2023 IEEE International Conference on Robotics and Automation (ICRA 2023)
May 29 - June 2, 2023. London, UK

Pose Quality Prediction for Vision Guided Robotic Shoulder Arthroplasty

Morgan Windsor [1], Jing Peng[1], Ashish Gupta[3], Peter Pivonka[2], and Michael J Milford[1]

***Abstract*— Surgical assistive robots offer the potential for drastically improved patient outcomes through more accurate, more repeatable surgical procedures like shoulder arthroplasty operations. Existing robotic systems typically rely on optical marker tracking and require invasive marker attachment for localization, complicating the surgical workflow and patient recovery. But moving towards a markerless system is very challenging, both because of the absolute difficulty and the large variation in localization conditions across thousands of surgical procedures. In this paper we propose an alternative approach: rather than try to create a "perfect" and fully generalizable markerless localization system, instead create a reliable and trustworthy localization system that is able to continually self-assess the likely quality of its localization estimates, and act accordingly. We propose a lightweight method for predicting vision-based pose estimation performance using internal pipeline artifacts (without needing external ground truth from a marker-based system). Using extensive real robot experiments with challenging actual imagery from surgery, we demonstrate our prediction system accurately self-characterizes the localization system's performance across a wide range of localization conditions, and demonstrate that this prediction system generalizes to a range of surgical conditions. We then show how online performance prediction can drive active robot navigation that minimizes localization error, reducing target pose estimation error by 96.1% for rotation and 96.7% for translation compared to rejected alternative trajectories.**

I. INTRODUCTION

A key factor in patient outcomes following shoulder arthroplasty is correct glenoid implant positioning [1]. Implant placement during surgery is a complex and challenging process and implants are often not positioned according to plan [2]. Robotic assistive systems have been shown to improve the accuracy of implant positioning in other surgical operations including hip, knee, and spine procedures [3]. However, existing commercial systems rely on optical markers for patient pose estimation, substantially impeding the surgical workflow [3], with additional problems caused by the chance of complications from invasive marker attachment to patient bones [4]. For this reason, markerless approaches are attractive, but face challenges of harsh lighting (Fig. 1), sparse bone features, and frequent bone occlusion. In this work we study the problem of patient-to-robot localization. In recognition of the extreme challenges, here we propose a new localization system targeting the glenoid that is able to predict its own performance, enabling both active navigation strategies to mitigate localization issues, and the possibility of handover to humans where necessary.

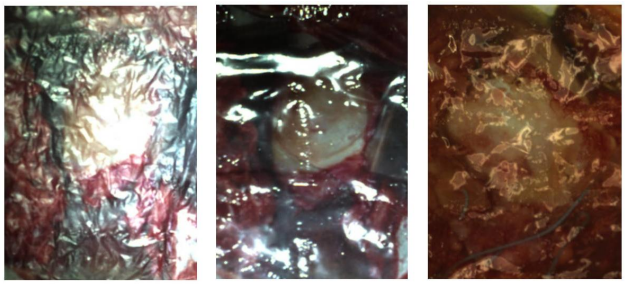

Fig. 1. Limited bone exposure and harsh lighting conditions, shown simulated here on actual surgical images captured during shoulder arthroplasty, present some of the major challenges for vision-based guidance systems in the surgical setting.

We make the following specific contributions:

1) A new lightweight method for predicting the quality of individual target pose estimates calculated using local feature matching, that leverages a rapid, consensus-based prediction.
2) A framework for active navigation of the robot tool-point towards the operating zone that continually assesses and re-assesses likely localization performance, and plans and changes navigation paths accordingly.
3) Extensive experiments under a range of challenging conditions involving variable lighting and target obscuration, that demonstrate both the predictive capabilities of the localization system self-assessment processes, and the ability of the robot to plan routes that minimize localization errors.

The remainder of the paper covers relevant previous research and the opportunity in this problem domain to substantially improve surgical outcomes and quality of life of patients. We present the various components of our approach, our experimental setup and results, and conclude with a discussion of the key findings and future work.

This research was partially supported by the QUT Centre for Robotics and funding from ARC Laureate Fellowship FL210100156 to M. Milford. The work of M. Windsor and J.Peng is supported by an ARC Industrial Transformation Training Centre (ITTC) for Joint Biomechanics grant IC190100020, M. Windsor is also supported by an Australian Government Research Training Program (RTP) Scholarship.

[1]M. Windsor, J. Peng, and M. Milford are with the QUT Centre for Robotics, School of Electrical Engineering and Robotics at the Queensland University of Technology. morgan.windsor@hdr.qut.edu.au, jing.peng@qut.edu.au, michael.milford@qut.edu.au

[2]P. Pivonka is with the QUT Centre for Biomedical Technologies, School of Mechanical, Medical, and Process Engineering at the Queensland University of Technology. peter.pivonka@qut.edu.au

[3]A. Gupta is with the Queensland Unit for Advanced Shoulder Research and Greenslopes Private Hospital, Brisbane, QLD, Australia. ashish@qoc.com.au

II. BACKGROUND

Here we present an overview of the shoulder arthroplasty procedure, outline existing Computer Assisted Orthopedic Surgery (CAOS) systems, 6-degree-of-freedom (6DOF) pose estimation techniques, and related work in localization system integrity.

