MRI have limitations such as cost, restricted field of view, and metal artefacts, which make it difficult to observe joints dynamically and provide clinicians with an improved understanding of PFJ interactions to improve patient management and outcomes. Combining 2D ultrasound with motion tracking for 3D reconstruction of the joint using semantic segmentation and position registration can be a solution. However, the need for expensive external infrastructure to estimate the trajectories of the scanner remains the main limitation to implementing 3D bone reconstruction from handheld ultrasound scanning clinically. In this paper, we assess two markerless methods - Visual-Inertial Odometry (VIO) and deep learning-based inertial-only odometry - for tracking a handheld ultrasound scanner. We present a low-cost method for 3D reconstruction of bone. We compare these methods against a motion-capture baseline, and the resulting 3D reconstructions demonstrate that using VIO or deep inertial odometry can reconstruct the PFJ for visualisation and further measurements from handheld ultrasound scans. The visual-inertial method performs as well as the motion capture method, with average reconstruction errors of $1.25\,\mathrm{mm}$ and $1.21\,\mathrm{mm}$, respectively. The VIO method is the first infrastructure-free method for 3D reconstruction of bone from handheld ultrasound scanning and has the potential to be used in a wide range of clinical applications, such as patella tracking.**

***Index Terms*—3D freehand ultrasound, Deep Learning, Motion estimation, Inertial measurement unit**

This work was supported by the EPSRC ORCA Robotics Hub (EP/R026173/1), the EU H2020 Project THING (Grant ID 780883), a Royal Society University Research Fellowship (Fallon) and Versus Arthritis MedTech POC Grant (Mellon & Tu)

I. Introduction

Approximately 20% of patients experience chronic pain after undergoing knee replacement surgery, despite the procedure having generally positive outcomes [1]. This is a significant issue that affects between 10-33% of Total Knee Replacement (TKR) patients, as studies have shown [2], [3]. Pain after TKR is often caused by abnormal patellofemoral joint kinematics, as many studies have reported [4], [5]. Unfortunately, there is no reliable and accurate method that can be used to routinely assess knee kinematics, particularly after joint replacement.

However, ultrasound has been shown to have enormous potential for bone visualization and is a low-cost and radiation-free alternative to MRI and X-ray imaging (Nazarian2008). Consequently, there has been significant interest in using ultrasound for 3D reconstruction of bone [6]–[9]. If this tool can be deployed to routinely assess knee kinematics after joint replacement, particularly those related to the patellofemoral joint, it could lead to improved diagnosis and potentially more reliable treatments for dissatisfied patients. It is possible to reconstruct 3D images from a series of 2D ultrasound images with tracked scanner position and orientation. However, current approaches require external infrastructure to facilitate device tracking, such as a motion capture system [6], [8], electromagnetic markers [9], or a robot arm for scanning [7]. Unfortunately, these methods come with significant infrastructure requirements that increase costs and limit the usefulness of the tool. As a result, clinical implementation is hindered due to the significant barriers posed by such requirements

"State estimation" is a well-established research area with techniques for tracking the position and orientation of robots without external infrastructure in the field of robotics, [10]. Visual Odometry (VO) is one of the research areas within this