A. Shoulder Arthroplasty

Shoulder arthroplasty is an orthopedic surgical procedure to replace both the ball and socket in the main joint in the shoulder, the glenohumeral joint, with prosthetic implants. The procedure is used to treat conditions such as arthritis, bone loss [5], and irreparable rotator cuff tears [6]. The volume of shoulder arthroplasty procedures performed each year is rapidly increasing; between 2011 and 2017 there was a 103.7% increase to 104,575 procedures in the United States alone [7]. Accurate positioning of the glenoid, or socket, component during shoulder arthroplasty is a key determinant of surgical outcomes, both in terms of implant longevity and post-operative joint function [1]. Key to this process is placement of the guide-wire, also called a K-wire, which is drilled into the glenoid and used to guide the reamers and drills that prepare the bone and define the final implant placement [8]. Achieving accurate placement of the guide-wire is challenging given limited bone exposure [9], large variability between patients [1], and the mobility of the glenoid during surgery. Surgeons often have difficulty reproducing planned implant placement using manual instrumentation techniques [2], [10] with up to 68.6% of implants defined as malpositioned, being greater than 4 millimeters or 10 degrees deviated from the surgical plan.

B. Computer and Robot Assisted Orthopedic Surgery

CAOS systems include planning software, Patient Specific Instruments (PSI), surgical navigation systems, and surgical robots [11]. Current orthopedic surgical robots such as the Mako (Stryker), ROSA (Zimmer-Biomet), and Navio (Smith & Nephew) systems are commercialized for hip, knee, and spine procedures [11], but are not available for use in shoulder arthroplasty. Current guidance systems typically rely on optical tracking markers, which slows down surgery due to manual registration processes [3]. Currently used markers require additional or larger incisions to expose bone for mounting that may introduce complications such as coracoid fracture [4]. Markerless systems for knee procedures are yet to be commercialized [12]–[16] and have not been applied to robot guidance.

C. Pose Estimation

While pose estimation in the context of surgery is a relatively narrow research area, 6DOF pose estimation has been widely studied in robotics and computer vision. Techniques have been developed for different sensor modalities such as RGB images, point clouds, and combinations of the two [17]. Pose estimation techniques vary considerably from 3D bounding box detectors [18]–[20], common in autonomous vehicle applications, to direct 6DOF pose estimators using approaches such as template matching [21], [22], deep-learning methods [23]–[30], and traditional geometric and algorithmic techniques [31]–[35].

While deep learning-based approaches have become dominant in many application areas, their need for extensive training data is especially problematic in the surgical domain given practical challenges of gathering extensive training data on a patient-by-patient basis. This limitation has begun to be addressed through methods to synthetically generate training data [16], [36].

Another key factor is the likelihood that these robots will continue to work collaboratively with surgeons rather than independently, placing a higher emphasis on trustworthiness and self-assessment capability rather than absolute performance.

D. Localization Integrity

Localization or positioning integrity can be defined as the level of trust we have that the localization estimate of a system is within an acceptable pre-defined tolerance [37]. In safety critical applications such as robotic surgery, the ability of a system to self-assess its performance and act based on that assessment is as important as its absolute performance. Substantial progress has been made in fields like satellite navigation and aviation, with growing interest for autonomous vehicles [38] [37] and generalized autonomous systems [39], but there is generally a lack of robot localization-specific integrity methods for visual-localization systems [40]–[42].

III. APPROACH

Here we present our approach to improving patient-to-robot localization performance of a vision-guided surgical robot, both through predicting the quality of target pose estimates, and active navigation. We detail our overall system architecture and the underlying pose estimation pipeline, describe the prediction factors used for target pose quality prediction, the prediction model itself, our consensus-based approach to localization performance prediction, and our method for robot trajectory selection.

A. System Overview

Our system consists of two components, localization system performance prediction and target viewpoint selection. The architecture of our system and its interaction with the underlying pose estimation pipeline is shown in Fig. 2.

B. Target Pose Estimation Pipeline

Our system builds on an underlying image feature based pose estimation pipeline which localizes targets relative to the robot. Offline, we construct a target feature model, M, consisting of image feature descriptors extracted from target reference images, and 3D coordinates in the target reference frame, $p_M^i = [x_i, y_i, z_i]$, mapped using known geometry. During online pose estimation, observed image feature descriptors and 2D image coordinates, $p_O^j = [u_j, v_j]$, are extracted from frames captured with a calibrated camera. Observed feature descriptors are matched against the

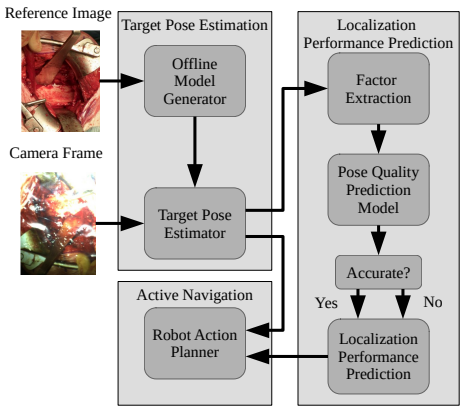

Fig. 2. Our system extracts information from an underlying pose estimation pipeline to predict the quality of individual pose estimates. We filter out pose estimates with predicted high error and use the rate of filtering at a single camera position to predict the overall localization system performance at that point which informs motion planning.

model descriptors, and the matches used to generate a list of 2D-3D correspondences. Finally, pose recovery is performed by solving the Perspective-n-Point Problem with random sampling consensus applied for outlier rejection (PnP-RANSAC).

C. Individual Pose Estimate Quality Prediction Factors

Our overall prediction system is based on the core principle of a consensus estimate based on multiple observations of the target from any particular robot position. For each of these individual estimates, we extract prediction factors from the pose estimation pipeline, for every frame where a target pose is recovered, and assign a label of *accurate* or *inaccurate* based on it being within the required tolerance.

We select the following factors based on observed correlations between each and increased target pose estimation error:

1) The number of features returned as RANSAC inliers.

$$f_1 = n \tag{1}$$

2) The spread of the inlier feature positions in the target object reference frame, calculated using a least squares linear regression on P_M^i for all inliers, finding the minimum distance from the regression fit to each feature coordinate, d_{min}, and returning the standard deviation.

$$f_2 = \sigma([d_{min}^i]; i = 1, ...n) \tag{2}$$

3) The mean re-projection error of the inlier features, calculated as the Euclidean distance between each observed feature coordinate, p_O^i, and the corresponding model feature, re-projected onto the image plane using the estimated target pose, $p_r^i = [u_r, v_r]$.

$$f_3 = \frac{1}{n}\sum_{i=1}^{n}\sqrt{(p_O^i - p_r^i)^2} \tag{3}$$

D. Individual Pose Estimate Quality Prediction Model

We construct our target pose quality prediction model as a Support Vector Machine (SVM) that outputs a binary label vector, $y \in \{0,1\}^m$ where $y = 1$ for the accurate class and $y = 0$ for the inaccurate class. The feature vector of our model is constructed from the factors described in III-C as $x^{(i)} = [f_1^{(i)}, f_2^{(i)}, f_3^{(i)}]$. Feature normalization is performed by feature clipping f_3 to the same value as the RANSAC re-projection error threshold parameter used during pose estimation. The SVM is trained with a radial basis function kernel and regularization parameter $C = 1$.

E. Consensus-Based Localization Performance Prediction

We perform localization system performance prediction at discrete camera positions using the predicted labels of individual target pose estimates. We perform target pose estimation over multiple images collected from a single camera position, predict the label for each pose and calculate $R = n_a/n_f$, the ratio of individual predicted accurate pose estimates, n_a, to frames where target pose estimation was attempted, n_f. Higher values for R indicate better predicted localization system performance.

F. Target Approach Trajectory Sampling and Selection

We propose a sample-and-predict technique to enable a robot to navigate through a target approach trajectory with low localization error, and, in cases where it is unable to maintain a required localization performance level, to identify and flag this and cease navigation.

With our technique, as shown in Fig. 3, we sample a series of pre-defined potential starting positions. We rank each position first using R values; we rank multiple positions with equal R values using the mean of the predictor model decision function calculated at each. We re-position the robot at the highest ranked starting position and begin to approach the target, regularly stopping to re-evaluate the predicted localization system performance calculating a new predicted R value. If R calculated at any location is lower than a specified threshold, we search the alternative approach directions at the current distance from the target, in order of ranking, until an acceptable position is identified or the search space is exhausted. If an acceptable approach is identified the robot re-positions and continues the target approach, otherwise it signals a localization failure and waits for user intervention.

IV. EXPERIMENTAL METHODOLOGY

Here we describe the experimental testing environment and equipment, explain specific implementation details, and describe the processes used in data collection and live navigation experiments.

A. Testing Environment

Testing and data collection was conducted with a camera equipped robot using varying combinations of pose estimation target, lighting conditions, and target obscuration. Hardware consisted of an IDS U3-3060CP-C camera rigidly

mounted to the final link of a Franka-Emika Panda robot. Software was implemented using RoboStack [43]. To evaluate the general utility of our prediction system, we applied our approach to four different pose estimation configurations, using different combinations of SIFT [44] and KAZE [45] feature detectors, and EPnP [34] and SQPnP [35] algorithms for pose recovery, implemented with OpenCV libraries [46].

Four varied actual surgical images of a human glenoid during shoulder arthroplasty procedures were selected for use as pose estimation targets [47]–[50]. To represent a subset of actual surgical challenges such as fluid obscuration and varied and harsh lighting we introduced additional perturbations including obscuring semi-opaque plastic materials and varied lighting positions from an authentic medical examination lamp. Ground truth poses of the target were determined using a grid of ArUco markers [51] placed under the target images. We calculated standard relative translation, E_t and rotation E_r pose estimation errors between the ground truth target pose and the estimated target pose within the camera reference frame.

B. Prediction and Characterization Evaluation Experiments

To collect data for model training and predictor evaluation we comprehensively sampled the robot workspace by moving the camera through a predefined 3D grid of poses representing the operational workspace of the robot in a shoulder surgery context. The grid comprised six layers of 7x7 grids. Each layer in the grid traced over a section of a sphere centered on the target as shown in Fig. 3, with the camera continually orientated towards the target. Layers in the sampling pattern were spaced between 0.35 and 0.6 meters from the localization target with 50 millimeters between each. Camera position spacing changed in each layer maintaining a 5 degree target viewing angle change between positions. The pose of the camera, relative to the robot base, and 20 image frames were recorded at each position.

Five data collection runs were completed with varying surgical targets and lighting conditions. Example images of the target (as viewed from the system camera) are shown in Fig. 1.