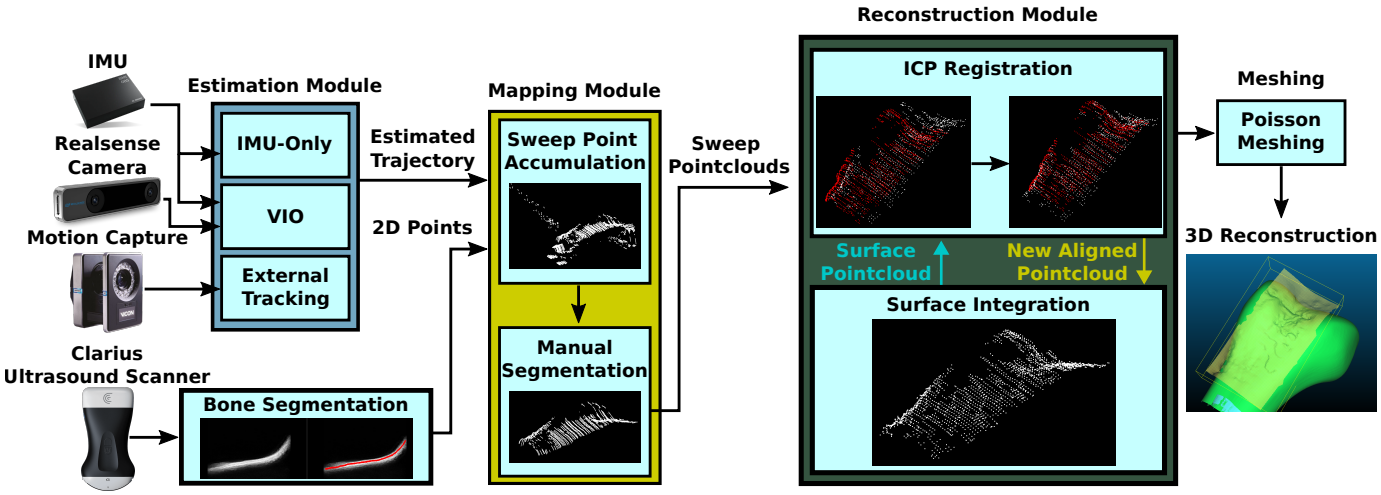

Fig. 1. Flow diagram of complete reconstruction pipeline. The estimation module computes the scanning trajectory using either IMU-only, VIO or motion capture external tracking. Separately, 2D ultrasound points are segmented for the presence of bone. These points are projected in 3D using the estimated trajectory and accumulated into separate point clouds for each sweep over the bone. After manually segmenting out the surface, the point clouds are aligned using ICP and integrated into the surface estimate. Finally, once all the sweeps have been integrated, the final surface is sampled and converted to a mesh using the Poisson method.

field, which involves tracking visual features in the environment using a camera to estimate motion [11]. This approach is often combined with Inertial Measurement Units (IMUs), which are small, low-cost sensors capable of measuring acceleration and rotation rates. When these two technologies are combined, the resulting system is called Visual-Inertial Odometry (VIO), and it has superior robustness to visual challenges, such as darkness or motion blur.

Recently, there have been several works applying deep learning to inertial data to enable low-drift, IMU-only odometry [12]–[14]. These methods train neural networks to learn motion models of, for example, pedestrians or robots. The network can then predict displacement or velocity from inertial data. The purpose of this study was to utilize robotic state estimation techniques to track a handheld ultrasound scanner, with the objective of exploring the feasibility of infrastructure-free methods for generating a 3D reconstruction of the bone surface from ultrasound imaging. Additionally, we aimed to develop a motion model of the scanner by training a neural network to learn from only IMU sensing. Finally, we compared the effectiveness of these two methods against a baseline established by external motion capture tracking of the scanner.

II. Methods

We evaluated three different methods for tracking the ultrasound scanner: motion capture, visual-inertial odometry (VIO), and deep learning IMU-only odometry. An OptiTrack system was used to capture motion of the ultrasound scanner with reflective marker. For VIO, we used a factor graph-based approach [14] to combine data from a stereo camera and an IMU mounted on the scanner. To implement the IMU-only method, we trained a neural network to predict position increments based on IMU data and combined it with classical IMU integration, which was similar to previous work [15].

A. Experimental Setup

The experimental setup used is shown in Fig.2. An Intel Realsense T265 stereo camera, which includes a Bosch BMI085 IMU, was mounted to a handheld ultrasound scanner (L7-38, Clarius Mobile Health Corporation, Burnaby, BC, Canada) with a 3D-printed mount. We used two phantom models of the same distal femur, with the patella in two different positions. The first model had the patella in the position of knee flexion, while the second had the patella in the position of knee extension. Our objective is to reconstruct these two models to visualize the different positions of the patella. To ensure that the bone models remained stationary during scanning, we bolted them to a metal plate. The phantoms were encased the models in a gelatine medium to enable ultrasound scanning. Infrared markers were placed on the bone and the scanner to enable motion capturing. All the devices were connected to a single PC running a Robotic Operation System to record all the sensor data and ensure that all the timestamps were recorded from a common clock.

B. Data Set Collection and Neural Network Training

A total of 150 scans were taken, out of which 130 involved the user slowly panning across the phantom while repeating for different starting positions and scanning directions. The scanning path was short and contained either 2 or 6 sweeps. The remaining 20 scans were longer and involved the user moving laterally while panning to cover the entire phantom. In all scans, the user picked up the scanner from a fixed cradle and returned it to the same position at the end.