C. Pose Estimate Quality Prediction Model Training

We processed the robot workspace sampling data collected in Section IV-B through four configurations of the implemented pose estimation pipeline recording E_t, E_r, and the factors described in Section III-C. Each frame with a returned pose estimate is assigned a ground truth label of accurate where E_t is less than 4 millimeters and E_r less than 10 degrees, otherwise it is labelled inaccurate. These threshold values were selected as domain appropriate, being the positioning errors that define a malpositioned glenoid implant in shoulder arthroplasty [2] [10].

A training set was generated from 20% of the labelled data of one experimental configuration selected through random sampling, maintaining the ratio of class labels between the training set and complete dataset. The pose quality prediction model described in Section III-D was fit with the training set and used for pose quality prediction in all subsequent experiments - providing an explicit test of the generalization capability of the system to different lighting conditions and surgical targets.

D. Active Robot Experiments

We completed a series of active experiments using the process shown in Fig. 3. During each experiment five initial starting positions were sampled at a distance of 0.6 meters above the localization target and re-sampling was performed every 50 millimeters during target approach. The threshold for acceptable predicted localization performance set at $R = 0.9$, and pose estimate quality prediction was performed with a model pre-trained as previously described.

Following individual active trials the robot workspace was re-sampled along each potential path to provide data for comparison between selected and alternate trajectories. The localization target and lighting conditions were changed between each trial.

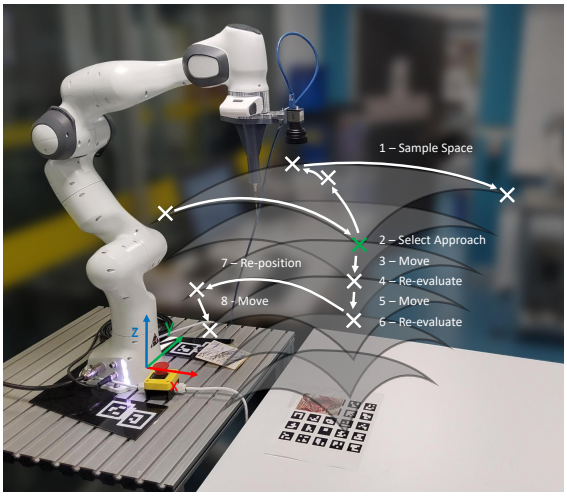

Fig. 3. During active robot navigation experiments the workspace is partly sampled (1), and an initial approach direction selected (2). The robot begins its target approach from the selected position (3), stopping regularly to re-evaluate the localization system performance (4). If performance is predicted to be acceptable the current target approach is maintained (5), otherwise (6), the workspace is again sampled until an acceptable approach is identified (7, 8) or the need for user intervention is signaled.

V. RESULTS

Here we present the results of experiments in four key areas: the effect on localization performance from the prediction system, the accuracy of the prediction system itself, characterization of the prediction accuracy over the entire robot workspace compared to ground truth, and comparison of the localization performance along navigation paths selected and rejected with the predictive system in active navigation trials.

A. Localization Improvement Through Pose Estimate Filtering

To assess the effect on the localization system of our approach, we evaluated localization performance with various

TABLE I

LOCALIZATION PERFORMANCE IMPROVEMENT THROUGH POSE ESTIMATION FILTERING.

System Configuration	Dataset	Median Rotation Error (Deg)		Median Translation Error (mm)	
		Baseline	Ours	Baseline	Ours
KAZE + EPnP	A	1.20	**0.97**	1.4	**1.1**
	B	1.49	**1.09**	2.0	**1.4**
	C	2.06	**1.72**	2.7	**2.3**
	D	2.75	**2.29**	2.7	**2.2**
	E	9.74	**3.61**	9.8	**3.5**
KAZE + SQPnP	A	1.72	**1.15**	1.5	**1.2**
	B	1.26	**0.97**	1.7	**1.4**
	C	1.89	**1.60**	2.3	**1.9**
	D	2.86	**2.23**	2.3	**1.9**
	E	11.52	**4.18**	10.3	**4.0**
SIFT + EPnP	A	1.20	**1.03**	1.4	**1.2**
	B	1.03	**0.92**	1.7	**1.5**
	C	1.66	**1.38**	2.5	**2.1**
	D	1.20	**1.09**	1.8	**1.7**
	E	4.41	**1.78**	5.4	**1.8**
SIFT + SQPnP	A	0.63	**0.52**	1.2	**1.0**
	B	0.92	**0.80**	1.6	**1.4**
	C	1.55	**1.32**	2.2	**1.8**
	D	1.03	**0.86**	1.3	**1.2**
	E	4.70	**1.72**	5.4	**1.8**

TABLE II

COMPARISON OF PRECISION SCORES FOR MATCHED RECALL IN POSE QUALITY PREDICTION.

System Configuration	Dataset	Matched Recall	Precision (WRG)	Precision (Ours)
KAZE + EPnP	A	0.88	0.84	**0.97**
	B	0.92	0.71	**0.92**
	C	0.93	0.67	**0.81**
	D	0.95	0.61	**0.71**
	E	0.8	0.27	**0.53**
KAZE + SQPnP	A	0.97	0.73	**0.86**
	B	0.92	0.77	**0.94**
	C	0.94	0.72	**0.87**
	D	0.96	0.65	**0.76**
	E	0.8	0.24	**0.48**
SIFT + EPnP	A	0.92	0.86	**0.96**
	B	0.89	0.83	**0.94**
	C	0.87	0.71	**0.85**
	D	0.87	0.78	**0.81**
	E	0.59	0.4	**0.81**
SIFT + SQPnP	A	0.92	0.88	**0.98**
	B	0.89	0.85	**0.95**
	C	0.88	0.77	**0.89**
	D	0.89	0.9	**0.97**
	E	0.64	0.42	**0.84**

TABLE III

POSE ESTIMATION ERROR AND ACCURATE FRAMES RETURNED ALONG ACCEPTED AND REJECTED PATHS

	Median Rotation Error (Deg)	Median Translation Error (mm)	Poses Within Tolerance (%)
Selected Path	**0.64**	**1.3**	**92.8**
Rejected Path	16.21	40.0	37.2
Relative Difference	96.1%	96.7%	149%

localization targets, lighting conditions, and pose estimation pipeline configurations. We calculate the median pose estimation error over the entire sampled space for each test configuration and compare the results both with and without pose estimation filtering. As shown in Table I the use of our approach to filter out predicted inaccurate pose estimates resulted in significant reductions in median localization error, with all but one dataset brought to within acceptable limits. Across all datasets our method resulted in a relative 26.8% mean reduction in rotation error and 25.9% mean reduction in translation error. The performance improvements provided by our system were significant across both variations in scene and underlying pose estimation system configuration, showing the potential for our system to generalize between target textures and lighting conditions.

B. Pose Quality Prediction

While localization performance is improved, we provide further analysis here of the performance of the prediction system itself, using the same comprehensive workspace sampling data as Section V-A. Due to the sparsity of quantitative benchmarks in this narrow field our comparison baseline is a weighted naive system tuned to predict accurate / inaccurate labels for each pose estimate at the same rate as our predictive system, referred to as Weighted Random Guesser (WRG). Using a parameter sweep, we equalize the recall scores of the baseline with our system to enable direct comparison in Table II. Our prediction model outperforms the WRG across 100% of the tested datasets with a relative mean increase in precision of 32.6% at matched recall.

C. Robot Workspace Characterization

To move beyond summary statistics, we also characterize the localization prediction performance over the workspace of the robot, once again with the workspace exploration data from Section V-A. We generate heatmaps of the calculated R values across the workspace for each dataset and pose estimation configuration. Results for Dataset A, B, and C are shown in Fig. 4.

D. Active Localization Evaluation

Here we present results from a series of active experiments as described in Section IV-D. The automatically selected target approach trajectory and R_{gt} values for the sampled positions in the robot workspace are shown in Fig. 5. We compare the median pose estimation error and percentage of pose estimates within tolerance along the traversed trajectory and the trajectories rejected by the system based on initial sampling in Table III. Trials where no initial path was directly rejected (Trial 3), or no pose estimates were generated by the underlying pipeline while sampling the rejected path (Trial 2) were excluded from these calculations.

Across all trials our system consistently selected approach trajectories that returned high rates of accurate individual pose estimates, maintained localization accuracy within our specified tolerance and below that of the rejected trajectories. In terms of specific examples of system functionality, during Trial 3 our system correctly identified the need to transition between target approach angles to maintain performance, while in Trial 5 our system correctly identified that localization performance of all possible routes was below acceptable limits, flagged this outcome, and ceased operation.