The data was divided into two sets, 80% for training and 20% for validation. A single model was trained using all data for IMU-only estimation. The Adam optimizer was used with a learning rate of 10^{-6} over 2000 epochs and the model minimized validation error. The training process took about 6 hours on an NVIDIA GeForce RTX 3080 Ti GPU. Two

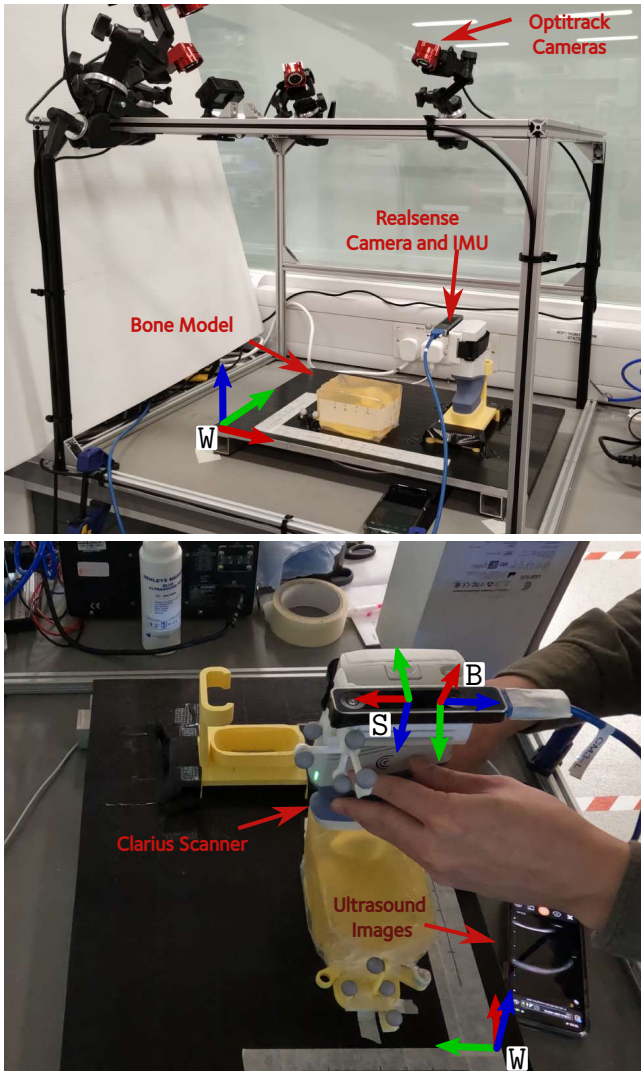

Fig. 2. Experimental setup for collecting ultrasound dataset. A Realsense stereo camera is mounted on the Clarius ultrasound scanner and a mount holds the scanner in place. The human operator picks up the scanner and passes it over the bone model several times before returning it to the mount.

scan types were used for both phantoms, labelled as Scan Type A for cross-sections (x direction) and Scan Type B for longitudinal(y direction) sections.

C. Tracking method

Both the VIO and IMU-only methods were applied to estimate the Maximum-A-Posteriori (MAP) state trajectory of a handheld ultrasound scanner in batch. The state estimator was implemented using the factor-graph library GTSAM [16] and built on the work from [15] while the neural network was the same as in [14] and developed using the PyTorch library.

The stereo camera was sampled at 30 Hz with one in three frames selected as a keyframe. Therefore, in the VIO method, nodes are added to the graph at 10 Hz, with IMU integration factors linking each consecutive pair.

IMU data was recorded at 200 Hz. For the neural network, the data samples were buffered in windows of 1 s ($N = 200$).

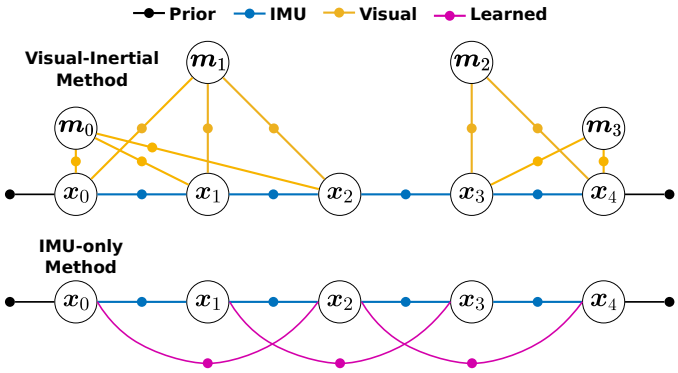

Fig. 3. Two example factor graph frameworks for five states $j = 0, 1, 2, 3, 4$. On top is the VIO method which uses prior, visual and IMU factors. On the bottom is the learned inertial odometry method which only uses IMU and learned factors. The four types of factors are: prior factors constraining the start and end pose to be equivalent, IMU integration factor from [17], visual landmark factors from [15] and our learned displacement factor from [14].