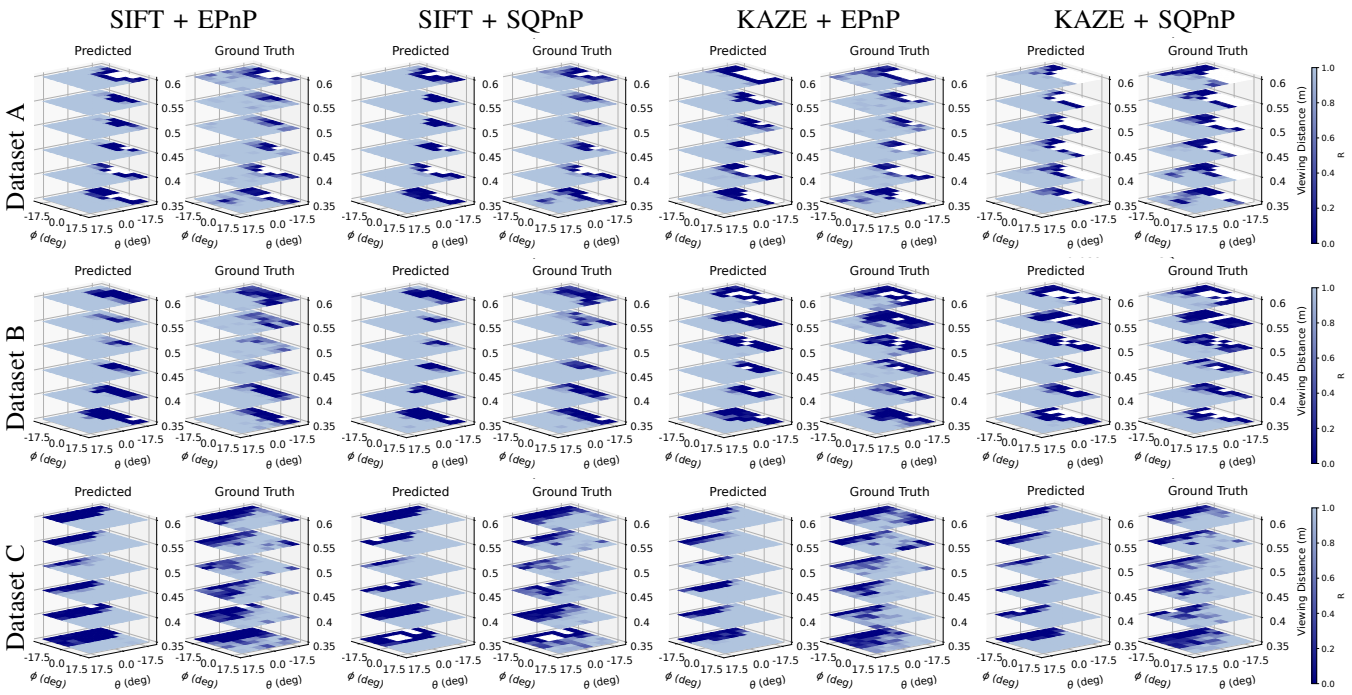

Fig. 4. The workspace heatmaps of R across all tests show close alignment between predicted and ground truth values. Camera angle θ is rotation about the x-axis of the robot base, camera angle ϕ is rotation about y-axis of the robot, both are measured relative to the camera oriented vertically downwards. Positions where the pose estimation pipeline failed to return any pose estimate across all sampled frames are indicated with white squares.

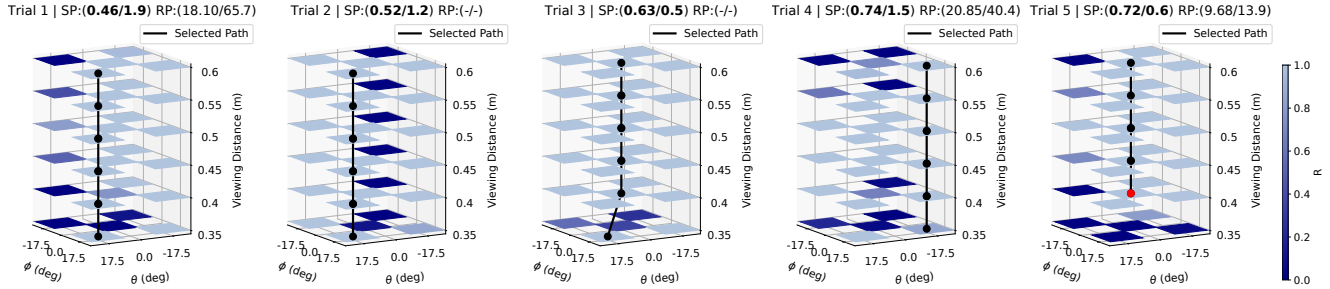

Fig. 5. Title format is Trial #, Selected Path (SP) & Rejected Path(s), metrics are Rotation Error (degrees)/ Translation Error (mm). Camera paths selected using our method consistently result in low localization error trajectories. Our method correctly identifies the need for approach trajectory deviation in Trial 3 and that the localization system is unable to provide acceptable performance in Trial 5.

VI. CONCLUSIONS

We have presented an approach to improving localization performance of a vision-guided surgical robot by introducing a new prediction-based approach that enables prediction of the likely localization performance throughout the robot's operating workspace, which not only directly improves localization precision-recall performance, but also facilitates active navigation to minimize localization issues. While not yet achieving the sub millimeter and sub degree accuracy generally targeted by CAOS systems this approach shows promising performance improvements over the baseline localization system.

Future work will expand our approach into four key areas of opportunity. In this work we build our system on a local image feature based pose estimation pipeline, but future work could expand our approach to point cloud-based target pose estimation methods, which are increasingly popular in surgery specific pose estimation [12]–[16]. The active navigation system used in our method relied on discrete sampling, already producing good results, but future work could extend this to a more continuous sampling approach, or variable discreteness sampling. While the results here demonstrate good generalization capability from parameters trained on a single operating configuration, future work could explore online fine tuning of parameters in a semi-self-supervised manner, providing improved performance specific to each individual operating scenario. Finally, where our results show improvement in localization performance metrics future work should assess performance in terms of an application specific task e.g transitioning experiments to anatomic phantoms and measuring the final positioning error of a robot inserted guide-wire.

REFERENCES

[1] T. M. Gregory, A. Sankey, B. Augereau, E. Vandenbussche, A. Amis, R. Emery, and U. Hansen, "Accuracy of Glenoid Component Placement in Total Shoulder Arthroplasty and Its Effect on Clinical and Radiological Outcome in a Retrospective, Longitudinal, Monocentric Open Study," *PLoS ONE*, vol. 8, no. 10, p. e75791, Oct 2013.

[2] G. Villatte, A.-S. Muller, B. Pereira, A. Mulliez, P. Reilly, and R. Emery, "Use of Patient-Specific Instrumentation (PSI) for glenoid component positioning in shoulder arthroplasty. A systematic review and meta-analysis," *PLOS ONE*, vol. 13, no. 8, p. e0201759, Aug 2018.

[3] A. F. Chen, G. S. Kazarian, G. W. Jessop, and A. Makhdom, "Robotic Technology in Orthopaedic Surgery," *Journal of Bone and Joint Surgery*, vol. 100, no. 22, pp. 1984–1992, Nov 2018.

[4] G. Porcellini, L. Tarallo, M. Novi, F. Spiezia, and F. Catani, "Technology applications in shoulder replacement," *Journal of Orthopaedics and Traumatology*, vol. 20, no. 1, pp. 20–21, 2019.

[5] A. Mata-Fink and J.-E. Bell, "Indications and Preoperative Evaluation for Anatomic Shoulder Arthroplasty," in *Anatomic Shoulder Arthroplasty: Strategies for Clinical Management*, A. D. Armstrong and A. M. Murthi, Eds. Cham: Springer, 2016, pp. 1–13.

[6] G. Fama and A. Pozzuoli, "History of Reverse Shoulder Arthroplasty," in *Reverse Shoulder Arthroplasty: Current Techniques and Complications*, S. Gumina, F. A. Grassi, and P. Paladini, Eds. Cham: Springer, 2019, pp. 3–23.

[7] E. R. Wagner, K. X. Farley, I. Higgins, J. M. Wilson, C. A. Daly, and M. B. Gottschalk, "The incidence of shoulder arthroplasty: rise and future projections compared with hip and knee arthroplasty," *Journal of Shoulder and Elbow Surgery*, vol. 29, no. 12, pp. 2601–2609, Dec 2020.

[8] J. Levy, "Standard Glenoid Replacement," in *Anatomic Shoulder Arthroplasty: Strategies for Clinical Management*, A. D. Armstrong and A. M. Murthi, Eds. Cham: Springer, 2016, pp. 97–110.

[9] A. P. Dekker and A. A. Tambe, "Navigation in Shoulder Arthroplasty," *Journal of Arthroscopy and Joint Surgery*, vol. 8, no. 1, pp. 35–43, Jan 2021.

[10] B. S. Schoch, E. Haupt, T. Leonor, K. W. Farmer, T. W. Wright, and J. J. King, "Computer navigation leads to more accurate glenoid targeting during total shoulder arthroplasty compared with 3-dimensional preoperative planning alone," *Journal of Shoulder and Elbow Surgery*, vol. 29, no. 11, pp. 2257–2263, Nov 2020.

[11] F. Picard, A. H. Deakin, P. E. Riches, K. Deep, and J. Baines, "Computer assisted orthopaedic surgery: Past, present and future," *Medical engineering & physics*, vol. 72, pp. 55–65, Oct 2019.

[12] P. Rodrigues, M. Antunes, C. Raposo, P. Marques, F. Fonseca, and J. P. Barreto, "Deep segmentation leverages geometric pose estimation in computer-aided total knee arthroplasty," *Healthcare Technology Letters*, vol. 6, no. 6, pp. 226–230, Dec 2019.

[13] I. Félix, C. Raposo, M. Antunes, P. Rodrigues, and J. P. Barreto, "Towards markerless computer-aided surgery combining deep segmentation and geometric pose estimation: application in total knee arthroplasty," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 9, no. 3, pp. 271–278, 2021.

[14] H. Liu and F. R. Y. Baena, "Automatic markerless registration and tracking of the bone for computer-assisted orthopaedic surgery," *IEEE Access*, vol. 8, pp. 42 010–42 020, 2020.

[15] X. Hu, H. Liu, and F. R. Y. Baena, "Markerless navigation system for orthopaedic knee surgery: A proof of concept study," *IEEE Access*, vol. 9, pp. 64 708–64 718, 2021.

[16] X. Hu, A. Nguyen, and F. R. y. Baena, "Occlusion-robust visual markerless bone tracking for computer-assisted orthopedic surgery," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–11, 2022.

[17] C. Sahin, G. Garcia-Hernando, J. Sock, and T.-K. Kim, "A review on object pose recovery: From 3D bounding box detectors to full 6D pose estimators," *Image and Vision Computing*, vol. 96, p. 103898, Apr 2020.

[18] B. Li, W. Ouyang, L. Sheng, X. Zeng, and X. Wang, "Gs3d: An efficient 3d object detection framework for autonomous driving," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019.

[19] Y. Wang, W.-L. Chao, D. Garg, B. Hariharan, M. Campbell, and K. Q. Weinberger, "Pseudo-lidar from visual depth estimation: Bridging the gap in 3d object detection for autonomous driving," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 8445–8453.

[20] B. Xu and Z. Chen, "Multi-level fusion based 3d object detection from monocular images," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 2345–2353.

[21] S. Hinterstoisser, V. Lepetit, S. Ilic, P. Fua, and N. Navab, "Dominant orientation templates for real-time detection of texture-less objects," in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2010, pp. 2257–2264.

[22] E. Muñoz, Y. Konishi, V. Murino, and A. Del Bue, "Fast 6d pose estimation for texture-less objects from a single rgb image," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2016, pp. 5623–5630.

[23] Y. Xiang, T. Schmidt, V. Narayanan, and D. Fox, "Posecnn: A convolutional neural network for 6d object pose estimation in cluttered scenes," *arXiv preprint arXiv:1711.00199*, 2017.

[24] W. Kehl, F. Manhardt, F. Tombari, S. Ilic, and N. Navab, "Ssd-6d: Making rgb-based 3d detection and 6d pose estimation great again," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, Oct 2017.

[25] K. Park, T. Patten, and M. Vincze, "Pix2pose: Pixel-wise coordinate regression of objects for 6d pose estimation," in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, October 2019.

[26] R. König and B. Drost, "A hybrid approach for 6dof pose estimation," in *European Conference on Computer Vision (ECCV)*. Springer, 2020, pp. 700–706.

[27] S. Peng, Y. Liu, Q. Huang, X. Zhou, and H. Bao, "Pvnet: Pixel-wise voting network for 6dof pose estimation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 4561–4570.