Neural network inference was done at a rate of 30 Hz, which we empirically found to give the best performance. Therefore, for each learned displacement factor, $k = j - 30$, meaning each factor linked two states 1 s apart. Fig 3 shows the learned displacement factor connecting every second node. In practice, because $k = j - 30$, the factors connect every 30th node, i.e., node 0 to 30, 1 to 31 etc. Therefore, the first 1 s of IMU data is needed before the neural network can start producing outputs.

In this work, we assumed each trajectory has a "loop closure," meaning the scanner trajectory ends with the same position and orientation as at the start. We also assume zero velocity at the start and end (as shown in Fig.3). Two prior factors are used, one for the first pose and one for the last pose. We use standard IMU pre-integration theory to connect consecutive nodes in the graph, as in [17]. This relates to the pose, velocity, and biases between two consecutive states.

External cameras such as OptiTrack can be used for motion estimation, with an expected error of less than 1 mm when calibrated correctly. However, occlusions can cause "jumps" in the estimated motion, leading to errors in position estimation. To mitigate this issue, we apply a median filter with a window size of 9 to the camera data, which is recorded at a rate of 200 Hz. The filtered trajectory is used as a reference point for comparing other estimation methods.

D. Reconstruction

The reconstruction was adopted from a previous work [8] with weighted projective fusion for point cloud integration and surface reconstruction. The final output mesh can be directly compared with ground truth meshes of bone phantom models.

1) Surface segmentation: The bone surface was segmented from ultrasound images using the method in [18]. The segmented pixels are then projected into 3D as points (as shown in Fig4). With the trajectory estimate from the tracking module, the projected points are accumulated into point clouds according to the "sweep" number over the bone. Finally, the reconstructed point clouds are reviewed and cleaned by

removing mislabelled points. Several point clouds of the bone surface corresponding to a separate sweep were carried on for further evaluation.

2) Alignment of Point Cloud Sweeps: For VIO and IMU-only methods there is incremental drift while estimating the scanner's pose. Knowing that we are scanning the same bone over several sweeps, we can correct this drift and align the sweep points using Iterative Closest Point (ICP) [19] with a open-source library *libpointmatcher* [20]. ICP were constrained to position and yaw (rotation about the gravity axis) only, considering the acceleration, pitch and roll estimate from the IMU were highly accurate over a short time period of a single scan. Although errors caused by tracking drift while moving across the bone cannot be corrected as the sweeps are treated as rigid point clouds, we assume that the drift incurred over this small distance is not enough to severely degrade the reconstruction.

While the Optitrack-estimated sensor pose is not subject to drift, aligning point clouds also slightly improves reconstruction for this configuration.

3) Surface Reconstruction and Meshing: The open source library Open3D [21] were used to perform Truncated Signed Distance Fields (TSDF) representation and integration algorithm [22] to fuse the aligned point clouds into smooth mesh surface. For TSDFs are a volumetric representation where voxel values correspond to the signed distance from the surface boundary, being positive outside the surface and negative when penetrating inside the surface. TSDF integration was computed with $2\,\mathrm{mm}$ resolution and $1\,\mathrm{cm}$ truncation with $1\,\mathrm{mm}$ voxel filtering and Point-to-Plane matching to improve robustness to outlier points.

Alignment and surface reconstruction were done in a loop whereby point clouds are first aligned, then integrated into the surface. A point cloud sampling of the TSDF surface is then extracted and used to align the next point cloud. Once all the sweeps have been aligned and integrated the final surface, then the point cloud is converted to a mesh using the Poisson method [23].

III. Results and Discussion

A. Tracking Results

TableI shows the level of estimation error of each tracking method compared to the Optitrack trajectory, which was considered the ground truth. Visual-inertial and deep learning inertial values are presented with and without the loop closure constraint from the start and end of prior factors.

Both the Absolute Pose Error (APE) and Relative Pose Error (RPE) translation components [24] were computed. APE gives the average error of each trajectory compared to the ground truth trajectory at any point in time. RPE, on the other hand, shows the rate of drift over a given distance. As shown in Table I, RPE was computed over $20\,\mathrm{mm}$ segments, indicating that the value represents the average drift for every $20\,\mathrm{mm}$ travelled.

Adding the loop closure slightly decreases the average position error (APE) in most cases except for Bone 2 Scan A, but it has no effect on the RPE. This is because the VIO method is already highly accurate and drifts an average of $1.8\,\mathrm{mm}$ per $20\,\mathrm{mm}$. In contrast, the IMU-only methods have larger errors than VIO. Even though adding the loop closure does reduce both RPE and APE for most trajectories, it remains higher than VIO.

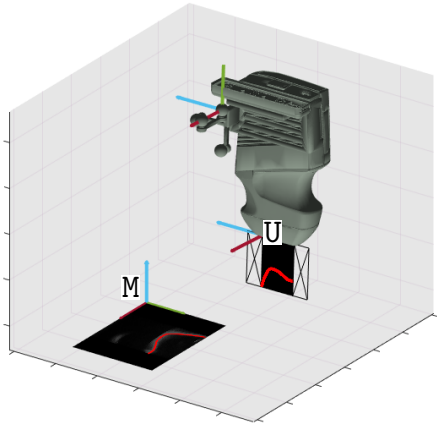

Fig. 4. Ultrasound images collected through the Clarius API are in a size of 640×480. The origin of the ultrasound image frame `M` was defined as pixel (0,0). A fixed, known transformation from `M` into the ultrasound scanner's frame `U` was applied to all segmented pixels to project them into 3D points. Finally, the 3D points in `U` were transformed into the global coordinate system using the selected odometry source.

TABLE I
Position Estimation Error [mm]

Method	**APE**				**RPE** (20 mm)			
Bone	Bone 1		Bone 2		Bone 1		Bone 2	
Scan Type	A	B	A	B	A	B	A	B
VIO	8.7	3.8	3.5	4.2	2.3	1.9	1.6	1.2
VIO - Loop	8.3	3.8	4.9	4.2	2.3	1.9	1.6	1.2
IMU-Only	53.2	124.4	112.2	18.9	10	14.7	14.6	9.7
IMU-Only - Loop	49.0	41.9	132.7	34.7	9.9	11.8	15.3	9.7

The trajectories for all six scans for Bone1 (Fig.5) and Bone2 (Fig.6) were plotted and showed a clear difference in the magnitude of error between VIO and IMU. The VIO methods were almost perfectly aligned with Optitrack, whereas the IMU-Only method had more significant errors. However, for five of the six trajectories, the method largely followed the Optitrack trajectory with a slight offset. It is evident that the network has learned the scanning motion and is providing useful information to the estimator. In the reconstruction step, we corrected these errors using ICP.

B. Reconstruction Results

Table II shows the mean and worst distance errors of each method when compared to the ground truth mesh model of the phantom bone. The motion capture method and the visual-inertial method had average errors of $1.21\,\mathrm{mm}$ and $1.25\,\mathrm{mm}$ respectively, with only a 3% difference between them. However, the IMU-only method had an average error of $1.85\,\mathrm{mm}$, which was 35% higher than the motion capture method. This is because the IMU-only method has higher drift rates

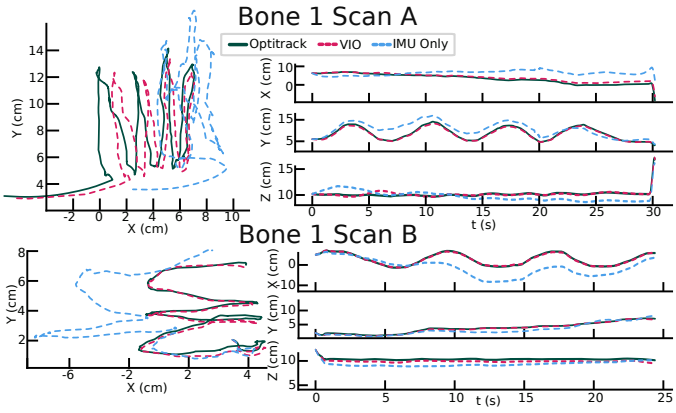

Fig. 5. Trajectories for the different algorithms while scanning Bone 1. Visual-inertial and learned inertial methods both use the loop closure constraint. In these plots, the trajectory is cropped to start when the scanner is in contact with the bone and cropped to end when scanning is finished. The first pose is aligned with ground truth in position and yaw, not roll or pitch.

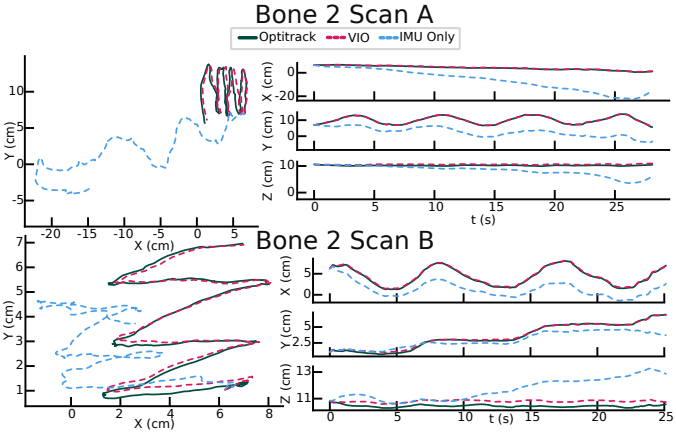

Fig. 6. Trajectories for the different algorithms while scanning Bone 2. Visual-inertial and learned inertial methods both use the loop closure constraint. Alignment is the same as for Fig. 5.