[28] Z. Li, G. Wang, and X. Ji, "Cdpn: Coordinates-based disentangled pose network for real-time rgb-based 6-dof object pose estimation," in *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019, pp. 7677–7686.

[29] T. Hodan, D. Barath, and J. Matas, "Epos: Estimating 6d pose of objects with symmetries," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 11 703–11 712.

[30] S. Shi, C. Guo, L. Jiang, Z. Wang, J. Shi, X. Wang, and H. Li, "Pv-rcnn: Point-voxel feature set abstraction for 3d object detection," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 10 529–10 538.

[31] A. S. Mian, M. Bennamoun, and R. A. Owens, "Automatic correspondence for 3D modeling: an extensive review," *International Journal of Shape Modeling*, vol. 11, no. 02, pp. 253–291, Dec 2005.

[32] B. Drost, M. Ulrich, N. Navab, and S. Ilic, "Model globally, match locally: Efficient and robust 3d object recognition," in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2010, pp. 998–1005.

[33] J. Vidal, C.-Y. Lin, and R. Marti, "6D pose estimation using an improved method based on point pair features," in *2018 4th International Conference on Control, Automation and Robotics (ICCAR)*. IEEE, Apr 2018, pp. 405–409.

[34] V. Lepetit, F. Moreno-Noguer, and P. Fua, "EPnP: An accurate O(n) solution to the PnP problem," *International Journal of Computer Vision*, vol. 81, no. 2, pp. 155–166, Feb 2009.

[35] G. Terzakis and M. Lourakis, "A consistently fast and globally optimal solution to the perspective-n-point problem," in *European Conference on Computer Vision (ECCV)*. Springer, 2020, pp. 478–494.

[36] J. Cartucho, S. Tukra, Y. Li, D. S. Elson, and S. Giannarou, "Visionblender: a tool to efficiently generate computer vision datasets for robotic surgery," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 9, no. 4, pp. 331–338, 2021.

[37] J. Al Hage, P. Xu, P. Bonnifait, and J. Ibanez-Guzman, "Localization Integrity for Intelligent Vehicles Through Fault Detection and Position Error Characterization," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 4, pp. 2978–2990, Apr 2022.

[38] G. D. Arana, O. A. Hafez, M. Joerger, and M. Spenko, "Localization safety validation for autonomous robots," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 6276–6281.

[39] A. Gautam, T. Whiting, X. Cao, M. A. Goodrich, and J. W. Crandall, "A Method for Designing Autonomous Robots that Know Their

Limits," in *2022 International Conference on Robotics and Automation (ICRA)*. IEEE, May 2022, pp. 121–127.
[40] C. Zhu, M. Meurer, and C. Günther, "Integrity of visual navigation—developments, challenges, and prospects," *NAVIGATION: Journal of the Institute of Navigation*, vol. 69, no. 2, 2022.
[41] Y. D. Yasuda, L. E. G. Martins, and F. A. Cappabianco, "Autonomous visual navigation for mobile robots: A systematic literature review," *ACM Computing Surveys (CSUR)*, vol. 53, no. 1, pp. 1–34, 2020.
[42] H. Carson, J. J. Ford, and M. Milford, "Predicting to Improve: Integrity Measures for Assessing Visual Localization Performance," *IEEE Robotics and Automation Letters*, pp. 1–8, 2022.
[43] T. Fischer, W. Vollprecht, S. Traversaro, S. Yen, C. Herrero, and M. Milford, "A robostack tutorial: Using the robot operating system alongside the conda and jupyter data science ecosystems," *IEEE Robotics and Automation Magazine*, 2021.
[44] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International journal of computer vision*, vol. 60, no. 2, pp. 91–110, 2004.
[45] P. F. Alcantarilla, A. Bartoli, and A. J. Davison, "Kaze features," in *European Conference on Computer Vision (ECCV)*, 2012, pp. 214–227.
[46] G. Bradski, "The OpenCV Library," *Dr. Dobb's Journal of Software Tools*, 2000.
[47] D. Oudet, "Reverse Shoulder Arthroplasty for Massive Rotator Cuff Tears," *International Journal of Orthopaedics*, vol. 6, no. 1, pp. 1016–1031, 2019.
[48] B. S. Schoch, R. U. Hartzler, and J. W. Sperling, "Procedure 9 - Total Shoulder Arthroplasty," in *Operative Techniques: Shoulder and Elbow Surgery*, 2nd ed., D. H. Lee and R. J. Neviaser, Eds. Philadelphia, PA: Elsevier, 2019, pp. 102–112.
[49] S. S. Goldberg, J.-E. Bell, H. J. Kim, S. F. Bak, W. N. Levine, and L. U. Bigliani, "Hemiarthroplasty for the rotator cuff-deficient shoulder," *Journal of Bone and Joint Surgery*, vol. 90, no. 3, pp. 554–559, 2008.
[50] M. Crowther, "Total shoulder replacement: Lima smr reverse shoulder for proximal humeral fracture surgical technique," Oct 2021, Accessed: Aug 8, 2022. [Online]. Available: https://www.orthoracle.com/library/lima-smr-reverse-total-shoulder-replacement-for-proximal-humeral-fracture/
[51] S. Garrido-Jurado, R. Muñoz-Salinas, F. Madrid-Cuevas, and M. Marín-Jiménez, "Automatic generation and detection of highly reliable fiducial markers under occlusion," *Pattern Recognition*, vol. 47, no. 6, pp. 2280–2292, 2014.