Fig.7 and Fig.8 present the reconstructed models of Bone 1 and Bone 2 for each scanning type. The colourised mesh represents the distance error from the ground truth model, and the surface reconstruction is overlaid on top of the ground truth model for each bone. It is important to note that the mesh reconstructions may contain some errors, such as a flaring out of the mesh at the edges, which is observable in Bone 1 Scan A with OptiTrack (Fig.7).

In this study, we demonstrated a slightly higher error rate compared to other published works, without requiring expensive external infrastructure. Reference [9] used EM trackers to track the scanner's movement on the skin and reported an average error of $1.05\,\mathrm{mm}$ while scanning the distal femur and proximal tibia of 15 cadaveric specimens. Reference [7] used a robotic arm to perform the scanning, and they reported an average error of $0.85\,\mathrm{mm}$ across two distal femurs (one human model and one bovine) using their best results, with an average worst error of $9.06\,\mathrm{mm}$.

Other reconstruction methods can cover a wider area of the bone, but they may introduce additional problems. For example, using a statistical shape model like in [9] can lead to biased reconstructions. Similarly, submerging the sensor in water as in [7] is not a practical solution. In contrast, our method involves using a general-purpose scanner held by a human and does not require an SSM. This approach makes our method more practical and versatile.

C. Future Work

In this work, we have demonstrated that tracking the pose of an ultrasound probe with outward-facing visual-inertial odometry is sufficiently accurate such that, when paired with ICP-based reconstruction, it can be used for markerless, hand-held ultrasound reconstruction. Future work on developing this method should focus on the hardware, possibly integrating the camera with the ultrasound scanner.

The learning-based IMU tracking method shows promise, but currently, it has a high drift rate, which makes it impractical. However, using more prior information may help to make the IMU-only method useful. Since the drift rate is quite high, it may be better to accumulate point clouds over shorter distances, rather than multiple sweeps. This could make the ICP-based matching more challenging, but it could be mitigated by using an SSM. While incorporating an SSM into the pipeline could introduce a bias into the reconstruction, it may be attractive if the goal is to estimate the position of the joint, assuming no abnormalities in the shape of the bone.

We additionally investigated if the amount of available data significantly affected the results. We tested models trained with half as much data and found that RPE only increased by 2.3%. Using a quarter data increases error by 15%. Therefore, it is unlikely that more training data is needed to improve our IMU-only method. Instead, possibly new neural network architectures should be explored.

IV. Conclusion

In conclusion, this paper presents a novel approach for markerless 3D reconstruction of bone using a handheld ultrasound scanner. Our method combines recent robotic estimation techniques, including visual-inertial odometry and learning inertial motion tracking to estimate the pose of the ultrasound scanner. We have shown that when using VIO, it is possible to reconstruct the bone with errors comparable to methods using external tracking infrastructure. This tool can then be used to visualize the position of the patella with different degrees of flexion. We also investigated the usability of a learning-based inertial tracking approach based on deep learning. We have shown the method has promise and that a neural network is able to learn scanning motions; however, drift is likely too high to be immediately useful. Instead, we believe with the addition of SSM, an IMU-only method for 3D reconstruction could be possible.

References

[1] V. Wylde, A. Beswick, J. Bruce, A. Blom, N. Howells, and R. Gooberman-Hill, "Chronic pain after total knee replacement," *EFORT Open Reviews*, vol. 3, no. 8, pp. 461–470, 2018.

TABLE II
MEAN/WORST RECONSTRUCTION ERROR [MM]

Method	**Bone 1**		**Bone 2**	
Sweep Type	A	B	A	B
Optitrack	0.99 / 7.71	1.00 / 7.50	1.53 / 9.95	1.76 / 5.11
VIO	0.85 / 5.2	1.10 / 3.93	1.70 / 6.43	1.80 / 6.32
IMU-Only	1.8 / 7.4	1.51 / 7.67	2.56 / 10.03	1.98 / 7.93

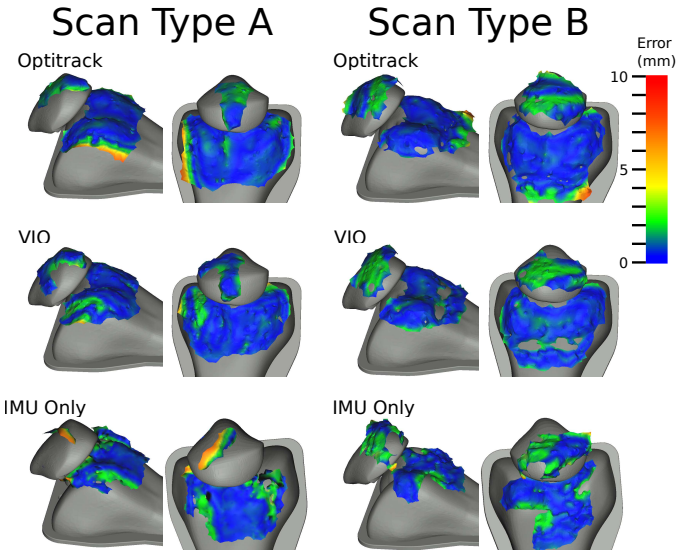

Fig. 7. Reconstructions of Bone 1 which is in 40° flexion. We show results for each estimation methods and for both scanning types A and B.

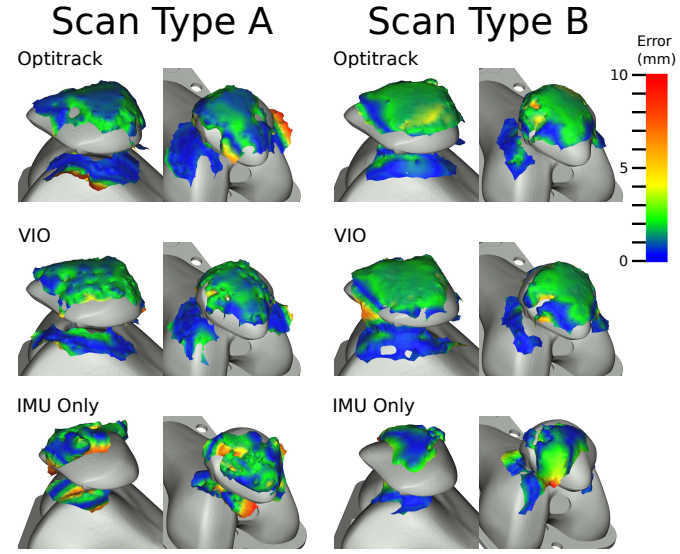

Fig. 8. Reconstructions of Bone 2 which is in full extension. The patella is now directly over the distal femur. We show results for each estimation methods and for both scanning types A and B.

[2] K. K. Petersen, O. Simonsen, M. B. Laursen, T. A. Nielsen, S. Rasmussen, and L. Arendt-Nielsen, "Chronic postoperative pain after primary and revision total knee arthroplasty," *The Clinical journal of pain*, vol. 31, no. 1, pp. 1–6, 2015.

[3] P. J. van der Wees, J. J. Wammes, R. P. Akkermans, J. Koetsenruijter, G. P. Westert, A. van Kampen, G. Hannink, M. de Waal-Malefijt, and B. W. Schreurs, "Patient-reported health outcomes after total hip and knee surgery in a dutch university hospital setting: results of twenty years clinical registry," *BMC musculoskeletal disorders*, vol. 18, no. 1, pp. 1–10, 2017.

[4] C. Belvedere, A. Ensini, A. Leardini, V. Dedda, A. Feliciangeli, F. Cenni, A. Timoncini, P. Barbadoro, and S. Giannini, "Tibio-femoral and patello-femoral joint kinematics during navigated total knee arthroplasty with patellar resurfacing," *Knee Surgery, Sports Traumatology, Arthroscopy*, vol. 22, pp. 1719–1727, 2014.

[5] R. S. Pulavarti, V. V. Raut, and G. J. McLauchlan, "Patella denervation in primary total knee arthroplasty–a randomized controlled trial with 2 years of follow-up," *The Journal of Arthroplasty*, vol. 29, no. 5, pp. 977–981, 2014.

[6] E. Świątek-Najwer, K. Otto, P. Krowicki, K. Krysztoforski, P. Keppler, and J. Kozak, "3d bone shape modelling basing on dataset recorded by ultrasound free-hand navigated probe," in *Information Technologies in Biomedicine, Volume 4*, E. Piętka, J. Kawa, and W. Wieclawek, Eds. Cham: Springer International Publishing, 2014, pp. 45–56.

[7] W. Kerr, P. Rowe, and S. G. Pierce, "Accurate 3d reconstruction of bony surfaces using ultrasonic synthetic aperture techniques for robotic knee arthroplasty," *Computerized Medical Imaging and Graphics*, vol. 58, pp. 23–32, 2017.

[8] R. Jia, P. Monk, D. Murray, J. A. Noble, and S. Mellon, "CAT & MAUS: A novel system for true dynamic motion measurement of underlying bony structures with compensation for soft tissue movement," *J Biomech*, vol. 62, pp. 156–164, 2017.

[9] M. R. Mahfouz, E. E. Abdel Fatah, J. M. Johnson, and R. D. Komistek, "A novel approach to 3d bone creation in minutes," *The Bone & Joint Journal*, vol. 103-B, no. 6 Supple A, pp. 81–86, 2021.

[10] T. D. Barfoot, *State Estimation for Robotics*. Cambridge University Press, 2017.

[11] R. Mur-Artal and J. D. Tardós, "ORB-SLAM2: an open-source SLAM system for monocular, stereo and RGB-D cameras," *IEEE Transactions on Robotics*, vol. 33, no. 5, pp. 1255–1262, 2017.

[12] C. Chen, X. Lu, A. Markham, and N. Trigoni, "IONet: Learning to Cure the Curse of Drift in Inertial Odometry," *32nd AAAI Conference on Artificial Intelligence, AAAI 2018*, pp. 6468–6476, 2018.

[13] W. Liu, D. Caruso, E. Ilg, J. Dong, A. I. Mourikis, K. Daniilidis, V. Kumar, and J. Engel, "TLIO: Tight learned inertial odometry," *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 5653–5660, 2020.

[14] R. Buchanan, M. Camurri, F. Dellaert, and M. Fallon, "Learning inertial odometry for dynamic legged robot state estimation," in *Proceedings of the 5th Conference on Robot Learning*, vol. 164, 2021, pp. 1575–1584.

[15] D. Wisth, M. Camurri, and M. Fallon, "Vilens: Visual, inertial, lidar, and leg odometry for all-terrain legged robots," *IEEE Transactions on Robotics*, vol. 39, no. 1, pp. 309–326, 2022.

[16] F. Dellaert, "Factor Graphs and GTSAM: A Hands-on Introduction," Georgia Institute of Technology, Tech. Rep. September, 2012. [Online]. Available: https://smartech.gatech.edu/handle/1853/45226

[17] C. Forster, L. Carlone, F. Dellaert, and D. Scaramuzza, "On-Manifold Preintegration for Real-Time Visual-Inertial Odometry," *IEEE Transactions on Robotics*, vol. 33, no. 1, pp. 1–21, 2017.

[18] R. Jia, S. J. Mellon, S. Hansjee, A. P. Monk, D. W. Murray, and J. A. Noble, "Automatic bone segmentation in ultrasound images using local phase features and dynamic programming," in *IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, 2016, pp. 1005–1008.

[19] F. Pomerleau, F. Colas, and R. Siegwart, "A review of point cloud registration algorithms for mobile robotics," *Foundations and Trends in Robotics*, vol. 4, pp. 1–104, 05 2015.

[20] F. Pomerleau, F. Colas, R. Siegwart, and S. Magnenat, "Comparing ICP Variants on Real-World Data Sets," *Autonomous Robots*, vol. 34, no. 3, pp. 133–148, 2013.

[21] Q.-Y. Zhou, J. Park, and V. Koltun, "Open3d: A modern library for 3d data processing," 2018. [Online]. Available: https://arxiv.org/abs/1801.09847

[22] R. A. Newcombe, S. Izadi, O. Hilliges, D. Molyneaux, D. Kim, A. J. Davison, P. Kohi, J. Shotton, S. Hodges, and A. Fitzgibbon, "KinectFusion: Real-time dense surface mapping and tracking," in *10th*

IEEE International Symposium on Mixed and Augmented Reality, 2011, pp. 127–136.

[23] M. Kazhdan, M. Bolitho, and H. Hoppe, "Poisson Surface Reconstruction," in *Symposium on Geometry Processing*, A. Sheffer and K. Polthier, Eds. The Eurographics Association, 2006.

[24] R. Kümmerle, B. Steder, C. Dornhege, M. Ruhnke, G. Grisetti, C. Stachniss, and A. Kleiner, "On measuring the accuracy of slam algorithms," *Autonomous Robots*, vol. 27, no. 4, pp. 387–407, 2009.